NCGC National Clinical Guideline Centre

Delirium: diagnosis, prevention and management

Clinical Guideline 103

July 2010

Commissioned by the National Institute for Health and Clinical Excellence

Published by the National Clinical Guideline Centre (formerly the National Collaborating Centre for Nursing and Supportive Care) at The Royal College of Physicians, 11 St Andrews Place, Regent's Park, London, NW1 4BT
First published 2010

Update information

January 2023: We have reviewed the evidence and made new recommendations on assessing and diagnosing delirium. These recommendations are marked [2023].

March 2019: Olanzapine has been removed from recommendation 1.6.4 because the clinical need can now be met by a licensed product. The footnote to this recommendation stated that haloperidol and olanzapine do not have UK marketing authorisation for delirium treatment. However, haloperidol does now have marketing authorisation. Therefore, the footnote has been removed because it no longer applies to haloperidol. Recommendation 1.5.1 has also been amended because a delirium diagnosis is now confirmed using DSM-V.

For the 2023 and 2019 updates, see the amended recommendations at: www.nice.org.uk/guidance/cg103

This full guideline preserves evidence reviews and committee discussions from the original 2010 version.

CONTENTS

APPENDICES (in separate files):

APPENDIX A: Scope

APPENDIX B: Declarations of Interest

APPENDIX C: Clinical questions and literature search

APPENDIX D: Included studies

APPENDIX E: Methodological quality

APPENDIX F: Multivariate risk factors

APPENDIX G: Excluded studies (part A) and excluded studies reference list (part B)

APPENDIX H: Recommendations for future research

APPENDIX I: DSM-IV and scales for measuring delirium

APPENDIX J: Health economics appendix

APPENDIX K: Forest Plots

Guideline development group

Prof John Young (Chair)	Honorary Consultant Geriatrician, Bradford Teaching Hospitals Foundation NHS Trust.
Dr David Anderson	Consultant in Old Age Psychiatry, Mersey Care NHS Trust.
Ms Melanie Gager	Sister in Critical Care Follow Up, Royal Berkshire Hospital, Reading.
Dr Jim George	Consultant Physician, Cumberland Infirmary, Carlisle.
Ms Jane Healy	Senior Clinical Practice Facilitator, UCLH NHS Foundation Trust, London.
Ms Wendy Harvey (nee Tomlinson)	Home Manager, MHA Care Group
Dr Anne Hicks	Consultant in Emergency Medicine, Plymouth Hospitals NHS Trust.
Dr John Holmes	Senior Lecturer - Liaison Psychiatry of Old Age, Institute of Health Sciences, University of Leeds.
Ms Emma Ouldred	Dementia Nurse Specialist, King's College Hospital NHS Foundation Trust, London.
Dr Najma Siddiqi	Consultant Psychiatrist, Bradford District Care Trust, West Yorkshire.
Mr Gordon Sturmey	Patient Member, Critpal (Intensive Care Society) (until August 2008)
Ms Beverley Tabernacle	Nurse Consultant, Salford Royal Foundation Trust (until January 2009)
Ms Rachel White	Patient Member
Mr Matt Wiltshire	Patient member (from November 2008)
Dr Andrew Clegg (non-voting member)	SpR in Geriatric and General Medicine, Bradford Royal Infirmary, West Yorkshire.
Dr Anayo Akunne	NCGC Health Economist
Dr Ian Bullock (voting member)	NCGC Chief Operating Officer
Ms Sarah Davis (voting member)	NCGC Senior Health Economist (until December 2009)
Dr Bernard Higgins	NCGC Clinical Director
Mr Paul Miller	NCGC Senior Information Specialist
Ms Lakshmi Murthy	NCGC Research Fellow
Dr Rachel O'Mahony	NCGC Senior Research Fellow (delirium guideline project manager August 2009 – March 2010)
Ms Jill Parnham	NCGC Operations Director
Dr Silvia Rabar	NCGC Project Manager (from March 2010)
Dr Fulvia Ronchi	NCGC Senior Project Manager (April - August 2009)
Dr Maggie Westby (voting member)	NCGC Clinical Effectiveness Lead

Acknowledgements

The development of this guideline was greatly assisted by the following people:

Ms Taryn Krause (NCGC Senior Project Manager); Dr Kathy DeMott (NCGC Research Fellow); Ms Nahara Martinez (NCGC Research Fellow); Dr Sharon Swain (NCGC Senior Research Fellow); Ms Tamara Diaz, Ms Abigail Jones, Ms Fatema Limbada, Ms Emma Nawrocki and Ms Jaymeeni Solanki (NCGC Project Co-ordinators); Dr Sarah King, Dr Yoon Loke, Dr Jacoby Patterson, and Mr John Rathbone (external reviewers); Ms Christine Sealey (NICE Commissioning Manager until September 2008); Ms Susan Latchem (NICE Commissioning Manager from September 2008 until January 2010); Ms Sarah Dunsdon (NICE Commissioning Manager from January 2010); Ms Emilene Coventry (NICE Senior Editor).

Abbreviations and acronyms

ADL	Activities of Daily Living
AGU	Acute Geriatric Unit
AMT	Abbreviated Mental Test
ANOVA	Analysis of variance
APACHE	Acute Physiology and Chronic Health Evaluation (severity of illness classification system)
ARDS	Acute respiratory distress syndrome
ASA	American Society of Anesthesiologists (score for illness severity)
ASE	Attention Screening Examination
BEHAVE-AD	Behavioural Pathology in Alzheimer's Disease Rating
BNF	British National Formulary
CABG	Coronary artery bypass grafting
CAM	Confusion Assessment Method
CAM-ICU	Confusion Assessment Method for the ICU
CCA	Cost-consequences analysis
CD	Compact disc
CDR	Clinical Dementia Rating scale
CDT	Clock Drawing Test
CEA	Cost-effectiveness analysis
c.f.	Confer (refer to)
CGBRS	Crichton Geriatric Behavioural Rating Scale
CGI	Clinical global impression scale
CGI-GI	Clinical global impression scale: global improvement item
CGI-SI	Clinical global impression scale: severity of illness item
CHF	Chronic heart failure

CI / 95% CI	Confidence interval / 95% confidence interval
CIPFA	Chartered Institute of Public Finance and Accountancy
CNS	Central nervous system
COPD	Chronic obstructive pulmonary disease
CT	Computed tomography
CUA	Cost-utility analysis
DH	Department of Health
DI	Delirium Index
DRS / DRS-98 or DRS-R-98	Delirium Rating Scale / DRS-revised-98
DSA	Deterministic Sensitivity Analysis
DSI	Delirium Symptom Interview
DSM (DSM III, III-R or IV)	Diagnostic and Statistical Manual of Mental Disorders (edition III, III-R or IV)
ED	Emergency Department
EQ-5D	EuroQol-5D
FCEs	Finished Consultant Episodes
FIM	Functional Independence Measure
GA	General anaesthesia
GDG	Guideline Development Group
GI	Gastrointestinal
GP	General Practitioner
GRADE	Grading of Recommendations Assessment, Development and Evaluation
HES	Hospital Episode Statistics
HR	Hazard Ratio
HRQoL	Health-related quality of life
HT / 5-HT / 5-HT3	5-hydroxytryptamine / 5-hydroxytryptamine 3

HTA	Health technology assessment
Hx	History (in appendices)
ICD-10	International Classification of Diseases, 10th edition
ICU-DSC	Intensive Care Unit- Delirium Screening Checklist
ICER	Incremental cost-effectiveness ratio
ICU	Intensive Care Unit
IQR	Interquartile range
INMB	Incremental Net Monetary Benefit
IQCODE	Informant Questionnaire on Cognitive Decline in the Elderly
IRR	Inter-rater reliability
K	Cohen's kappa
ITT	Intention to treat
LOS	Length of Stay
LR^+	Positive likelihood ratio
LTC	Long-term care
LY	Life-year
MD	Mean difference
MDAS	Memorial Delirium Assessment Scale
MDC	Major diagnostic category
MI	Myocardial infraction
MMSE	Mini-Mental State Examination
MRI	Magnetic resonance imaging
MTI	Multicomponent Targeted Interventions
NCGC	National Clinical Guidelines Centre
NH	Nursing Home
NHS	National Health Service
NHSEED	The NHS Economic Evaluation Database

NICE	National Institute for Health and Clinical Excellence
NINDS-AIREN	National Institute of Neurological Disorders and Stroke and Association Internationale pour la Recherché et l'Enseignement en Neurosciences
NNT	Number needed to treat
NPV	Negative predictive value
NSAID	Non-steroidal anti-inflammatory drug
OBS	Organic Brain Syndrome
OECD	Organisation for Economic Co-operation and Development
OR	Odds ratio
PASA	NHS Purchasing and Supply Agency
PCA	Patient controlled analgesia
PICO	Framework incorporating patients, interventions, comparison and outcome
POPS	Proactive care of older people undergoing surgery
PPP	Purchasing Power Parity
PPV	Positive predictive value
p.r.n	Pro re nata
PSA	Probabilistic sensitivity analysis
PSS	Personal Social Services
PSSRU	Personal Social Services Research Unit
QALY	Quality-adjusted life year
QUADAS	Quality assessment tool for diagnostic accuracy studies
RASS	Richmond Agitation Sedation Scale
RCT	Randomised controlled trial
RFs	Risk factors
ROC	Receiver operating characteristic
RR	Relative risk
SD	Standard deviation

SDC	Saskatoon Delirium Checklist
SE	Standard error
SICU	Surgical Intensive Care Unit
SPC	Summary of product characteristics
SPMSQ	Short Portable Mental Status Questionnaire
SR	Systematic review
TICS	Telephone interview for cognitive status
VAS	Visual analogue scale

Glossary of Terms

Abstract	Summary of a study, which may be published alone or as an introduction to a full scientific paper.
Acute confusional state	A synonymous term for delirium.
Algorithm (in guidelines)	A flow chart of the clinical decision pathway described in the guideline, where decision points are represented with boxes, linked with arrows.
Allocation concealment	The process used to prevent advance knowledge of group assignment in a RCT. The allocation process should be impervious to any influence by the individual making the allocation, by being administered by someone who is not responsible for recruiting participants.
AMT (Abbreviated Mental Test)	A quick and easy to use screening test to detect cognitive impairment.
Anticholinergic	A group of drugs which inhibit the transmission of parasympathetic nerve impulses and inhibit the brain neurotransmitter acetylcholine.
Antipsychotic	Also known as neuroleptic drugs, these are a class of psychoactive drugs.
Applicability	The degree to which the results of an observation, study or review are likely to hold true in a particular clinical practice setting.
Arm (of a clinical study)	Sub-section of individuals within a study who receive one particular intervention, for example placebo arm.
Association	Statistical relationship between two or more events, characteristics or other variables. The relationship may or may not be causal.
Atypical antipsychotic	These are the second-generation antipsychotics. They are chemically different from and have different side effects than the older ‘typical’ antipsychotic medications.
Baseline	The initial set of measurements at the beginning of a study (after run-in period where applicable), with which subsequent results are compared.
Before-and-after study	A study that investigates the effects of an intervention by measuring particular characteristics of a population both before and after taking the intervention, and assessing any change that occurs.

Bias	Systematic (as opposed to random) deviation of the results of a study from the 'true' results that is caused by the way the study is designed or conducted.
Blinding	Keeping the study participants, caregivers, researchers and outcome assessors unaware about the interventions to which the participants have been allocated in a study.
Cardio-aspirin	Lower dose treatment with aspirin to reduce the occurrence of vascular disease.
Carer (caregiver)	Someone other than a health professional who is involved in caring for a person with a medical condition.
Case-control study	Comparative observational study in which the investigator selects individuals who have experienced an event (For example, developed a disease) and others who have not (controls), and then collects data to determine previous exposure to a possible cause.
Case-series	Report of a number of cases of a given disease, usually covering the course of the disease and the response to treatment. There is no comparison (control) group of patients.
Clinical efficacy	The extent to which an intervention is active when studied under controlled research conditions.
Clinical effectiveness	The extent to which an intervention produces an overall health benefit in routine clinical practice.
Clinical question	In guideline development, this term refers to the questions about treatment and care that are formulated to guide the development of evidence-based recommendations.
Clinician	A healthcare professional providing direct patient care, for example doctor, nurse or physiotherapist.
Cochrane Review	The Cochrane Library consists of a regularly updated collection of evidence-based medicine databases including the Cochrane Database of Systematic Reviews (reviews of randomised controlled trials prepared by the Cochrane Collaboration).
Cognitive impairment	Difficulty with memory, thinking, concentration and ability to read and write.
Cohort study	A retrospective or prospective follow-up study. Groups of individuals to be followed up are defined on the basis of presence or absence of exposure to a suspected risk factor or intervention. A cohort study can be comparative, in which case two or more groups are selected on the basis of differences in their exposure to the agent of interest.
Comorbidity	Co-existence of more than one disease or an additional disease

	(other than that being studied or treated) in an individual.
Comparability	Similarity of the groups in characteristics likely to affect the study results (such as health status or age).
Concordance	This is a recent term whose meaning has changed. It was initially applied to the consultation process in which doctor and patient agree therapeutic decisions that incorporate their respective views, but now includes patient support in medicine taking as well as prescribing communication. Concordance reflects social values but does not address medicine-taking and may not lead to improved adherence.
Confidence interval (CI)	A range of values for an unknown population parameter with a stated 'confidence' (conventionally 95%) that it contains the true value. The interval is calculated from sample data, and generally straddles the sample estimate. The 'confidence' value means that if the method used to calculate the interval is repeated many times, then that proportion of intervals will actually contain the true value.
Confounding	In a study, confounding occurs when the effect of an intervention on an outcome is distorted as a result of an association between the population or intervention or outcome and another factor (the 'confounding variable') that can influence the outcome independently of the intervention under study.
Confusion Assessment Method (CAM)	An assessment tool that has been validated to help detect delirium that is carried out by means of a clinical interview.
Control group	A group of patients recruited into a study that receives no treatment, a treatment of known effect, or a placebo (dummy treatment) - in order to provide a comparison for a group receiving an experimental treatment, such as a new drug.
Cost benefit analysis	A type of economic evaluation where both costs and benefits of healthcare treatment are measured in the same monetary units. If benefits exceed costs, the evaluation would recommend providing the treatment.
Cost-consequences analysis (CCA)	A type of economic evaluation where various health outcomes are reported in addition to cost for each intervention, but there is no overall measure of health gain.
Cost-effectiveness analysis (CEA)	An economic study design in which consequences of different interventions are measured using a single outcome, usually in 'natural' units (For example, life-years gained, deaths avoided, heart attacks avoided, cases detected). Alternative interventions are then compared in terms of cost per unit of effectiveness.

Cost-effectiveness model	An explicit mathematical framework, which is used to represent clinical decision problems and incorporate evidence from a variety of sources in order to estimate the costs and health outcomes.
Cost-utility analysis (CUA)	A form of cost-effectiveness analysis in which the units of effectiveness are quality-adjusted life-years (QALYs).
Data synthesis	A generic term to describe methods used for summarising (comparing and contrasting) evidence into a clinically meaningful conclusion in order to answer a defined clinical question. This can include systematic review (with or without meta-analysis), other quantitative methods or qualitative and narrative summaries.
Decision analysis	An explicit quantitative approach to decision making under uncertainty, based on evidence from research. This evidence is translated into probabilities, and then into diagrams or decision trees which direct the clinician through a succession of possible scenarios, actions and outcomes.
Decision problem	A clear specification of the interventions, patient populations and outcome measures and perspective adopted in an evaluation, with an explicit justification, relating these to the decision which the analysis is to inform.
Discounting	Costs and perhaps benefits incurred today have a higher value than costs and benefits occurring in the future. Discounting health benefits reflects individual preference for benefits to be experienced in the present rather than the future. Discounting costs reflects individual preference for costs to be experienced in the future rather than the present.
Dominance	An intervention is said to be dominated if there is an alternative intervention that is both less costly and more effective.
Dosage	The prescribed amount of a drug to be taken, including the size and timing of the doses.
DSM III, III-R or IV	Diagnostic and Statistical Manual of Mental Disorders (edition III, III-R or IV). Diagnostic test used to diagnose delirium.
Economic evaluation	Comparative analysis of alternative health strategies (interventions or programmes) in terms of both their costs and consequences.
Effect (as in effect measure, treatment effect, estimate of effect, effect size)	The observed association between interventions and outcomes or a statistic to summarise the strength of the observed association.
Effectiveness	See ‘Clinical effectiveness’.

Efficacy	See 'Clinical efficacy'.
End of life care	People in the last few days of their life
Epidemiological study	The study of a disease within a population, defining its incidence and prevalence and examining the roles of external influences (For example, infection, diet) and interventions.
EQ-5D (EuroQol-5D)	A standardised instrument used to measure a health outcome. It provides a single index value for health status.
Evidence	Information on which a decision or guidance is based. Evidence is obtained from a range of sources including randomised controlled trials, observational studies, expert opinion (of clinical professionals and/or patients).
Exclusion criteria (literature review)	Explicit standards used to decide which studies should be excluded from consideration as potential sources of evidence.
Exclusion criteria (clinical study)	Criteria that define who is not eligible to participate in a clinical study.
Extended dominance	If Option A is both more clinically effective than Option B and has a lower cost per unit of effect, when both are compared with a do-nothing alternative then Option A is said to have extended dominance over Option B. Option A is therefore more efficient and should be preferred, other things remaining equal.
Extrapolation	In data analysis, predicting the value of a parameter outside the range of observed values.
Extrapyramidal	Pertaining to the tissues and structures outside the cerebrospinal pyramidal tracts of the brain that are associated with movement of the body, excluding motor neurons, the motor cortex, and the corticospinal and corticobulbar tracts.
Follow-up	Observation over a period of time of an individual, group or initially defined population whose appropriate characteristics have been assessed in order to observe changes in health status or health-related variables.
Generalisability	The extent to which the results of a study based on measurement in a particular patient population and/or a specific context hold true for another population and/or in a different context. In this instance, this is the degree to which the guideline recommendation is applicable across both geographical and contextual settings. For instance, guidelines that suggest substituting one form of labour for another should acknowledge that these costs might vary across the country.
Gold standard	See 'Reference standard'.

GRADE / GRADE profile	A system developed by the GRADE Working Group to address the shortcomings of present grading systems in healthcare. The GRADE system uses a common, sensible and transparent approach to grading the quality of evidence. The results of applying the GRADE system to clinical trial data are displayed in a table known as a GRADE profile.
Harms	Adverse effects of an intervention.
Health economics	The study of the allocation of scarce resources among alternative healthcare treatments. Health economists are concerned with both increasing the average level of health in the population and improving the distribution of health.
Health-related quality of life (HRQoL)	A combination of an individual's physical, mental and social well-being; not merely the absence of disease.
Heterogeneity	Or lack of homogeneity. The term is used in meta-analyses and systematic reviews when the results or estimates of effects of treatment from separate studies seem to be very different – in terms of the size of treatment effects or even to the extent that some indicate beneficial and others suggest adverse treatment effects. Such results may occur as a result of differences between studies in terms of the patient populations, outcome measures, definition of variables or duration of follow-up.
Hyperactive delirium	Subtype of delirium characterised by people who have heightened arousal and can be restless, agitated or aggressive.
Hypoactive delirium	Subtype of delirium characterised by people who become withdrawn, quiet and sleepy.
Hypothesis	A supposition made as a starting point for further investigation.
Imprecision	Results are imprecise when studies include relatively few patients and few events and thus have wide confidence intervals around the estimate of effect.
Incident delirium	Newly occurring case(s) of delirium
Inclusion criteria (literature review)	Explicit criteria used to decide which studies should be considered as potential sources of evidence.
Incremental analysis	The analysis of additional costs and additional clinical outcomes with different interventions.
Incremental cost	The mean cost per patient associated with an intervention minus the mean cost per patient associated with a comparator intervention.

Incremental cost effectiveness ratio (ICER)
The difference in the mean costs in the population of interest divided by the differences in the mean outcomes in the population of interest for one treatment compared with another.

$$ICER = \frac{(Cost_A - Cost_B)}{(Effectiveness_A - Effectiveness_B)}$$

Incremental net benefit (INB)
The value (usually in monetary terms) of an intervention net of its cost compared with a comparator intervention. The INB can be calculated for a given cost-effectiveness (willingness to pay) threshold. If the threshold is £20,000 per QALY gained then the INB is calculated as: (£20,000 x QALYs gained) – Incremental cost.

Index
In epidemiology and related sciences, this word usually means a rating scale, for example, a set of numbers derived from a series of observations of specified variables. Examples include the various health status indices, and scoring systems for severity or stage of cancer.

Indirectness
The available evidence is different to the clinical question being addressed, in terms of PICO (population, intervention, comparison and outcome).

Intention to treat analysis (ITT)
A strategy for analysing data from a randomised controlled trial. All participants are included in the arm to which they were allocated, whether or not they received (or completed) the intervention given to that arm. Intention-to-treat analysis prevents bias caused by the loss of participants, which may disrupt the baseline equivalence established by randomisation and which may reflect non-adherence to the protocol.

Intervention
Healthcare action intended to benefit the patient, for example, drug treatment, surgical procedure, psychological therapy.

Intraoperative
The period of time during a surgical procedure.

Length of stay
The total number of days a participant stays in hospital.

Licence
See 'Product licence'.

Life-years gained
Mean average years of life gained per person as a result of the intervention compared with an alternative intervention.

Likelihood ratio
The likelihood ratio combines information about the sensitivity and specificity. It tells you how much a positive or negative result changes the likelihood that a patient would have the disease. The likelihood ratio of a positive test result (LR+) is sensitivity divided by 1- specificity.

Literature review
An article that summarises the evidence contained in a number

	of different individual studies and draws conclusions about their findings. It may or may not be systematically researched and developed.
Long-term care	Residential care in a home that may include skilled nursing care and help with everyday activities. This includes nursing homes and residential homes.
Loss to follow-up	Also known as attrition. The loss of participants during the course of a study. Participants that are lost during the study are often call dropouts.
Markov model	A method for estimating long-term costs and effects for recurrent or chronic conditions, based on health states and the probability of transition between them within a given time period (cycle).
Meta-analysis	A statistical technique for combining (pooling) the results of a number of studies that address the same question and report on the same outcomes to produce a summary result. The aim is to derive more precise and clear information from a large data pool. It is generally more reliably likely to confirm or refute a hypothesis than the individual trials.
Mini-Mental State Examination (MMSE)	A commonly used instrument for screening cognitive function. It is not suitable for making a diagnosis but can be used to indicate the presence of cognitive impairment.
Multidisciplinary team	A team of healthcare professionals with the different clinical skills needed to offer holistic care to people with complex problems such as delirium.
Multivariate model	A statistical model for analysis of the relationship between two or more predictor (independent) variables and the outcome (dependent) variable.
Negative predictive value (NPV)	[In screening/diagnostic tests:] A measure of the usefulness of a screening/diagnostic test. It is the proportion of those with a negative test result who do not have the disease, and can be interpreted as the probability that a negative test result is correct. It is calculated as follows: NPV = Number with a negative test who do not have disease/Number with a negative test.
Number needed to treat (NNT)	The number of patients that who on average must be treated to prevent a single occurrence of the outcome of interest.
Observational study	Retrospective or prospective study in which the investigator observes the natural course of events with or without control groups; for example, cohort studies and case–control studies.

Odds ratio	A measure of treatment effectiveness. The odds of an event happening in the treatment group, expressed as a proportion of the odds of it happening in the control group. The 'odds' is the ratio of events to non-events.
Outcome	Measure of the possible results that may stem from exposure to a preventive or therapeutic intervention. Outcome measures may be intermediate endpoints or they can be final endpoints. See 'Intermediate outcome'.
P-value	The probability that an observed difference could have occurred by chance, assuming that there is in fact no underlying difference between the means of the observations. If the probability is less than 1 in 20, the P value is less than 0.05; a result with a P value of less than 0.05 is conventionally considered to be 'statistically significant'.
Placebo	An inactive and physically identical medication or procedure used as a comparator in controlled clinical trials.
Polypharmacy	The use or prescription of multiple medications.
Positive predictive value (PPV)	In screening/diagnostic tests:] A measure of the usefulness of a screening/diagnostic test. It is the proportion of those with a positive test result who have the disease, and can be interpreted as the probability that a positive test result is correct. It is calculated as follows: PPV = Number with a positive test.
Postoperative	Pertaining to the period after patients leave the operating theatre, following surgery.
Post-test probability	For diagnostic tests. The proportion of patients with that particular test result who have the target disorder (post test odds/[1 + post-test odds]).
Power (statistical)	The ability to demonstrate an association when one exists. Power is related to sample size; the larger the sample size, the greater the power and the lower the risk that a possible association could be missed.
Preoperative	Pertaining to the period before surgery commences.
Pre-test probability	For diagnostic tests. The proportion of people with the target disorder in the population at risk at a specific time point or time interval. Prevalence may depend on how a disorder is diagnosed.
Prevalent delirium	Cases of delirium that are present at the first assessment of the person; it cannot be determined when the delirium began.

Primary care	Healthcare delivered to patients outside hospitals. Primary care covers a range of services provided by general practitioners, nurses and other healthcare professionals, dentists, pharmacists and opticians.
Primary outcome	The outcome of greatest importance, usually the one in a study that the power calculation is based on.
Product licence	An authorisation from the MHRA to market a medicinal product.
Prognosis	A probable course or outcome of a disease. Prognostic factors are patient or disease characteristics that influence the course. Good prognosis is associated with low rate of undesirable outcomes; poor prognosis is associated with a high rate of undesirable outcomes.
Prospective study	A study in which people are entered into the research and then followed up over a period of time with future events recorded as they happen. This contrasts with studies that are *retrospective.*
Publication bias	Also known as reporting bias. A bias caused by only a subset of all the relevant data being available. The publication of research can depend on the nature and direction of the study results. Studies in which an intervention is not found to be effective are sometimes not published. Because of this, systematic reviews that fail to include unpublished studies may overestimate the true effect of an intervention. In addition, a published report might present a biased set of results (e.g. only outcomes or sub-groups where a statistically significant difference was found.
Quality of life	See 'Health-related quality of life'.
Quality-adjusted life year (QALY)	An index of survival that is adjusted to account for the patient's quality of life during this time. QALYs have the advantage of incorporating changes in both quantity (longevity/mortality) and quality (morbidity, psychological, functional, social and other factors) of life. Used to measure benefits in cost-utility analysis. The QALYs gained are the mean QALYs associated with one treatment minus the mean QALYs associated with an alternative treatment.
Quantitative research	Research that generates numerical data or data that can be converted into numbers, for example clinical trials or the national Census which counts people and households.
Quick Reference Guide	An abridged version of NICE guidance, which presents the key priorities for implementation and summarises the recommendations for the core clinical audience.

Randomisation	Allocation of participants in a research study to two or more alternative groups using a chance procedure, such as computer-generated random numbers. This approach is used in an attempt to ensure there is an even distribution of participants with different characteristics between groups and thus reduce sources of bias.
Randomised controlled trial (RCT)	A comparative study in which participants are randomly allocated to intervention and control groups and followed up to examine differences in outcomes between the groups.
RCT	See 'Randomised controlled trial'.
Receiver operated characteristic (ROC) curve	A graphical method of assessing the accuracy of a diagnostic test. Sensitivity Is plotted against 1-specificity. A perfect test will have a positive, vertical linear slope starting at the origin. A good test will be somewhere close to this ideal.
Reference standard	The test that is considered to be the best available method to establish the presence or absence of the outcome – this may not be the one that is routinely used in practice.
Relative risk (RR)	The number of times more likely or less likely an event is to happen in one group compared with another (calculated as the risk of the event in group A/the risk of the event in group B).
Remit	The brief given by the Department of Health and Welsh Assembly Government at the beginning of the guideline development process. This defines core areas of care that the guideline needs to address.
Reporting bias	See publication bias.
Resource implication	The likely impact in terms of finance, workforce or other NHS resources.
Retrospective study	A retrospective study deals with the present/ past and does not involve studying future events. This contrasts with studies that are *prospective.*
Secondary outcome	An outcome used to evaluate additional effects of the intervention deemed a priori as being less important than the primary outcomes.
Selection bias	A systematic bias in selecting participants for study groups, so that the groups have differences in prognosis and/or therapeutic sensitivities at baseline. Randomisation (with concealed allocation) of patients protects against this bias.

Selection criteria	Explicit standards used by guideline development groups to decide which studies should be included and excluded from consideration as potential sources of evidence.
Sensitivity	Sensitivity or recall rate is the proportion of true positives which are correctly identified as such. For example in diagnostic testing it is the proportion of true cases that the test detects. See the related term 'Specificity'
Sensitivity analysis	A means of representing uncertainty in the results of economic evaluations. Uncertainty may arise from missing data, imprecise estimates or methodological controversy. Sensitivity analysis also allows for exploring the generalisability of results to other settings. The analysis is repeated using different assumptions to examine the effect on the results. One-way simple sensitivity analysis (univariate analysis): each parameter is varied individually in order to isolate the consequences of each parameter on the results of the study. Multi-way simple sensitivity analysis (scenario analysis): two or more parameters are varied at the same time and the overall effect on the results is evaluated. Threshold sensitivity analysis: the critical value of parameters above or below which the conclusions of the study will change are identified. Probabilistic sensitivity analysis: probability distributions are assigned to the uncertain parameters and are incorporated into evaluation models based on decision analytical techniques (For example, Monte Carlo simulation).
Significance (statistical)	A result is deemed statistically significant if the probability of the result occurring by chance is less than 1 in 20 ($p < 0.05$).
Specificity	The proportion of true negatives that a correctly identified as such. For example in diagnostic testing the specificity is the proportion of non-cases incorrectly diagnosed as cases. See related term 'Sensitivity'. In terms of literature searching a highly specific search is generally narrow and aimed at picking up the key papers in a field and avoiding a wide range of papers.
Stakeholder	Those with an interest in the use of the guideline. Stakeholders include manufacturers, sponsors, healthcare professionals, and patient and carer groups.
Subsyndromal delirium	A person who has some, but not all, the features of delirium.

Systematic review	Research that summarises the evidence on a clearly formulated question according to a pre-defined protocol using systematic and explicit methods to identify, select and appraise relevant studies, and to extract, collate and report their findings. It may or may not use statistical meta-analysis.
Treatment allocation	Assigning a participant to a particular arm of the trial.
Typical antipsychotic	These are sometimes referred to as first generation antipsychotics because they are the older medications used to treat psychotic symptoms. They were not called "typical" until the newer generation of these drugs (the 'atypical antipsychotics') were developed.
Univariate	Analysis which separately explores each variable in a data set.
Utility	A measure of the strength of an individual's preference for a specific health state in relation to alternative health states. The utility scale assigns numerical values on a scale from 0 (death) to 1 (optimal or 'perfect' health). Health states can be considered worse than death and thus have a negative value.

1 Introduction

1.1 What is a guideline?

Our clinical guidelines are recommendations for the care of individuals in specific clinical conditions or circumstances within the National Health Service (NHS) – from prevention and self-care through primary and secondary care to more specialised services. We base our clinical guidelines on the best available research evidence, with the aim of improving the quality of healthcare. We use predetermined and systematic methods to identify and evaluate the evidence relating to specific clinical questions.

Clinical guidelines can:

- provide recommendations for the treatment and care of people by health professionals
- be used to develop standards to assess the clinical practice of individual health professionals
- be used in the education and training of health professionals
- help patients to make informed decisions
- improve communication between patient and health professional

While guidelines assist the practice of healthcare professionals, they do not replace their knowledge and skills.

We produce our guidelines using the following steps:

1. Guideline topic is referred to the National Institute for Health and Clinical Excellence (NICE) from the Department of Health
2. Stakeholders register an interest in the guideline and are consulted throughout the development process.
3. The scope is prepared by the National Clinical Guideline Centre (NCGC)
4. The NCGC establish a guideline development group (GDG)
5. A draft guideline is produced after the group assesses the available evidence and makes recommendations
6. There is a consultation on the draft guideline.
7. The final guideline is produced.

The NCGC and NICE produce a number of versions of this guideline:

- the **full guideline** contains all the recommendations, plus details of the methods used and the underpinning evidence

- the **NICE guideline** presents the recommendations from the full version in a format suited to implementation by health professionals and NHS bodies
- the **quick reference guide** presents recommendations in a suitable format for health professionals
- information for the public (**'understanding NICE guidance'**) is written using suitable language for people without specialist medical knowledge.

This version is the full version. The other versions are available from NICE (www.nice.org.uk).

1.2 The need for this guideline

Delirium, sometimes called 'acute confusional state' is a common clinical syndrome characterised by disturbed consciousness and a change in cognitive function or perception that develops over a short period of time (usually 1-2 days).

Although the clinical presentation of delirium differs considerably from patient to patient, there are several characteristic features that help make the diagnosis. The standard criteria for delirium, are described in the 'Diagnostic and Statistical Manual of Mental Disorders' [DSM-IV] (1994):

- disturbance of consciousness (i.e., reduced clarity of awareness of the environment) with reduced ability to focus, sustain, or shift attention.
- a change in cognition (such as memory deficit, disorientation, language disturbance) or the development of a perceptual disturbance that is not better accounted for by a pre-existing, established, or evolving dementia.
- the disturbance develops over a short period of time (usually hours to days) and tends to fluctuate during the course of the day.
- there is evidence from the history, physical examination, and laboratory findings that: (1) the disturbance is caused by the direct physiological consequences of a general medical condition, (2) the symptoms in criteria (a) and (b) developed during substance intoxication, or during or shortly after, a withdrawal syndrome, or (3) the delirium has more than one aetiology".

Features of delirium are recent onset of fluctuating awareness, impairment of memory and attention, and disorganised thinking. Additional features may include hallucinations and disturbance of sleep-wake cycle. There are three clinical subtypes of delirium: hyperactive (characterised by hallucinations, delusions, agitation, and disorientation); hypoactive, which is particularly easy to miss in clinical practice (characterised by sleepy state, uninterested in activities of living, often unrecognised or labelled as dementia); or mixed (patients can move between the two subtypes). Delirium may be present when a person is admitted to hospital or long-term care (prevalent delirium) or it may develop during a hospital admission or residential stay in long-term care (incident delirium). It can be difficult to distinguish between delirium and dementia, and some people may have both conditions (delirium on dementia).

Delirium is a common but serious condition that is associated with poor outcomes. However, it can be prevented and treated if dealt with urgently.

There is a need for guidance to improve methods of appropriate identification, diagnosis, prevention and management of delirium. Failure to diagnose delirium, or misdiagnosis (mainly as dementia), can lead to medical emergencies being missed (ie. appropriate assessment and treatment may be omitted) and inappropriate treatment being given. Delirium is often preventable and improvements in care practices and other treatments are needed. The improved management of delirium has the potential to generate cost savings.

1.3 Remit

The following remit was received from the Department of Health in October 2007 as part of NICE's 17th wave programme of work.

The Department of Health asked the Institute:

> *Remit: "To prepare a clinical guideline on the diagnosis, prevention and management of delirium"*

1.4 What the guideline covers

This guideline covers adult patients (18 years and older) in a hospital setting and adults (18 and older) in long-term residential care. The guideline addresses: modifiable risk factors ('clinical factors') to identify people at risk of developing delirium; diagnosis of delirium in acute, critical and long-term care; as well as pharmacological and non-pharmacological interventions for a) reducing the incidence of delirium and its consequences, and b) to reduce the severity, duration and consequences of delirium in people who develop the condition.

Further details of the scope of the guideline can be found in Appendix A.

1.5 What the guideline does not cover

This guideline does not cover children and young people (under the age of 18 years), people receiving end-of-life care, people with intoxication and/or withdrawing from drugs or alcohol, and people with delirium associated with these states.

1.6 Who developed this guidance

This guideline was commissioned by NICE and developed initially by the National Collaborating Centre for Nursing and Supportive Care (NCC-NSC) which under merger status became part of the National Clinical Guideline Centre (NCGC). The

NCGC was formed on the 1st April 2009 and is one of four national collaborating centres (Cancer, Women and Children's Health, Mental Health and the NCGC) funded by NICE and comprises a partnership between a variety of academic, professional and patient-based organisations. As a multidisciplinary centre we draw upon the expertise of the healthcare professions and academics and ensure the involvement of patients in our work. Further information on the centre and our partner organisations can be found at our website (www.ncgc.ac.uk).

NICE funds the NCGC and thus supported the development of this guideline. The guideline development group was convened by the NCGC and chaired by Professor John Young in accordance with guidance from NICE.

The group met every 6-8 weeks during the development of the guideline. At the start of the guideline development process, all GDG members declared interests including consultancies, fee-paid work, share-holdings, fellowships and support from the healthcare industry. At all subsequent GDG meetings, members declared arising conflicts of interest, which were also recorded (Appendix B).

Members are either required to withdraw completely or for part of the discussion if their declared interest makes it appropriate, however this was not deemed necessary for any group members on this guideline.

Staff from the NCGC provided methodological support and guidance for the development process. They undertook systematic searches, retrieval and appraisal of the evidence and drafted the guideline. The glossary to the guideline contains definitions of terms used by staff and the GDG.

1.7 Related NICE guidance

NICE has developed or in the process of developing the following guidance (details available from www.nice.org.uk), some of which has been referred to in this guideline:

- Acutely ill patients in hospital: recognition of and response to acute illness in adults in hospital. NICE clinical guideline 50 (2007). Available from www.nice.org.uk/CG050.
- Infection control: prevention of healthcare-associated infection in primary and community care NICE clinical guideline 2 (2003). Available from www.nice.org.uk/CG2. This guideline is currently being updated.
- Nutrition support in adults: Nutrition support in adults: oral nutrition support, enteral tube feeding and parenteral nutrition. NICE clinical guideline 32 (2006). Available from www.nice.org.uk/CG032.
- Dementia: supporting people with dementia and their carers in health and social care. NICE clinical guideline 42 (2006). Available from www.nice.org.uk/CG042.
- Drug misuse: opioid detoxification. NICE clinical guideline 52 (2007). Available from www.nice.org.uk/CG0452.

- Surgical site infection – prevention and treatment of surgical site infection. NICE clinical guideline 74 (2008). Available from www.nice.org.uk/CG074.
- Schizophrenia – core interventions in the treatment and management of schizophrenia in primary and secondary care (update). NICE clinical guideline 82 (2009). Available from www.nice.org.uk/CG082.
- Alzheimer's disease - donepezil, galantamine, rivastigmine (review) and memantine for the treatment of Alzheimer's disease. NICE technology appraisal 111 (2007). Available from www.nice.org.uk/TA111.
- Schizophrenia - the clinical effectiveness and cost effectiveness of newer atypical antipsychotic drugs for schizophrenia. NICE technology appraisal 43 (2002). Available from www.nice.org.uk/TA43.
- Parkinson's disease – national clinical guideline for diagnosis and management in primary and secondary care. NICE clinical guideline 35 (2006). Available from www.nice.org.uk/CG035.
- Violence – the short-term management of disturbed/violent behaviour in in-patient psychiatric settings and emergency departments. NICE clinical guideline 25 (2005). Available from www.nice.org.uk/CG025.
- Alcohol use disorders: diagnosis and clinical management of alcohol-related physical complications. NICE clinical guideline 100 (2010). Available from www.nice.org.uk/CG100
- Alcohol dependence and harmful alcohol use. NICE clinical guideline. Publication expected February 2011.
- Falls: the assessment and prevention of falls in older people. NICE clinical guideline 21 (2004). Available from www.nice.org.uk/CG021.

2 Methodology

This guideline was commissioned by NICE and developed in accordance with the guideline development process outlined in 'The guidelines manual' (NICE 2009).

2.1 Developing the clinical questions

Clinical questions were developed to guide the literature searching process and to facilitate the development of recommendations by the GDG. They were drafted by the technical team and refined and validated by the GDG. The questions were based on the scope (Appendix A).

The full list of clinical questions addressed by the guideline is summarised in table 2.1 below:

Table 2.1: full list of clinical questions

Question	Relevant Chapter
Diagnosis	
What are the symptoms that indicate a person may have delirium?	6
What is the diagnostic accuracy of practical diagnostic tests compared with the reference standard DSM IV, to identify delirium in people in hospital and long-term care settings?	6
What are the diagnostic criteria that must be fulfilled to identify that a person has delirium?	6
Prognosis	
What are the risk factors for delirium?	7 and 8
What are the precipitating factors for delirium?	7
What are the consequences of delirium in terms of morbidity and mortality in a person in hospital or long-term care ?	9
Interventions	
Prevention of delirium in a hospital setting	
What are the most clinical and cost effective and safe pharmacological interventions for the prevention of delirium in people in hospital?	11A and 14
What are the most clinical and cost effective single-component, non-pharmacological interventions for the prevention of delirium in people in hospital?	10A
What are the most clinical and cost effective multicomponent interventions for the prevention of delirium in people in hospital?	10B
Prevention of delirium in a long-term care setting	
What are the most clinical and cost effective and safe pharmacological interventions for the prevention of delirium in people in long-term care?	11B and 14
What are the most clinical and cost effective single-component, non-pharmacological interventions for the prevention of delirium in people in long-term care?	10A
What are the most clinical and cost effective multicomponent interventions for the prevention of delirium in people in long-term care?	10B
Treatment of delirium in a hospital setting	
What are the most clinical and cost effective and safe pharmacological interventions for treating people with delirium in hospital?	13 and 14
What are the most clinical and cost effective single-component, non-pharmacological interventions for treating people with delirium in hospital?	No studies found
What are the most clinical and cost effective multicomponent interventions for treating people with delirium in hospital?	12
Treatment of delirium in a long-term care setting	
What are the most clinical and cost effective and safe pharmacological	13 and 14

interventions for treating people with delirium in long-term care?	
What are the most clinical and cost effective single-component, non-pharmacological interventions for treating people with delirium in long-term care?	No studies found
What are the most clinical and cost effective multicomponent interventions for treating people with delirium in long-term care?	No studies found
Patient information	
What information should be given to people at risk of developing delirium, or people with delirium, and their families or carers?	15
Other	
What is the prevalence of delirium in different hospital settings and in long-term care?	5

From these clinical questions, the technical team produced review questions and protocols to address these questions. The protocols are reported in the clinical effectiveness review methods section (2.3).

2.2 Searching the literature

2.2.1 Clinical literature search

The search strategies and the databases searched are presented in detail in Appendix C. All searches were carried out on the following core databases: Medline, Embase, Cinahl and The Cochrane Library. Additional databases were searched for individual reviews as appropriate.

Databases were searched using relevant subject headings and free-text terms. Where appropriate, study design filters were applied. Non-English language studies and abstracts were not reviewed initially, with the exception of studies translated for Cochrane reviews, but the GDG directed that a search was carried out for any RCT, regardless of language.

Searches were initially performed for articles published since 1994, the publication date of the DSM-IV which is the reference standard for the diagnosis of delirium. Following guidance from the GDG, a further search back to 1987 was carried out in order to retrieve studies using the earlier *Diagnostic and Statistical Manual III (Revised)* (DSMIII-R) as the reference standard.

All searches were updated to 17th August 2009. Hand-searching was not undertaken following NICE advice that exhaustive searching on every guideline review topic is not practical or efficient (Mason 2002). Reference lists of articles were checked for studies of potential relevance.

2.2.2 Sifting process

Once the search had been completed, the following sifting process took place:

- 1st sift: one reviewer sifted the title/abstract for articles that potentially met the eligibility criteria.

- 2nd sift: full papers were ordered that appeared relevant and eligible or where relevance/eligibility was not clear from the abstract.
- 3rd sift: full papers were appraised that meet eligibility criteria. Generally, one reviewer appraised the papers using an inclusion criteria form, and this was checked where necessary by a second reviewer.

Once individual papers were retrieved, the articles were checked for methodological rigour (see section 2.4), applicability to the UK and clinical significance. Assessment of study quality concentrated on dimensions of internal validity and external validity. At this stage, some studies were excluded if the interventions were not licensed for use in the UK or they were not regularly used in the UK. Studies in which the interventions were obsolete were also excluded.

2.2.3 Economic literature search

Economic evidence was obtained from systematic searches of the following databases in accordance with the NICE Guidelines Manual (NICE 2009): Medline, Embase, the Health Technology Appraisals (HTA) database and the NHS Economic Evaluations Database (NHSEED. The latter two databases were searched via The Cochrane Library.

Detailed search strategies can be found in Appendix J.

2.3 Clinical effectiveness review methods

This section describes the methods of reviewing that are common to all reviews of intervention studies, to reviews of prognostic factors and to reviews of diagnostic test accuracy. Further specific details are given in the individual reviews.

2.3.1 Selection criteria: general

The following selection criteria were to be applied to studies to determine their suitability for inclusion in the reviews:

Types of studies

For intervention studies, the randomised trial (RCT) and quasi randomised trial (for example, allocation by alternation and date of birth) were to be the primary trial designs. Non-randomised studies could be included only if there was no other evidence, with preference given to large cohort studies and comparative non-randomised designs; case series or case reports were not included and before-and-after studies were considered cautiously for prevention studies only.

For prognostic factor reviews, RCTs comparing groups with different risk factors (e.g. types of surgery) and cohort studies (prospective and retrospective) investigating the incidence of delirium or the consequences of delirium were to be the main study designs. We note that, for some risk factors (e.g. age), the randomised trial cannot be used as the study design. If there were no cohort studies available, case-control studies and cross-sectional surveys could be considered, with allowance made for the fact that they have increased potential for bias.

For reviews of diagnostic test accuracy, the cross sectional study was to be the primary study design. Studies were to be those in which diagnoses obtained using a new (index) test were compared with 'true' diagnoses obtained using a reference standard, with both tests being carried out in the same patients. Case control studies were to be considered only in the absence of cross sectional studies.

Studies were to be excluded if there were fewer than 20 patients in each arm for comparative studies and if there were fewer than 20 patients overall for cohort studies. We did not restrict the size of the studies of diagnostic test accuracy.

Studies were limited to the English language, initially, with the exception of studies translated for Cochrane reviews, but the GDG directed that a search was carried out for any RCT, regardless of the language.

Types of participants

For intervention studies, reviews were to be carried out separately to address interventions for prevention and treatment of delirium. Separate reviews were also done in the two main population groups: patients in a hospital setting and people in long-term care.

For prognostic factor reviews, the populations were not to be treated separately, although it was noted which population was concerned.

Reviews of diagnostic test accuracy are sensitive to the population, so long-term care, hospital setting and intensive care unit (ICU) were to be treated separately.

For all reviews, participants were to be adults (18 years and older) who were:

- Patients in a hospital setting, including surgical, medical, ICU, Accident and Emergency departments, and those in mental health settings
- In long-term care settings

Studies including children or young people were to be considered if the mean age was 18 years or older. Studies in the community could be included as indirect evidence for the long-term care population.

Excluded populations were to be:

- Children and young people (younger than 18 years).
- People receiving end-of-life care.
- People with intoxication and/or those who are withdrawing from drugs or alcohol, and/or (treatment intervention reviews) people with delirium associated with these states

For the treatment intervention reviews: participants were to have delirium. Delirium is defined according to criteria described in the DSM-IV (1994) (see Appendix I). Typically, delirium is diagnosed by examining changes in cognitive function, and this is linked to the DSM-IV criteria. Validated instruments, based on the operational application of the DSM-IV or DSM-III-R diagnostic criteria, are given in Appendix I.

2.3.2 Selection criteria: reviews of interventions

Types of intervention

The interventions considered varied across reviews. Interventions could be pharmacological or non-pharmacological (e.g. haloperidol, music therapy).

Pharmacological interventions were to be restricted to those licensed for use in the UK, but these drugs were not necessarily those indicated for delirium (there are no drugs for delirium in the British National Formulary (BNF)). Pharmacological reviews were to be carried out by class rather than by individual drug, but drugs within a class were to be reported as subgroups (e.g. atypical antipsychotics: olanzapine and risperidone).

Different doses, regimens and routes of delivery were to be permitted and studies were to be initially combined in analyses, regardless of these features.

Types of comparisons

The following comparisons were to be included:

i. Delirium intervention (**A**) versus placebo
ii. **A** versus usual care/no intervention
iii. **A** plus second intervention (**X**) versus **X** alone
iv. Within a class of interventions, **A1.1** versus **A1.2**
v. Across classes of interventions: **A1** versus **A2**

In analyses, comparisons (i) and (ii) could be combined, but (iii) was to be treated separately because of possible drug interactions.

2.3.3 Types of outcome measures

For studies of interventions for the prevention of delirium, the primary outcome was to be incidence of delirium. All included types and severities of delirium were to be combined. For reviews of patients in hospital, the primary outcome was to be measured during the hospital stay.

For the incidence of delirium, studies should report that the DSM-IV or the DSM- III-R and validated scales associated with them were used (see Appendix I). Other acceptable methods could include a structured clinical interview.

Secondary outcomes were to be:

- Duration of delirium
- Severity of delirium
- Length of stay in hospital
- Incidence of dementia or cognitive impairment
- Number of patients discharged to new long-term care placement (for studies in a hospital setting)
- Mortality
- Quality of life (patient)
- Quality of life (carer)
- Activities of daily living
- Use of psychotropic medication
- Incidence of post traumatic stress disorder
- Admission to hospital (for long-term care studies)

For studies of interventions for the treatment of delirium, the primary outcomes were to be:

- Duration of delirium
- Complete response (number recovered from delirium)

Secondary outcomes:

- Severity of delirium
- Length of stay
- Incidence of dementia / cognitive impairment
- Number of patients discharged to new long-term care placement (for those in hospital)
- Mortality
- Number of patients with persisting delirium
- Quality of life (patient)
- Quality of life (carer)

For all intervention reviews, other outcome measures to be recorded were:

- Adverse effects associated with the intervention (e.g. extrapyramidal symptoms).

2.3.4 Selection criteria: reviews of prognostic factors

Two types of prognostic factor reviews were carried out, investigating prognostic factors for delirium, and studying the consequences of delirium for people with delirium.

Prognostic (risk) factors

The risk factors to be considered for delirium are listed at the start of that review (section 7.2.1).

For the consequences of delirium review, the risk factor was to be one of:

- Incident delirium (although prevalent delirium was also acceptable)
- Persistent delirium: this was defined after McAvay (2006) as 'delirium in patients who met the full criteria for delirium at the discharge interview, or who had full delirium during the hospitalisation and partial symptoms at discharge'.
- Severity of delirium

Types of outcome measures

For the risk factors review, the following outcomes were to be included:

- Incidence of delirium
- Incidence of persistent delirium
- Severity of delirium
- Duration of delirium

For the consequences review, the following outcomes were to be included:

- Dementia/Cognitive impairment
- Progression of dementia
- Discharge to care home (for people who were in hospital)
- Falls
- Hospital admission (for people who were in long-term care)
- Post discharge care
- Post traumatic stress disorder
- Pressure Ulcers
- Mortality
- Impact on carers
- Length of stay
- Quality of life for patients

2.3.5 Selection criteria: reviews of diagnostic test accuracy

Prior tests

No prior tests were to have been undertaken

The index test

The following index tests, including the people operating them, were to be examined, subdivided by setting:

Hospital:

- Abbreviated Mental test (AMT); anyone could do this test
- Clock-drawing; could be used by untrained nurses or volunteers

- Confusion Assessment Method [long version] (long CAM); should be carried out by trained healthcare professionals
- Confusion Assessment Method [short version] (short CAM); should be carried out by trained healthcare professionals
- Delirium Rating Scale (DRS-98); should be carried out by trained healthcare professionals
- Mini Mental State Examination (MMSE) or other cognitive assessment instrument: trained healthcare professionals.

ICU:

- CAM-ICU and Richmond Agitation Sedation Scale (RASS) (together); should be carried out by trained healthcare professionals

The reference standard

The reference standard was to be DSM-IV or ICD-10; carried out by a trained specialist. These systems are further described in Appendix I.

The target condition

The target condition was to be delirium; subsyndromal delirium was not to be included.

2.3.6 Outcomes

For studies of diagnostic test accuracy, the outcomes to be recorded were sensitivity, specificity, positive predictive value, negative predictive value, likelihood ratio, diagnostic odds ratio, pre- and post-test probabilities. These were to be calculated from raw data, and occasionally raw data could be back-calculated from test accuracy statistics.

2.3.7 Data extraction

Data from included studies were extracted by one reviewer for each review, and randomly checked by a second reviewer, and entered into a Microsoft Access relational database that had been especially designed for the guideline.

2.4 Appraising the evidence

2.4.1 Appraisal of methodological quality of intervention studies

For randomised trials, the following factors were considered in assessing the potential for bias:

- *A priori* sample size calculation
- Method of generation **of the randomisation sequence**
- Allocation concealment at randomisation:
    - The means of preventing the treatment assignment being known before the time of allocation
- Baseline comparability of treatment groups for relevant risk factors
- Patients stated to be blinded, especially for comparisons with placebo:
    - Blinding involves hiding the nature of the intervention from participants, clinicians and treatment evaluators after allocation has taken place
    - Blinding may be not be possible depending on the nature of the interventions
    - Blinding may be more important for some outcomes than others:

- Outcome assessor stated to be blinded:
- No missing data for each outcome:
    - Studies with at least 20% of data missing from any group were to be considered to be potentially biased, more so if there is a differential drop out from any one group or if the missing data is known to be significantly different from the remaining data
    - Those with moderate loss to follow up (20 to 50%) were to be considered in sensitivity analyses
    - Those with 50% or more patients missing from any one group were to be regarded as flawed and not analysed further (but would be included in the review)
- Intention to treat analysis:
    - Trial participants should be analysed in the groups to which they were randomised regardless of which (or how much) treatment they actually received, and regardless of other protocol irregularities **and**
    - All participants should be included regardless of whether their outcomes were actually collected

For non-randomised intervention studies, the following factors were considered in assessing the potential for bias; further details are given in The Cochrane Handbook for Systematic Reviews of Interventions (http://www.cochrane-handbook.org/ : Box

13.1.a: Some types of non-randomised study design used for evaluating the effects of interventions)

- Selection bias:
    - Account is taken of the confounding factors, either by design (e.g. matching or restriction to particular subgroups) or by methods of analysis (e.g. stratification or regression modelling with propensity scores or covariates)
    - Confounding factors for delirium intervention reviews that the GDG believed should be taken into consideration were: age, cognitive impairment, sensory impairment, polypharmacy
- Prospectiveness:
    - On the basis of identification of participants; baseline assessment and treatment allocation; assessment of outcomes
- Blinding (see RCTs)
    - Of patients
    - Of outcome assessors
- No loss to follow up (see RCTs)
- Intention to treat (see RCTs)

2.4.2 Appraisal of methodological quality of studies of prognostic factors

Cohort studies were assessed using criteria based on the Newcastle-Ottawa checklist and the NICE Guidelines Manual. Studies were considered to be of acceptable quality if the asterisked statement(s) for each criterion were met; otherwise their quality rating was downgraded.

The following criteria were taken into consideration to give an overall quality rating, with examples given for risk factors for the incidence of delirium – similar arguments apply for the consequences review:

- Representativeness of the exposed cohort:
    - Truly representative of the community e.g. random sample from the guideline's population*****
    - Somewhat representative of the community e.g. hospital patients only*

  - Selected group e.g. cardiac operations
  - No description of the derivation of the cohort or unclear.

- Selection of the non exposed cohort:
  - Drawn from the same community as the exposed cohort*
  - Drawn from a different source – e.g. compared with general population levels in epidemiological studies
  - No description of the derivation of the non exposed cohort or unclear.
- Ascertainment of exposure:
  - Measurement of risk factor using an adequate method (e.g. MMSE for dementia)*
  - Measurement of risk factor using a partly adequate method*
  - Measurement of risk factor using an inadequate method (e.g. retrospective examination of chart records)
  - No description.

- Demonstration that the outcome of interest was not present at the start of the study:
  - Yes (includes analyses that excluded patients with prevalent delirium)*
  - No.

- Prospectiveness:
  - Prospective study*
  - Retrospective study
  - Unclear.

- Comparability of cohorts on the basis of the design or analysis:
  - Cohorts balanced at baseline for important factors (see below)*

- Adjusted for confounding factors in the analysis and has at least 10 events per factor in the analysis*
- Study has at least 8 to 10 events per factor and analysis is adjusted for at least 3 of 4 relevant factors in the analysis*
- Study adjusts for some confounders (or keeps them constant): 2 of 4 included in the analysis
- Study has fewer than 8 to10 events per factor in the analysis
- Study does not adjust for confounders.

In cohort studies, the best way to adjust for confounders is to use regression methods to adjust for all the factors at once in a multivariate analysis. For validity, there should be at least ten patients for each factor in the regression equation for continuous outcomes, or at least ten patients having the event (e.g. delirium) per factor for dichotomous outcomes. However, if there are insufficient relevant factors taken into account, the quality of the study should be downgraded.

The relevant factors that had to be included in the analysis were decided *a-priori* by the GDG using consensus methods. For the non-pharmacological risk factors review for the incidence of delirium, they were: age; sensory impairment, dementia/cognitive impairment and polypharmacy. For the pharmacological risk factors review, polypharmacy was excluded. The relevant factors for each consequence of delirium are given in that review. To qualify as a well adjusted study, the analysis should include at least 3 out of 4 of these factors (or they should be kept constant).

- Ascertainment of outcome:
  - Measurement of delirium using an adequate method (e.g. DSMIV, CAM)*
  - Measurement of delirium using a partly adequate method (e.g. MMSE)
  - Measurement of delirium using an inadequate method (e.g. retrospective examination of chart records)
  - No description.

- Adequacy of follow up of cohorts:
  - Complete follow-up: all participants accounted for*

- Participants lost to follow-up unlikely to introduce bias: more than 80% follow up*
- Follow-up rate less than 80% and no description of those lost
- No statement.

All these factors were taken into consideration to give an overall quality rating.

2.4.3 Appraisal of methodological quality of studies of diagnostic test accuracy

For studies of diagnostic test accuracy, the study quality was assessed using a modified version of the 'QUADAS' list, with each item scored as yes, no or unclear (Whiting 2003). The following factors were considered in assessing the potential for bias:

- Representative spectrum: whether or not the patients had delirium and were representative of the population of the review.
  - Studies that recruited a group of healthy controls and a group known to have the target disorder were coded as 'no' on this item
- Clear description of selection criteria
- Reference standard likely to classify the target condition correctly
- Acceptable delay between tests: period between the reference standard and the index test was short enough to be reasonably sure that the target condition did not change between the 2 tests; for delirium, the GDG considered this to be about half a day

An overall assessment for each study was given of ++ (good), + (acceptable, with some reservations) and – (unacceptable).

2.4.4 Data synthesis for intervention trials

Meta-analysis of similar trials, where appropriate, was carried out using *The Cochrane Collaboration's* analysis software, Review Manager (Version 5). Trials were pooled using a fixed effects model and plotted on forest plots. Where there was significant heterogeneity, sensitivity analyses and subgroup analyses were carried out. Meta-regression was not considered for this guideline as there were fewer than ten studies in the meta-analyses.

For dichotomous studies, intention to treat analyses (including all participants according to their assigned groups) were used, when reported by the study authors, and failing that, available case analyses (all those reporting an outcome) as reported by the authors. When there were incomplete data reported (more than 20% missing in any one group), we carried out sensitivity analyses, excluding these studies.

When it was possible to combine studies, outcomes were summarised for dichotomous data using relative risks. Numbers needed to treat, with their 95% confidence intervals (95% CI) and the control group rate (range of rates) to which they apply, were calculated from the risk difference where appropriate. The number needed to treat (NNT) is the number of patients who would have to be treated for one to have an improved outcome.

For continuous data, weighted mean differences were used to summarise the pooled data, and where the studies had different scales, standardised mean differences were used. Sometimes it may be necessary to invert scales (e.g. if one has the maximum value meaning poor outcome and in another it means a good outcome).

Studies, in which one or more reported final values and others reported change scores, were combined if the scales used were the same, otherwise they were reported separately. If both final values and change scores were reported in a single study, the former were used. Summary statistics and their 95% confidence intervals were reported where sufficient detail allowed their calculation, together with the control group range.

Where there were differences between studies in the way the results were reported, for example, summary statistics only or raw data, the summary statistic (e.g. RR) and its standard error was calculated from 95% confidence intervals, and the studies combined using the generic inverse variance method in Review Manager. For continuous outcomes reporting the difference in means with a p-value, the standard error was also calculated.

Where possible, account was taken of unit of randomisation errors (e.g. cluster trials).

Results from RCTs and non-randomised studies were not combined, but were reported as subgroups. Generally non-randomised studies were not included if the RCT data were adequate, but if the RCTs were very small or of poor quality, non-randomised studies could be included to give supplementary information.

Heterogeneity between trials was assessed by visual inspection of forest plots, noting where there was poor overlap of horizontal lines, and by using statistical measures: the χ^2 test for heterogeneity and the level of inconsistency, I^2 ($I^2 = [(\chi^2 - df)/ \chi^2] \times 100\%$, where df is the degrees of freedom). We considered that there was heterogeneity if the heterogeneity p-value was less than 0.1 and/or I^2 was greater than 50%. Any heterogeneity was explored further (see subgroup analyses below) and unexplained heterogeneous results were not used as the basis for recommendations.

Stratifications

Separate reviews were carried out for prevention and treatment, and for setting (hospital and long-term care).

Combining studies

Studies were combined regardless of:

- medical or surgical patients
- ICU or not
- risk of delirium, including baseline levels of dementia (for prevention reviews)
- dose of intervention

In pharmacological reviews, all the drugs in a particular class were considered in the same review, with individual drugs considered as subgroups in meta-analysis.

Subgroup analyses

If there was heterogeneity, subgroup analyses were carried out to investigate it.

The following subgroups were considered:

- For prevention reviews: people at high risk of delirium, such as those with dementia, may be distinguished from lower risk groups.
- Patients in ICU
- Type of intervention
- Dose of intervention
- Illness severity

Sensitivity analyses

Sensitivity analyses were carried out to investigate assumptions within the analyses. These included the following:

- Methodological quality
- Other features specific to each review.

In terms of methodological quality, we paid particular attention to allocation concealment and loss to follow-up (missing data). We did not include studies with more than 50% missing data in the analyses. Otherwise we carried out sensitivity analyses on studies that had between 20 and 50% missing data in any group.

2.4.5 Data synthesis for prognostic factor reviews

Odds ratios or relative risks, with their 95% confidence intervals, from multivariate analyses were extracted from the papers, and standard errors were calculated from the 95% CIs. The log (odds ratio) with its standard error was then entered into the generic inverse variance technique of Review Manager 5. Studies were not combined in a meta-analysis because they were observational studies. Sensitivity analyses were carried out on the basis of study quality, and the results were represented on forest plots and reported as ranges.

2.4.6 Data synthesis for reviews of diagnostic test accuracy

For diagnostic test accuracy studies, 2 by 2 tables were constructed from raw data, which allowed calculation of sensitivity, specificity, positive predictive value, negative predictive value, likelihood ratio, diagnostic odds ratio, pre- and post-test probabilities. Some of this was done using an Access database, and Review Manager (version 5) was also used for the calculation of sensitivity and specificity and the representation of these in both forest plots and the receiver operating characteristic (ROC) space.

2.4.7 Grading evidence

The GRADE‡ scheme (GRADE working group 2004) was used to assess the quality of the evidence for each outcome not each study, using the approach described below, and evidence summaries across all outcomes were produced.

According to the GRADE scheme, evidence is classified as high, moderate, low or very low:

- High: further research is very unlikely to change our confidence in the estimate of effect
- Moderate: further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate
- Low: further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate
- Very low: any estimate of effect is very uncertain.

The procedure adopted when using GRADE was:

‡ GRADE – Grading of Recommendations Assessment, Development and Evaluation

- A quality rating was assigned, based on the study design, for example, RCTs started as high and observational studies as low.
- This rating was up- or down-graded according to specified criteria: study quality, consistency, directness, preciseness and reporting bias. These criteria are detailed below. Criteria were given a downgrade mark of −1 or –2 depending on the severity of the limitations.
- The downgrade/upgrade marks were then summed and the quality rating revised. For example, a decrease of –2 points for an RCT would result in a rating of 'low'.

Reasoning was explained for the downgrade marks.

The GRADE scheme was used for both RCTs and observational studies for pharmacological intervention studies and for single component non-pharmacological interventions. For non-pharmacological interventions, evidence profiles were not generated due to the complexity of the mulitcomponent interventions that were considered. Part of the interpretation of the interventions included a themed analysis and assessing evidence which incorporates a qualitative aspect is currently not undertaken within the GRADE scheme. In absence of a GRADE scheme, a narrative summary of the quality of evidence based on the quality appraisal criteria for randomised and non-randomised studies was presented in the review.

Evidence profiles were not generated for diagnostic studies as currently there is not a validated approach for summarisng a body of evidence for studies on diagnostic test accuracy. In absence of a GRADE scheme, a narrative summary of the quality of the evidence, based on the quality appraisal criteria from QUADAS was presented in the review.

Risk of bias

Risk of bias is assessed against standard criteria, depending on the study design. For randomised trials, we took into account: the adequacy of allocation concealment; blinding of participants for comparisons and outcomes susceptible to bias; attrition (missing data) and baseline comparability. A downgrade mark of –1 was given for inadequate or unclear allocation concealment and for a loss to follow-up of more than 20% in any one group or overall. Studies with more than 50% missing data were excluded from the analysis unless they were the only study, in which case they were given a downgrade mark of –2. If the evidence was a meta-analysis of several studies, we took into consideration the proportion and weighting of higher risk studies, and in some instances carried out sensitivity analyses disregarding these studies and giving a separate rating for the new meta-analysis.

Inconsistency

When several RCTs have widely differing estimates of treatment effect (heterogeneity or variability in results), the results are regarded as inconsistent. We defined this as a p-value for heterogeneity less than 0.1 and/or an I^2 value greater than 50%. Where this was the case, we gave a downgrade mark of −1. If the p-value was less than 0.1 and the I^2 value was greater than 80%, we gave a downgrade mark of −2. Where possible, we carried out pre-defined subgroup analyses to investigate heterogeneity and reported these results separately.

Indirectness

Directness refers to the extent to which the population, interventions, comparisons and outcome measures are similar to those defined in the inclusion criteria for the reviews. Indirectness is only relevant if there is a compelling reason to expect important differences in the size of the effect. For example, many interventions have more or less the same relative effects across patient groups, so extrapolation is possible and reasonable. There are various types of indirectness that can be found in studies, but most relevant to this guideline are:

- When the setting is different from those of the guideline, e.g. community setting, rather than long-term care
- When the method for assessment of delirium is partly adequate or inadequate

Imprecision

This is a rather subjective, but nevertheless important category. Evidence is considered to be imprecise if:

- There are sparse data (only a few events and they are uninformative).
- The confidence interval for the effect estimate is consistent with different conclusions, for example, both a clinically important effect (benefit or harm) and no clinically important effect; or the CI is consistent with important harms, no clinically important effect and important benefits. Precision requires the GDG to decide what are clinically important harms and benefits for that outcome measure. For dichotomous outcomes we used a relative risk reduction of 25% (RR of 1.25 or 0.75) to indicate the clinically important threshold. For continuous outcomes the GDG determined that the clinically important threshold for a difference between intervention groups was 0.5 days for a stay in ICU, 1 day for a stay in hospital, 1 day for duration of delirium, and a change of 20% on any of the scales used (linearity assumed).
- If the confidence interval did not cross either of the clinically important thresholds (i.e. precise rating), the sample size was taken into consideration. If there was a power calculation for that outcome and comparison, it was used to decide if a study was 'small', otherwise the optimal information size was calculated (or 300 events total was assumed).

High risk of publication bias

- Papers are more likely to be published if their results are statistically significant. The existence of publication bias in the studies in a meta-analysis can be investigated in a limited way using funnel plots, in which the standard error is plotted against the log odds ratio, the log relative risk or the mean difference. Asymmetry is indicative of publication bias. This method is usually only useful when there are at least five studies. Publication bias should also be considered if there is industry funding, but the GDG decided that industry sponsored studies should not be regarded as potentially biased for these outcomes.

2.5 Cost-effectiveness review methods

Information on cost-effectiveness is important for guideline development as it aids decision making on the application of intervention options in the different population groups considered in the guideline. It provides evidence on the cost and health impact of different intervention options considered during the process of guideline development. At the initial stage of the delirium guideline development, the health economist in conjunction with the GDG identified priority areas for cost-effectiveness evidence. The use of delirium prevention and treatment interventions in hospital and long-term care settings were identified as high priority areas for cost-effectiveness evidence. They were classified as high priority as the prevention and treatment of delirium would save NHS and PSS (Personal Social Services) resources as well as improve patients' health related quality of life. Information on the additional benefit associated with different strategies was also required. It was therefore necessary to look for health economic information on the intervention strategies and we started by reviewing published economic evaluations.

A systematic review was carried out to identify and appraise existing published economic evaluations that are relevant to the guideline's clinical questions. An article had to present a full or partial economic evaluation to be included in this review. A full economic evaluation compares all relevant cost and patient outcomes and uses these to estimate a single measure of incremental cost and benefits. The different forms of economic evaluation include cost-effectiveness, cost-utility, cost-benefit or cost-minimisation analysis. A partial economic evaluation only reports some of the relevant cost and patient outcomes. Studies reporting data from non-OECD (Organisation for Economic Co-operation and Development) member countries were excluded as these were felt to be less applicable to current practice in the UK. Publications that dealt with palliative care were removed as these were outside the scope of the guideline. For trial based economic evaluations, studies were excluded if they did not meet the inclusion criteria for the clinical effectiveness review.

We initially searched Medline, Embase, NHSEED and HTA databases starting from 1994 to June 2008. An economics filter was applied to the Medline and Embase searches to identify relevant economic literature. The search terms used in Medline are given in Appendix J. The economics and quality of life filter is as listed in Appendix J. The terms were suitably adapted for searches in Embase, NHSEED and HTA. A total of

755 publications were sifted by the Health Economist. Sifting was done by reading the title and abstract of the publications and full papers were ordered for any potential economic evaluations. We ordered 12 publications (Bracco 2007, Pitkala 2008, Rizzo 2001, Robinson 2002, The Medical and Health Research Council of the Netherlands [ongoing], Beaupre 2006, Heyman and Lombardo 1995, Caplan and Harper 2007, Pandharipande 2007, Rubin 2006, Webster 1999, Caplan 2006) and four of them were reviewed (Bracco 2007, Pitkala 2008, Rizzo 2001, Robinson 2002). The outcomes of interest were intervention and non-intervention costs, the incidence and severity of delirium, incidence of complete recovery from delirium, Quality-adjusted life year (QALY) measure, and delirium-attributable mortality rate. The four papers reviewed (Bracco 2007, Pitkala 2008, Rizzo 2001, Robinson 2002) are described under the relevant clinical questions (Appendix J).

None of the identified economic evaluations were directly applicable to the guideline population. None of the studies assessed costs from a UK NHS and PSS perspective and none measured health benefits in QALYs. None of the studies discounted future costs and outcomes appropriately and none carried out a robust sensitivity analysis on the results of the economic analysis. We carried out update searches up to August 2009 but did not identify further relevant economic evaluation studies. As there was a lack of high quality, relevant evidence on the cost-effectiveness of the interventions included in the guideline, it became necessary to develop an original economic evaluation model to determine the cost-effectiveness of strategies for the prevention and treatment of delirium in different care settings.

2.6 Cost-effectiveness modelling

The details of the economic model are described in Appendix J.

We developed original models for intervention strategies in hospital care settings but could not develop any models for prevention and treatment strategies in the long-term care setting. This was because there was a lack of evidence from the long-term care setting which could be used to construct a cost-effectiveness model. The evidence on the adverse consequences of delirium came from studies that were carried out in the hospital setting (chapter 9). The efficacy estimates of the interventions that we modelled came from studies carried out in hospital settings. Furthermore, the costing of the multicomponent interventions was based on the assumption that they were applied in the hospital. We were not confident that we could use this evidence to model the cost-effectiveness of these interventions in long-term care setting.

The outcomes of interest for the model were incremental cost and QALY gained. Costs were assessed from an NHS and PSS perspective. These outcomes were used to estimate the incremental cost-effectiveness ratio and net monetary benefit. Incremental net monetary benefit is defined below. Future costs and QALYs were discounted at a rate of 3.5% per annum. This is in line with the reference case advocated by NICE (NICE 2008 [manual on TA]).

In the base case analysis, the cost effectiveness of an intervention was determined using the threshold, £20,000 per QALY, and all interventions were compared to the usual care. If an intervention strategy costs less than the comparator and generates greater benefit it is described as being dominant and is unequivocally cost-effective. If the intervention is more effective but more costly, the incremental cost per QALY is estimated and compared to the cost-effectiveness threshold of £20,000 to £30,000 per QALY in line with the principles stated in the NICE Technology Appraisal Manual (NICE 2008 [manual on TA]). Another alternative to using incremental cost and QALYs to estimate cost-effectiveness is the use of the Incremental Net Monetary Benefit (INMB). The INMB is the monetary value of an intervention compared to an alternative for a specific cost-effectiveness threshold. It is calculated as

Cost-effectiveness Threshold * incremental QALY – incremental cost.

An intervention is cost-effective if it has an INMB that is greater than zero.

We constructed our model using the best available evidence and according to the NICE reference case for economic evaluation (NICE 2008 [manual on TA]).). We described explicitly the assumptions made in the model as well as the uncertainties in the model input parameters. The results of the model were interpreted by the GDG bearing the assumptions in mind. We used deterministic and probabilistic sensitivity analyses to explore the impact of the assumptions and uncertainties on the model results. We discussed the limitations of the model. Further details on the cost-effectiveness model are given in chapter 16. For those clinical questions which were not prioritised for an original economic evaluation the GDG considered the likely cost-effectiveness of the interventions by making a qualitative judgement on the likely costs, health benefits and potential harms of interventions.

2.7 Developing recommendations

Over the course of the guideline development process, the GDG was presented with the following:

- The clinical and economic evidence reviews. All evidence tables are in Appendices D, E,F and G.
- Forest plots of results from studies, including meta-analyses where appropriate. (Appendix K)
- A description of the methods and results of the cost-effectiveness analysis (chapter 16).

Recommendations were drafted on the basis of this evidence whenever it was available.

When clinical and economic evidence was poor or absent, the GDG proposed recommendations based on their expert opinion.

The GDG also developed a care pathway algorithm according to the recommendations (see section **Error! Reference source not found.**).

2.8 Research recommendations

When areas were identified for which good evidence was lacking, the guideline development group considered making recommendations for future research. Decisions about inclusion were based on factors such as:

- the importance to patients or the population
- national priorities
- potential impact on the NHS and future NICE guidance
- ethical and technical feasibility.

The GDG identified five high priority research recommendations (after discussion and voting). The full list of recommendations for future research, as well as those chosen as high priority, can be found in Appendix H.

2.9 Key priorities for implementation

To assist users of the guideline in deciding the order in which to implement the recommendations, the GDG identified ten key priorities for implementation. The decision was made after discussion and independent voting by the GDG. They selected recommendations that would:

- have a high impact on outcomes that are important to patients
- have a high impact on reducing variation in care and outcomes
- lead to a more efficient use of NHS resources
- promote patient choice
- promote equalities.

In doing this the GDG also considered which recommendations were particularly likely to benefit from implementation support. They considered whether a recommendation:

- relates to an intervention that is not part of routine care
- requires changes in service delivery
- requires retraining staff or the development of new skills and competencies
- highlights the need for practice to change
- affects and needs to be implemented across various agencies or settings (complex interactions)

- may be viewed as potentially contentious, or difficult to implement for other reasons.

2.10 Validation of the guideline

The first draft of this guideline was posted on the NICE website for an 8-week consultation between 11th November 2009 and 6th January 2010 and registered stakeholders were invited to comment. The GDG responded to comments and an amended version of the guideline was produced.

2.11 Disclaimer and funding

Healthcare providers need to use clinical judgement, knowledge and expertise when deciding whether it is appropriate to apply guidelines. The recommendations cited here are a guide and may not be appropriate for use in all situations. The decision to adopt any of the recommendations cited here must be made by the practitioner in light of individual patient circumstances, the wishes of the patient, clinical expertise and resources.

The National Clinical Guideline Centre disclaim any responsibility for damages arising out of the use or non-use of these guidelines and the literature used in support of these guidelines.

The National Collaborating Centre for Nursing and Supportive Care (now a part of the National Clinical Guideline Centre) were commissioned by the National Institutie for Health and Clinical Excellence to undertake the work on this guideline.

2.12 Updating the guideline

This guideline will be updated when in concordance with NICE guidelines manual (NICE 2009).

3 Key messages of the guideline

3.1 Key priorities for implementation

The GDG identified ten key priorities for implementation. The decision was made after discussion and voting by the GDG. The recommendations chosen by the GDG as key priorities for implementation are listed below. The numbering of the recommendations in parentheses is as per the NICE version of the guideline.

In addition the GDG wanted to highlight the importance of being aware of delirium and its consequences and so a prominent statement (THINK DELIRIUM) has been included below.

"THINK DELIRIUM"

Be aware that people in hospital or long-term care may be at risk of delirium. This can have serious consequences (such as increased risk of dementia and/or death) and, for people in hospital, may increase their length of stay in hospital and their risk of new admission to long-term care.

Risk factor assessment

When people first present to hospital or long-term care, assess them for the following risk factors. If any of these risk factors is present, the person is at risk of delirium.

- Age 65 years or older.
- Cognitive impairment (past or present) and/or dementia.[1] If cognitive impairment is suspected, confirm it using a standardised and validated cognitive impairment measure.

[1] If dementia is suspected, refer to further information on the diagnosis, treatment and care of people with dementia in 'Dementia: supporting people with dementia and their carers in health and social care' (NICE clinical guideline 42).

- Current hip fracture.

- Severe illness (a clinical condition that is deteriorating or is at risk of deterioration)[2].

[1.1.1]

Indicators of delirium: at presentation

At presentation, assess people at risk for recent (within hours or days) changes or fluctuations in behaviour. These may be reported by the person at risk, or a carer or relative. Be particularly vigilant for behaviour indicating hypoactive delirium (marked *). These behaviour changes may affect:

- Cognitive function: for example, worsened concentration*, slow responses*, confusion.

- Perception: for example, visual or auditory hallucinations.

- Physical function: for example, reduced mobility*, reduced movement*, restlessness, agitation, changes in appetite*, sleep disturbance.

- Social behaviour: for example, lack of cooperation with reasonable requests, withdrawal*, or alterations in communication, mood and/or attitude.

If any of these behavior changes are present, a healthcare professional who is trained and competent in diagnosing delirium should carry out a clinical assessment to confirm the diagnosis. **[1.2.1]**

Interventions to prevent delirium

Ensure that people at risk of delirium are cared for by a team of healthcare professionals who are familiar to the person at risk. Avoid moving people within and between wards or rooms unless absolutely necessary. **[1.3.1]**

Give a tailored multicomponent intervention package:

- Within 24 hours of admission, assess people at risk for clinical factors contributing to delirium

---

[2] For further information on recognising and responding to acute illness in adults in hospital see 'Acutely ill patients in hospital' (NICE clinical guideline 50).

- Based on the results of this assessment, provide a multicomponent intervention tailored to the person's individual needs and care setting as described in recommendations 1.3.3.1–1.3.3.10. **[1.3.2]**

The tailored multicomponent intervention package should be delivered by a multidisciplinary team trained and competent in delirium prevention. **[1.3.3]**

Diagnosis (specialist clinical assessment)

If indicators of delirium are identified, carry out a clinical assessment based on the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) criteria or short Confusion Assessment Method (short CAM) to confirm the diagnosis. In critical care or in the recovery room after surgery, CAM-ICU should be used. A healthcare professional who is trained and competent in the diagnosis of delirium should carry out the assessment. If there is difficulty distinguishing between the diagnoses of delirium, dementia or delirium superimposed on dementia, treat for delirium first. **[1.5.1]**

Ensure that the diagnosis of delirium is documented both in the person's hospital record and in their primary care health record. **[1.5.2]**

Treatment of delirium

Initial management

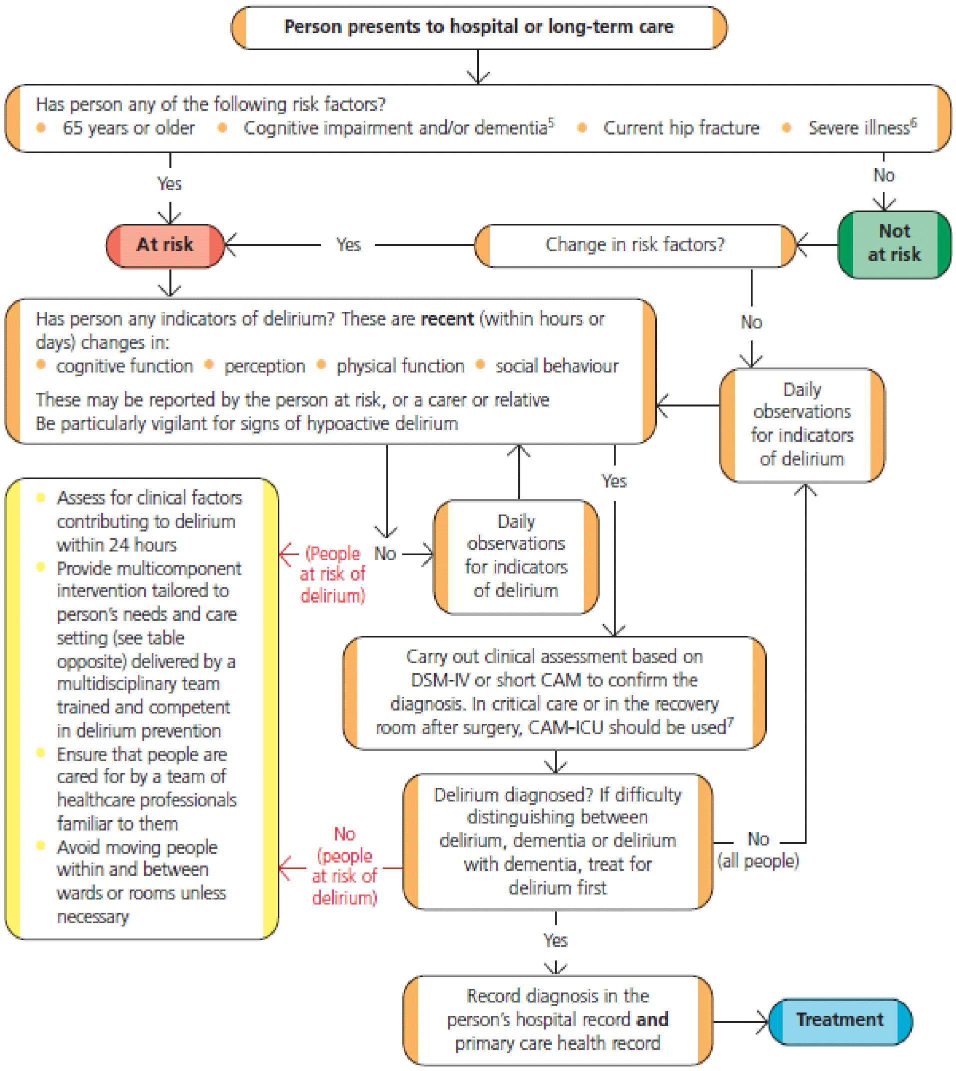
In people diagnosed with delirium, identify and manage the possible underlying cause or combination of causes. **[1.6.1]**

Ensure effective communication and reorientation (for example explaining where the person is, who they are, and what your role is) and provide reassurance for people diagnosed with delirium. Consider involving family, friends and carers to help with this. Provide a suitable care environment (see recommendation 1.3.1). **[1.6.2]**

Distressed people

If a person with delirium is distressed or considered a risk to themselves or others and verbal and non-verbal de-escalation techniques are ineffective or inappropriate, consider giving short-term (usually for 1 week or less) haloperidol[3] or olanzapine[3]. Start at the lowest clinically appropriate dose and titrate cautiously according to symptoms. **[1.6.4]**

Preventing and diagnosing delirium

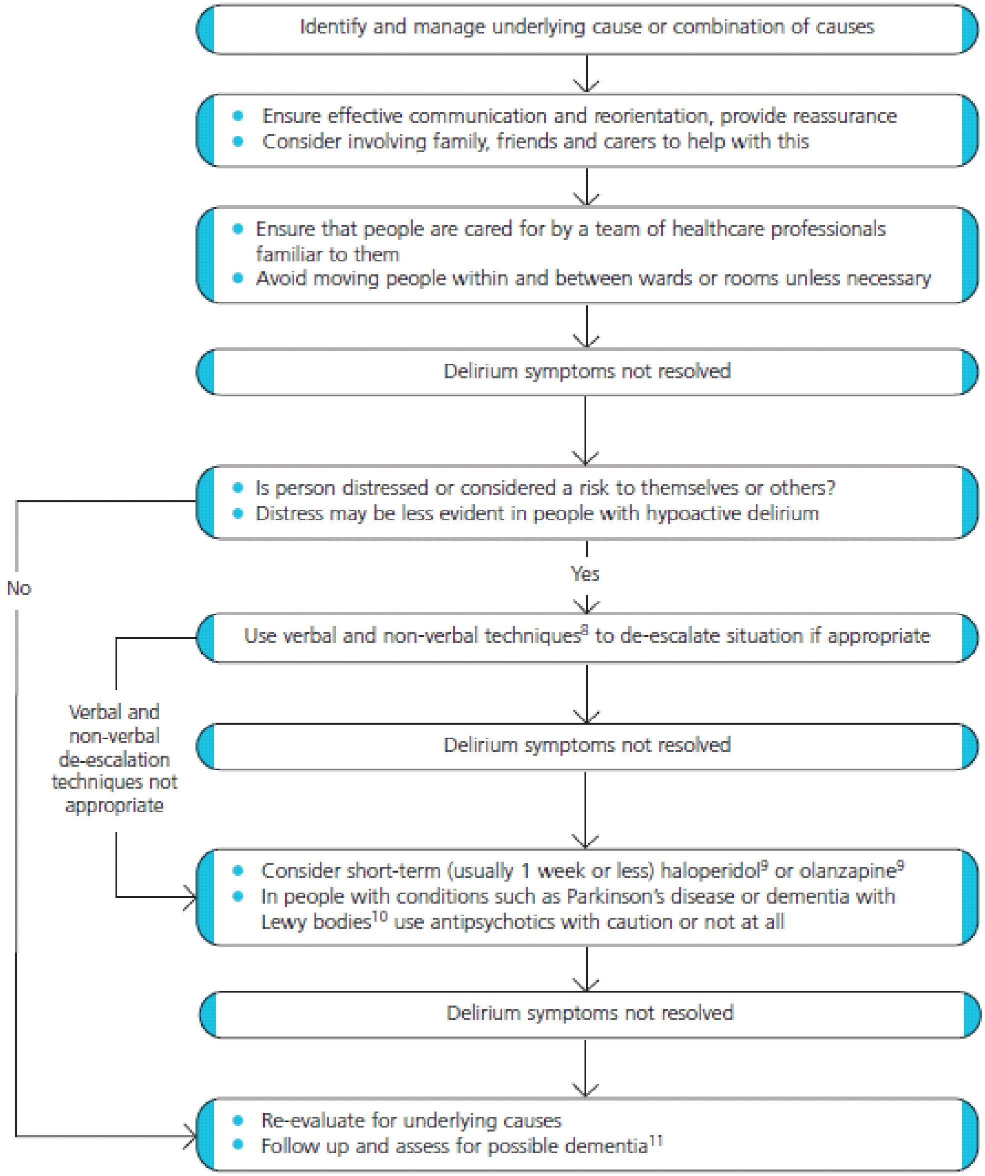


[5] If cognitive impairment is suspected, confirm using a standardised and validated cognitive impairment measure. If dementia is suspected, refer to 'Dementia: supporting people with dementia and their carers in health and social care' (NICE clinical guideline 42).
[6] For further information on recognising and responding to acute illness in adults in hospital see 'Acutely ill patients in hospital' (NICE clinical guideline 50).
[7] A healthcare professional trained and competent in the diagnosis of delirium should carry out this assessment.

---

[3] Haloperidol and olanzapine do not have UK marketing authorisation for this indication.

Treating delirium



[8] See 'Violence' (NICE clinical guideline 25).
[9] Haloperidol and olanzapine do not have UK marketing authorisation for this indication.
[10] For more information on the use of antipsychotics for these conditions, see 'Parkinson's disease' (NICE clinical guideline 35) and 'Dementia' (NICE clinical guideline 42).
[11] For more information on dementia see 'Dementia' (NICE clinical guideline 42).

Clinical indicators that can contribute to delirium

Factor	Preventive intervention
Cognitive impairment and/or disorientation	• Provide appropriate lighting and clear signage. A clock (consider providing a 24-hour clock in critical care) and a calendar should also be easily visible to the person at risk. • Reorientate the person. Explaining where they are, who they are, and what your role is. • Introduce cognitively stimulating activities (for example, reminiscence). • Facilitate regular visits from family and friends.
Dehydration and/or constipation	• Encourage the person to drink. Consider offering subcutaneous or intravenous fluids if necessary. • Seek advice if necessary when managing fluid balance in people with comorbidities (for example, heart failure or chronic kidney disease).
Hypoxia	• Assess for hypoxia and optimise oxygen saturation if necessary, as clinically appropriate.
Infection	• Look for and treat infection. • Avoid unnecessary catheterisation. • Implement infection control procedures in line with 'Infection control' (NICE clinical guideline 2).
Immobility or or limited mobility or	• Encourage people to: – mobilise soon after surgery – walk (provide walking aids if needed – these should be accessible at all times) • Encourage all people, including those unable to walk, to carry out active range-of-motion exercises.
Multiple medications	• Carry out a medication review for people taking multiple drugs, taking into account both the type and number of medications.
Pain	• Assess for pain. Look for non-verbal signs of pain, particularly in those with communication difficulties (for example, people with learning difficulties or dementia, or people on a ventilator or who have a tracheostomy). • Start and review appropriate pain management in any person in whom pain is identified or suspected.
Poor nutrition	• Follow the advice given on nutrition in 'Nutrition support in adults' (NICE clinical guideline 32). • If people have dentures, ensure they fit properly.
Sensory impairment	• Resolve any reversible cause of the impairment, such as impacted ear wax • Ensure working hearing and visual aids are available to and used by people who need them.

Factor	Preventive intervention
Sleep disturbance	• Promote good sleep patterns and sleep hygiene[4] by: – avoiding nursing or medical procedures during sleeping hours, if possible – scheduling medication rounds to avoid disturbing sleep, and – reducing noise to a minimum during sleep periods.

[4] For more information on good sleep hygiene, see 'Parkinson's disease' (NICE clinical guideline 35).

4 Summary of recommendations

"THINK DELIRIUM"

The GDG highlighted that all healthcare professionals should:

> *Be aware that people in hospital or long-term care may be at risk of delirium. This can have serious consequences (such as increased risk of dementia and/or death) and, for people in hospital, may increase their length of stay in hospital and their risk of new admission to long-term care.*

(NOTE: the numbering of the recommendations in parentheses is as per the NICE version of the guideline.)

Risk factor assessment

When people first present to hospital or long-term care, assess them for the following risk factors. If any of these risk factors is present, the person is at risk of delirium.

- Age 65 years or older.
- Cognitive impairment (past or present) and/or dementia.[5] If cognitive impairment is suspected, confirm it using a standardised and validated cognitive impairment measure.
- Current hip fracture.
- Severe illness (a clinical condition that is deteriorating or is at risk of deterioration)[6].

[1.1.1]

Observe people at every opportunity for any changes in the risk factors for delirium. **[1.1.2]**

---

[5] If dementia is suspected, refer to further information on the diagnosis, treatment and care of people with dementia, in 'Dementia: supporting people with dementia and their carers in health and social care' (NICE clinical guideline 42).

[6] For further information on recognising and responding to acute illness in adults in hospital see 'Acutely ill patients in hospital' (NICE clinical guideline 50).

Indicators of delirium: at presentation

At presentation, assess people at risk for recent (within hours or days) changes or fluctuations in behaviour. These may be reported by the person at risk, or a carer or relative. Be particularly vigilant for behaviour indicating hypoactive delirium (marked *). These behaviour changes may affect:

- Cognitive function: for example, worsened concentration*, slow responses*, confusion.
- Perception: for example, visual or auditory hallucinations.
- Physical function: for example, reduced mobility*, reduced movement*, restlessness, agitation, changes in appetite*, sleep disturbance.
- Social behaviour: for example, lack of cooperation with reasonable requests, withdrawal*, or alterations in communication, mood and/or attitude.

If any of these behaviour changes are present, a healthcare professional who is trained and competent in diagnosing delirium should carry out a clinical assessment to confirm the diagnosis. **[1.2.1]**

Interventions to prevent delirium

Ensure that people at risk of delirium are cared for by a team of healthcare professionals who are familiar to the person at risk. Avoid moving people within and between wards or rooms unless absolutely necessary. **[1.3.1]**

Give a tailored multicomponent intervention package:

- Within 24 hours of admission, assess people at risk for clinical factors contributing to delirium.
- Based on the results of this assessment, provide a multicomponent intervention tailored to the person's individual needs and care setting as described in recommendations 1.3.3.1–1.3.3.10. **[1.3.2]**

The tailored multicomponent intervention package should be delivered by a multidisciplinary team trained and competent in delirium prevention. **[1.3.3]**

[1.3.3.1] Address cognitive impairment and/or disorientation by:

- providing appropriate lighting and clear signage; a clock (consider providing a 24-hour clock in critical care) and a calendar should also be easily visible to the person at risk
- talking to the person to reorientate them by explaining where they are, who they are, and what your role is
- introducing cognitively stimulating activities (for example, reminiscence)
- facilitating regular visits from family and friends.

[1.3.3.2] Address dehydration and/or constipation by:

- ensuring adequate fluid intake to prevent dehydration by encouraging the person to drink – consider offering subcutaneous or intravenous fluids if necessary
- taking advice if necessary when managing fluid balance in people with comorbidities (for example, heart failure or chronic kidney disease).

[1.3.3.3] Assess for hypoxia and optimise oxygen saturation if necessary, as clinically appropriate.

[1.3.3.4] Address infection by:

- looking for and treating infection
- avoiding unnecessary catheterisation
- implementing infection control procedures in line with 'Infection control' (NICE clinical guideline 2).

[1.3.3.5] Address immobility or limited mobility through the following actions:

- Encourage people to:
    - mobilise soon after surgery
    - walk (provide appropriate walking aids if needed – these should be accessible at all times).

- Encourage all people, including those unable to walk, to carry out active range-of-motion exercises.

[1.3.3.6] Address pain by:

- assessing for pain
- looking for non-verbal signs of pain, particularly in those with communication difficulties (for example, people with learning difficulties or dementia, or people on a ventilator or who have a tracheostomy)
- starting and reviewing appropriate pain management in any person in whom pain is identified or suspected.

[1.3.3.7]

Carry out a medication review for people taking multiple drugs, taking into account both the type and number of medications.

[1.3.3.8] Address poor nutrition by:

- following the advice given on nutrition in 'Nutrition support in adults' (NICE clinical guideline 32)
- if people have dentures, ensuring they fit properly.

[1.3.3.9] Address sensory impairment by:

- resolving any reversible cause of the impairment, such as impacted ear wax
- ensuring hearing and visual aids are available to and used by people who need them, and that they are in good working order.

[1.3.3.10] Promote good sleep patterns and sleep hygiene[7] by:

- avoiding nursing or medical procedures during sleeping hours, if possible

[7] For more information on good sleep hygiene, see 'Parkinson's disease' (NICE clinical guideline 35).

- scheduling medication rounds to avoid disturbing sleep
- reducing noise to a minimum during sleep periods.

Indicators of delirium: daily observations

Observe at least daily, all people in hospital or long-term care for recent (within hours or days) changes or fluctuations in usual behaviour (see recommendation 1.2.1). These may be reported by the person at risk, or a carer or relative.

If any of these behaviour changes is present, a healthcare professional who is trained and competent in the diagnosis of delirium should carry out a clinical assessment to confirm the diagnosis. **[1.4.1]**

Diagnosis (specialist clinical assessment)

If indicators of delirium are identified, carry out a clinical assessment based on the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) criteria or short Confusion Assessment Method (short CAM) to confirm the diagnosis. In critical care or in the recovery room after surgery, CAM-ICU should be used. A healthcare professional who is trained and competent in the diagnosis of delirium should carry out the assessment. If there is difficulty distinguishing between the diagnoses of delirium, dementia or delirium superimposed on dementia, treat for delirium first. **[1.5.1]**

Ensure that the diagnosis of delirium is documented both in the person's hospital record and in their primary care health record. **[1.5.2]**

Treating delirium

Initial management

In people diagnosed with delirium, identify and manage the possible underlying cause or combination of causes. **[1.6.1]**

Ensure effective communication and reorientation (for example, explaining where the person is, who they are, and what your role is) and provide reassurance for people diagnosed with delirium. Consider involving family, friends and carers to help with this. Provide a suitable care environment (see recommendation 1.3.1). **[1.6.2]**

Distressed people

If a person with delirium is distressed or considered a risk to themselves or others, first use verbal and non-verbal techniques to de-escalate the situation. For more information on de-escalation techniques, see 'Violence' (NICE clinical guideline 25). Distress may be less evident in people with hypoactive delirium, who can still become distressed by, for example, psychotic symptoms. **[1.6.3]**

If a person with delirium is distressed or considered a risk to themselves or others and verbal and non-verbal de-escalation techniques are ineffective or inappropriate, consider giving short-term (usually for 1 week or less) haloperidol[8] or olanzapine[8]. Start at the lowest clinically appropriate dose and titrate cautiously according to symptoms. **[1.6.4]**

Use antipsychotic drugs with caution or not at all for people with conditions such as Parkinson's disease or dementia with Lewy-bodies.[9] **[1.6.5]**

If delirium does not resolve

For people in whom delirium does not resolve:

- Re-evaluate for underlying causes.
- Follow up and assess for possible dementia[10]. **[1.6.6]**

Information and support

Offer information to people who are at risk of delirium or who have delirium, and their family and/or carers, which:

- informs them that delirium is common and usually temporary
- describes people's experience of delirium

---

[8] Haloperidol and olanzapine do not have UK marketing authorisation for this indication.

[9] For more information on the use of antipsychotics for these conditions, see'Parkinson's disease' (NICE clinical guideline 35) and'Dementia' (NICE clinical guideline 42).

[10] For more information on dementia, see 'Dementia' (NICE clinical guideline 42).

- encourages people at risk and their families and/or carers to tell their healthcare team about any sudden changes or fluctuations in behaviour

- encourages the person who has had delirium to share their experience of delirium with the healthcare professional during recovery.

- advises the person of any support groups. **[1.7.1]**

Ensure that information provided meets the cultural, cognitive and language needs of the person. **[1.7.2]**

5 Epidemiology

CLINICAL QUESTION: What is the prevalence of delirium in different hospital settings and in long-term care?

5.1 Introduction

Delirium is a common clinical syndrome that can be found throughout the healthcare system. In order to understand more fully the clinical burden and associated health economic implications of delirium, it is necessary to first understand the epidemiology in terms of the occurrence of delirium within individual healthcare settings.

Operationalised diagnostic criteria for delirium have been formulated in the Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association 1980; American Psychiatric Association 1987; American Psychiatric Association 1994) (DSM III, DSM III-R and DSM-IV), and in the International Classification of Diseases (10th Edition) (World Health Organisation 1992) (ICD-10). There is good diagnostic agreement between DSM-IV and its predecessors, with DSM-IV identifying all patients diagnosed with delirium by DSM III and DSM III-R in one prospective cohort study of elderly hospital patients and nursing home residents (Laurila 2004, and section 6.6).

There is a notable disparity between the DSM and ICD-10 criteria for the diagnosis of delirium. The DSM-IV criteria are more inclusive in terms of diagnosis of delirium, with ICD-10 being relatively restrictive. In a cohort of elderly medical hospital patients and nursing home residents (mean age 88.4 years), 24.9% met the diagnostic criteria of DSM-IV, whilst only 10.1% of the same cohort were diagnosed with delirium when the diagnostic criteria of ICD-10 were applied (Laurila 2004). A comparison of the DSM-IV and ICD-10 criteria (table 5.1) reveals the ICD-10 criteria to include additional requirements for the diagnosis of delirium. In addition, the Laurila study (Laurila 2004) informs us that three cohorts were identified, those identified by DSM alone, ICD10 alone and both, and suggests that people who are identified using the ICD-10 criteria are different to the people identified using DSM. The stricter inclusion criteria and additional diagnostic requirements of ICD-10 have an associated impact on case detection and identify a cohort of patients who are more frequently dependent for care needs and more likely to be resident in the long-term care setting (Laurila 2004). Therefore we used the DSM-IV criteria as being the standard operational definition for delirium.

Table 5.1: DSM-IV and ICD-10 Diagnostic Criteria (American Psychiatric Association 1994; World Health Organisation 1992)

DSM-IV Diagnostic Criteria (American Psychiatric Association, 1994) **In order to be diagnosed with delirium, as a consequence of a general medical condition a patient must show all of the four features listed below:**	**ICD-10 Diagnostic Criteria (World Health Organisation, 1992)** **For a definite diagnosis, symptoms, mild or severe, should be present in each one of the following (five) areas:**
1. A disturbance of consciousness (i.e. reduced clarity of awareness of the environment) is evident, with reduced ability to focus, sustain or shift attention	a) Impairment of consciousness and attention (on a continuum from clouding to coma; reduced ability to direct, focus, sustain, and shift attention)
2. There is a change in cognition (such as memory deficit, disorientation, language disturbance) or the development of a perceptual disturbance that is not better accounted for by a pre-existing or evolving dementia.	b) Global disturbance of cognition (perceptual distortions, illusions and hallucinations – most often visual; impairment of abstract thinking and comprehension, with or without transient delusions, but typically with some degree of incoherence; impairment of immediate recall and of recent memory but with relatively intact remote memory; disorientation for time as well as, in more severe cases, for place and person)
3. The disturbance develops over a short period of time (usually hours to days) and tends to fluctuate during the course of the day.	
4. There is evidence from the history, physical examination, or laboratory findings that the disturbance is caused by the direct physiological consequences of a general medical condition.	
	c) Psychomotor disturbances (hypo- or hyperactivity and unpredictable shifts from one to the other; increased reaction time; increased or decreased flow of speech; enhanced startle reaction)
	d) Disturbance of the sleep-wake cycle (insomnia or, in severe cases, total sleep loss or reversal of the sleep-wake cycle; daytime drowsiness; nocturnal worsening of symptoms; disturbing dreams or nightmares, which may continue as hallucinations after awakening)
	e) Emotional disturbances, e.g. depression, anxiety or fear, irritability, euphoria, apathy, or wondering perplexity.

5.2 Terminology

Confusion can exist between the epidemiological terms **prevalence** and **incidence.** Prevalence represents the number of existing cases at a single point in time. Incidence represents the number of new cases that develop within a cohort over a defined period of time. The term ‘occurrence’ has been proposed as an alternative when there is ambiguity or overlap between the measurement of prevalence and incidence (Porta 2008).

Prevalent delirium in hospital therefore defines the presence of delirium at the point of admission to hospital. Incident delirium in hospital represents the development of delirium after hospital admission.

This is an important distinction to make as incident (new) cases of delirium are more likely to be amenable to strategies aimed at preventing the onset of delirium. It is therefore of key importance to provide *a priori* definitions of prevalence, incidence and occurrence rates with regard to delirium. Where it is not possible to use these definitions because of healthcare setting, alternatives will be considered, for example in the surgical setting, in which the concept of pre- and post-operative delirium is likely to hold importance.

As the emergency department represents a healthcare setting in which patients spend a short period of time prior to admission to the hospital bed base or discharge home, the concept of point prevalence is most relevant in this setting and incidence/occurrence rates will not be measured.

Long-term care represents the permanent residence of an individual, rather than respite care on a temporary basis. The concepts of point prevalence (prevalence at a single point in time) and period incidence (cumulative incidence over a defined period of time) are likely to be relevant in the long-term care setting.

5.2.1 *A priori* definitions

These *a priori* definitions form the basis for the review of study data and subsequent data categorisation:

Prevalent delirium: *The presence of delirium within the first 24 hours of admission to a healthcare setting (or the duration of the preoperative period within the surgical cohort).*

Incident delirium: *The development of delirium subsequent to the first 24 hours of admission (24 hours postoperatively in surgical cohorts), measured at least daily, until discharge from hospital or death.*

Occurrence rate: *Where study data reveal overlap between the a priori definitions of prevalent and incident data, or where the a priori conditions are not met, the term 'occurrence rate' will be used.*

Total Delirium: *Where there is more than one measure of rate of delirium available (e.g. both prevalent and incident delirium), or where occurrence rate represents data collected from healthcare admission to discharge, a fourth term, total delirium, will be summated to reflect the occurrence of delirium throughout the duration of stay.*

5.3 Selection criteria

Types of study

Prospective cohort and cross-sectional studies were to be included. Epidemiological data derived from the control arm of randomised clinical trials and case-control studies could be considered if there was evidence of reasonable representativeness of the sample. Retrospective studies were to be excluded.

Patient population & healthcare setting

Selection criteria for the patient population are defined in the methods section. Settings included are hospital and long-term care. In much of the guideline, the hospital patient population has been considered as a whole, but it is clear that this population is diverse and heterogeneous. For this epidemiological review, each healthcare setting was to be considered separately and data were to be grouped according to individual healthcare settings.

Studies were preferred if they were conducted in the UK. However, studies were to be included regardless of the country in which they were conducted, although the representativeness was to be taken into consideration in the analysis.

The DSM-IV criteria for delirium were to be the desired operational definition. As set out in the introduction, there is consistency between cases of delirium identified with DSM-IV versus DSM III-R and DSM III. Studies using a case definition based on the DSM-IV, DSM III-R or DSM III criteria [or a diagnostic tool validated against DSM-IV, DSM III-R or DSM III e.g. Confusion Assessment Method (CAM), DRS] were therefore to be included. As set out in the introduction (section 5.1), there is a notable disparity between cases of delirium that are identified with application of ICD-10 as compared with DSM-IV. Consequent to this, studies using the ICD-10 criteria for delirium were excluded from the epidemiological review.

Hospital Episode Statistics (HES)

Locally generated clinical coding data is collated nationally in the Hospital Episode Statistics (HES) database, the national statistical data warehouse for the NHS. Clinical coding of data is used for clinical research, epidemiological mapping and health resource allocation. A bespoke HES dataset was generated in order to assess the agreement between the epidemiological profile of delirium as determined by prospective cohort data and clinical coding data collated by the HES database.

5.4 Description of studies

Description of included and excluded papers together with study design are reported in table 5.2.

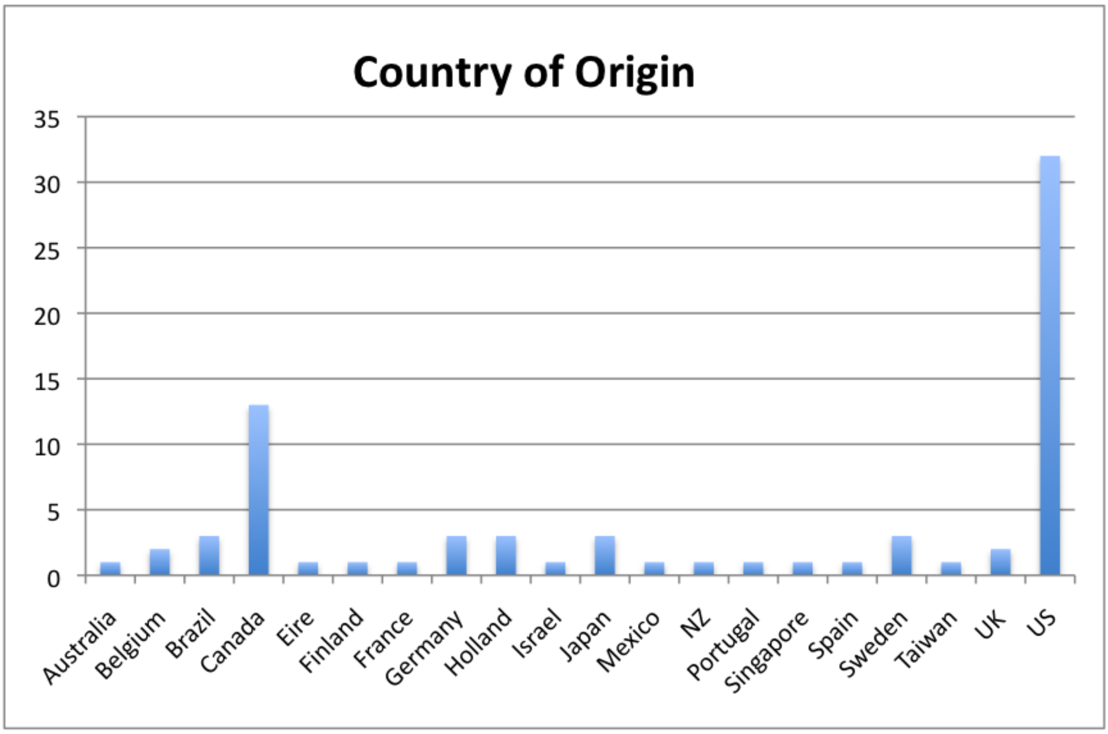
Table 5.2: study inclusion, exclusion and design

Papers	Details	Study
N=199 papers evaluated for inclusion		
N=124 papers excluded	Reasons for exclusion in Appendix G	
N=75 papers included	2 RCTs	Breitbart 1996; Cole 1994
	68 prospective cohort design	Adamis 2005; Angles 2008; Balas 2007; Benoit 2005; Bickel 2008; Brauer 2000; Caeiro 2004; Contin 2005; Dubois 2001; Edlund 1999; Edlund 2001; Edlund 2006; Ely 2001; Faezah 2008; Franco 2001; Furlaneto 2006; Galanakis 2001; Goldenberg 2006; Greene 2009; Hamann 2005; Henon 1999; Holden 2008; Holmes 2000; Inouye 1998; Inouye 1998; Inouye 1999; Jones 2006; Kagansky; Kawaguchi 2006; Koebrugge 2009; Koster 2008; Leslie 2005; Lin 2004; Marcantonio 1994; Martin 2000; McAlpine 2008; McCusker 2003; McNicoll 2003; Milbrandt 2004; Milisen 2001; Morrison 2003; Naughton 1995; Naughton 2005; O'Keefe 1999; Ouimet 2007; Pandharipande 2008; Patten 1997; Peterson 2006; Pisani 2006; Ramirez-Bermudez 2006; Roberts 2005; Robinson 2008; Robinson 2009; Rockwood 1999; Rolfson 1999; Rudolph 2005; Rudolph 2006; Rudolph 2007; Santana Santos 2005; Santos 2004; Sasajima 2000; Thomason 2005; Uldall 2000; van der Mast 1999; Van Rompaey 2009; Yoshimura 2004
	5 cross-sectional design	Elie 2000; Han 2009; Lewis 1995; Naughton 1995; Pitkala 2005

Eleven studies had fewer than 100 participants (Adamis 2005; Angles 2008; Edlund 2009; Goldenberg 2006; Koebrugge 2009; Milisen 2001; Robinson 2008; Rolfson 1999; Rudolph 2005; Rudolph 2006; Santana Santos 2005); 11 studies had more than 500 participants (Brauer 2000; Holmes 2000; Inouye 2008; Leslie 2005; Marcantonio 1994; McCusker 2003; Morrison 2003; Ouimet 2007; Peterson 2006; Rudolph 2007; Van Rompaey 2009) and the remaining 50 studies had between 100 and 500 participants.

The majority of included studies were of North American origin (figure 5.1), with only two studies based in the UK health service setting(Adamis 2005; Holmes 2000).

Figure 5.1: study by country of origin



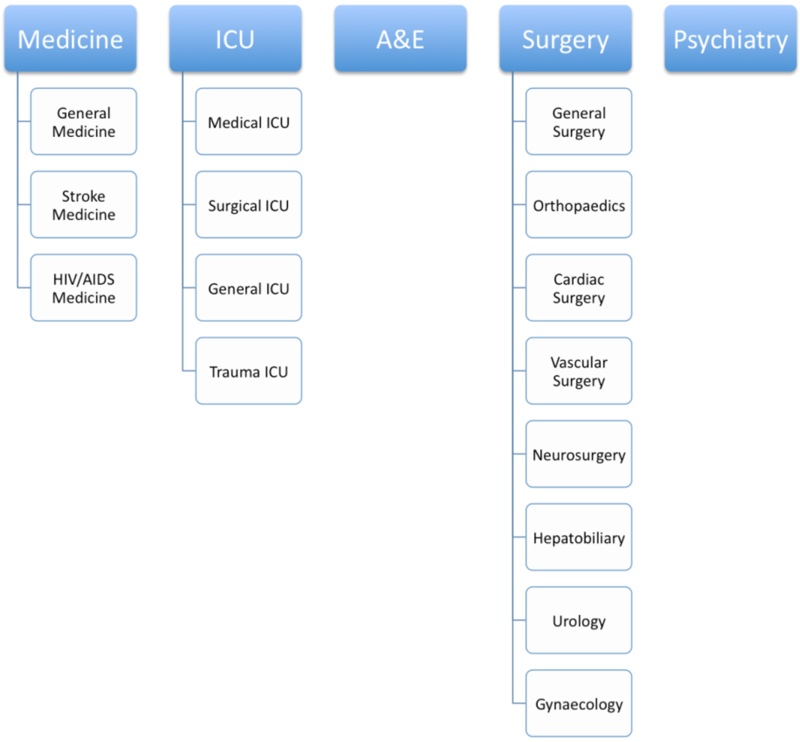
Thirty-eight studies selected adult patients with age cut-off points (Adamis 2005; Balas 2007; Bickel 2008; Brauer 2000; Breitbart 1996; Cole 1994; Edlund 2001; Edlund 2006; Elie 2000; Faezah 2008; Franco 2001; Furlaneto 2006; Galanakis 2001; Goldenberg 2006; Greene 2009; Han 2009; Henon 1999; Holden 2008; Holmes 2000; Inouye 1998; Inouye 1998; Inouye 1999; Jones 2006; Kagansky 2004; Koebrugge 2009; Leslie 2005; Lewis 1995; Marcantonio 1994; Martin 2000; McAlpine 2008; McNicoll 2003; Naughton 1995; Naughton 2005; Pisani 2006; Pitkala 2005; Rockwood 1999; Santos 2004; Santana Santos 2005). One study selected patients above the age of 40 years, three those above the 50 years, six selected patients above 60 years, 17 above 65 years, eight above 70 years and three studies selected patients above the age of 75 years.

Mean patient age varied between healthcare settings, with a higher mean age of study participants noted in the general medicine and long-term care cohorts (see Appendix D). A younger mean age of study participants was notable in the ICU, HIV/AIDS medicine and psychiatry settings.

Healthcare Setting

Studies were first assessed and grouped according to healthcare setting (Figure 5.2).

Figure 5.2: Hospital study populations grouped by healthcare setting



Where applicable, study populations were further categorised into, for example, acute and elective surgical patient groups. The long-term care setting was considered separately.

Both the ICU and acute stroke unit settings represent a form of enhanced specialist care within standard/usual care pathways. Thus, patients with ongoing delirium episodes may be admitted from the inpatient bed base to the ICU/acute stroke unit and therefore the occurrence rate can be a useful record of delirium rate for these specific healthcare settings. This model of ICU/acute stroke unit care is commonplace within the UK healthcare system.

5.5 Methodological quality of studies

The study cohort as a whole was assessed for representativeness on the grounds of the inclusion and exclusion criteria defined in each individual study. Inclusion and exclusion criteria were broadly similar between studies in each healthcare setting. Three studies (Andrew 2006; Edelstein 2004; Kakuma 2003) stated exclusion criteria showing that the study cohort was not representative of the population for that setting (see Appendix E). This is an important consideration for this epidemiology review, and these studies were therefore not analysed further.

One study (Andrew 2006) was in a long-term care setting whereby people with dementia were excluded from the cohort.
One study (Edelstein 2004) was in a hip fracture setting whereby only ambulatory home dwelling people were included in the cohort.
One study (Kakuma 2003) was in an emergency department setting whereby people presenting from long-term care were excluded from the participant cohort.

Fourteen studies listed dementia as an exclusion criterion (Andrew 2006; Bickel 2008; Contin 2005; Koebrugge 2009; Lin 2004; Roberts 2005; Rudolph 2007) or severe dementia (Franco 2001; Galanakis 2001; Han 2009; Kagansky 2004; Leslie 2005; Martin 2000; McNicoll 2003). However, as many of these studies were in the surgical and ICU setting, it was felt that the exclusion of people with dementia in these studies would not necessarily affect the representativeness of the study cohort.

As set out earlier, studies using the DSM-IV, DSM III-R or DSM III criteria (or a diagnostic tool validated against DSM-IV, DSM III-R or DSM III) were considered for inclusion. As delirium may often be present at admission and may be present for a short period of time with a tendency to fluctuate, included studies were appraised for quality on the basis of (1) an initial assessment for delirium within the first 24 hours of admission (post admission, preoperative period in the surgical studies) and (2) the frequency of subsequent assessments for delirium. Included studies were also appraised on the basis of sample size. These three criteria form the overall basis of the methodological quality assessment (Appendix E).

The relative importance of each quality criterion varies according to the type of epidemiological measurement. For example, prevalent delirium represents delirium within the first 24 hours of admission (preoperative period in the surgical cohort). With regard to this measure, the study size is therefore the key index. With regard to occurrence rate, the frequency of measurement of delirium and the study duration are potentially of greater importance.

Therefore, where studies recorded more than one measure of delirium (e.g. both prevalent delirium and occurrence rates), these were given separate quality assessments (Appendix E).

The studies were pragmatically and qualitatively grouped into high, medium and low quality on the basis of the quality criteria (Appendix E). Studies in which the sample size was small, in which the assessment of delirium was notably infrequent and/or the overall study length was short compared to the expected length of healthcare stay were considered to be at high risk of bias if a combination of these factors were present. Studies in which the methodology was unclear were also considered to lead to risk of bias. There was significant heterogeneity noted in frequency of assessment of delirium across all studies.

On the basis of these factors, four studies (Edlund 1999; Rudolph 2005; Santana Santos 2005; Van Rompaey 2009) were excluded from the overall results summary as they were felt to give potential for bias. These studies are highlighted in blue and given in italics in the study summary tables (Appendix D).

5.6 Results

Full data are given in Appendix D. Sixteen studies reported incidence or prevalence in different healthcare settings. Sixty-one studies report occurrence of delirium. The meaning of occurrence varied between studies and is shown in Appendix E under 'frequency of assessment' for each study. Three studies reported data for more than one setting:

- Pitkala 2005: General medicine (prevalence 32.6%); long-term care (15.9%)

- Bickel 2008: Orthopaedics acute hip fracture (occurrence 41%); orthopaedics elective surgery (12.5%)

- Galanakis 2001: Orthopaedics acute hip fracture (occurrence 40.5%); orthopaedics elective surgery (14.7%)

Summary data are reported by healthcare setting (table 5.2); in many healthcare settings the number of studies available for inclusion was limited, and the number ranged from 1 to 17 across all settings. Where more than one study is included, the median and range are given.

5.6.1 Sensitivity analysis

A sensitivity analysis was performed whereby the studies qualitatively graded as low quality were excluded from the dataset (table 5.6 – end of chapter). Removal of low quality studies led to significant change in a small number of cumulative results. Where this was the case, the sensitivity analysis results are preferred and these are shown in table 5.3 with the full results in square brackets. Exclusion of one low quality study with a low occurrence rate in the medical ICU setting led to a significant increase in the median (range) values for the occurrence of delirium, from 70.9 (22.4 – 83.3) to 80 (48 – 83.3). Following the sensitivity analysis, there was a decrease in the median (range) occurrence rate of delirium in the cardiac surgery setting, from 32 (13.5 – 50) to 21 (13.5 – 33.6), and an increase for the acute hip fracture setting. There was no apparent change in the rates of delirium in other healthcare settings when low quality studies were excluded. Where the only studies in a particular healthcare setting were low quality, this is indicated in the table.

Table 5.3: summary data by healthcare setting. (Full results are shown in red)

Healthcare setting	No. of studies	Age, years (median, range)	Prevalence % (median, range)	Incidence % (median, range)	Occurrence Rate % (median, range)	Total delirium % (median, range)
General Medicine	16	79.5 (76.4 – 83.6)	21.4 (18 – 32.6)	15.2 (12.5 – 17.9)	22 (5.7 – 42) [22 (3– 42)]	25 (15 – 42) [23.7 (15 – 42]
Stroke Medicine	2	69.3 (63.6 – 75)	12	No data available	24.3	24.3
HIV/AIDS Medicine	2	40.6 (39.2 – 42)	No data available	No data available	12 (12 – 12)	12 (12 – 12)
Medical ICU	7	56 (52.5 – 76)	36.6	24.4	80 (48 – 83.3) [70.9 (22.4 – 83.3)]	70.9 (48 – 83.3)
Surgical ICU	4	64 (64 – 75)	No data available	No data available	43.5 (29.8 – 70)	36.9 (29.8 - 44) [43 (29.8 – 44)]
Trauma ICU	1	44	No data available	No data available	59 (low quality)	No data available
General ICU	3	62.5 (61 – 64)	No data available	No data available	31.8 (19 – 45)	38.4 (31.8 – 45)
Emergency Department	4	79.9 (79.7 – 80.1)	9.8 (9.6 – 11.1)	No data available	No data available	9.8 (9.6 – 11.1)
General Surgery	5	68 (64.6 – 68.9)	No data available	No data available	11.4 (9 – 24)	No data available
Orthopaedics (Acute Hip Fracture)	10	79.8 (73.8 – 82.5)	22 (16.5 – 29.7)	30.3 (12.5 – 48.1)	28.3 (9.5 – 41) [17.4 (9.5 – 41)]	35 (29 – 68.1) [44.8 (29 – 41.1)]
Orthopaedics (Elective)	3	74.4 (73.8 – 74.9)	No data available	No data available	13.6 (12.5 – 14.7) [14.7 (12.5 – 22)]	No data available
Orthopaedics (Spinal Surgery)	1	59.2	No data available	No data available	3.8	No data available
Cardiac Surgery	5	68.8 (63 – 70)	No data available	No data available	21 (13.5 – 33.6) [32 (13.5 – 50)]	No data available
Vascular Surgery	2	71.6	No data available	No data available	31.1 (29.1 – 33)	No data available
Neurosurgery	1	No data available	No data available	No data available	14.9	14.9
Hepatobiliary	1	No data available	No data available	No data available	17	No data available
Urology	1	71.9	No data available	No data available	7	No data available
Gynaecology	1	No data available	No data available	No data available	17.5 (low quality)	No data available
Psychiatry	1	35.5	No data available	No data available	2.8	No data available
Long-term care	1	No data available	No data available	No data available	15.9 (low quality)	No data available

5.6.2 UK Data

Two included studies gave data on rates of delirium in the UK healthcare setting. The first, a prospective cohort study in a general medical setting with a sample size of 940 (Adamis 2005), recorded an occurrence rate of delirium of 37.3%. The second, a larger prospective cohort study in an orthopaedic setting with a sample size of 731 (Holmes 2000), recorded an occurrence rate of delirium of 14.8% (this study was considered to be of low quality). The limited number of studies available in UK healthcare settings leaves significant uncertainty as to the actual rates of delirium within the UK healthcare system.

5.6.3 Hospital Episode Statistics (HES)

In order to compare the epidemiological data with national clinical coding data, a bespoke dataset was requested from HES. The dataset provided information on the 2006 – 2007 total number of Finished Consultant Episodes (FCEs) of delirium (ICD code F05, delirium not induced by alcohol and other psychoactive) thus reflecting the scope of the guideline. The data were subcategorised by specialty and age (table 5.4).

Primary diagnoses represent the first of up to 14 diagnoses in the HES dataset and provide the main reason as to why the patient was in hospital. Subsequent to the primary diagnosis are up to 13 secondary diagnoses that record other diagnoses related to the episode. The bespoke delirium F05 dataset included both primary and secondary coded diagnoses of delirium, hence capturing all episodes of delirium in the UK healthcare setting in 2006 – 2007. It is likely that one episode of delirium corresponds to one patient having delirium. In order to calculate incidence of delirium as a percentage, the total number of FCEs in 2006 – 2007 (again split by specialty) was also requested. The latter is the record of the primary diagnoses only, which approximates to the number of admissions to each specialty. Therefore the HES delirium percentage is a reasonable reflection of the total delirium rate.

The dataset was split by age. The HES dataset captures episodes between the ages of 15 – 44 years followed by age 45 – 64 years. In order to provide a dataset that was representative of the mean age and inclusion criteria of the study cohort populations and in order that non-adult data was not introduced into the dataset, data were extracted from the HES dataset with a lower age limit of 45 years.

Table 5.4: Delirium Finished Consultant Episodes and Total Episodes by Specialty

Main Specialty	Delirium FCEs	Total Specialty FCEs	Total Delirium Episode Rate %
General Medicine	4706	2034768	0.23
Geriatric Medicine	3474	583506	0.59

Main Specialty	Delirium FCEs	Total Specialty FCEs	Total Delirium Episode Rate %
Critical Care	15	10240	0.15
Accident and Emergency	262	267476	0.10
Trauma and orthopaedics	204	652304	0.03
General Surgery	179	1041513	0.02
Adult Mental Illness	121	39839	0.30

5.6.4 Epidemiology data compared with coded HES data

HES data are generated over the course of the hospital admission. As discussed above, the proportion of episodes of delirium is very similar to the total rate of delirium in the study summary tables (Appendix D). In order to assess the reliability of the HES data, table 5.5 shows both the HES data and the appropriate median total delirium rate (from the sensitivity analyses) as reported by the epidemiological research studies and where total delirium rate was available.

Table 5.5: Comparison of Median Total Delirium Rates with HES Total Delirium Episode Rates

Main Specialty (number of included research studies)	Median (Range) Total Delirium Rate (Epidemiology Data) %	Total Delirium Episode Rate (HES data) %
General and Geriatric Medicine (6)	23.6 (15 – 42)	0.31
Critical Care (11)	48 (29.8 –83.3)	0.15
Accident and Emergency (4)	9.8 (9.6 – 11.1)	0.10
Trauma and orthopaedics (4)	44.8 (29 – 68.1)	0.03

There is a clear and significant disparity between the expected total delirium rates from a prospective cohort of patients admitted to hospital or long-term care as compared to the rates of delirium extracted from HES coding data. Less than one percent of the expected cases of delirium are identified by the coding process. There are also differences in the relative numbers of patients in the various healthcare settings, e.g. trauma & orthopaedic surgery has a similar level of delirium compared with general medicine in the studies, but the HES data show a much lower level for orthopaedic surgery. We recognise that some of the people identified by DSMIV may have had vascular dementia or dementia Lewy bodies, but the proportion of these groups is likely to be small. Even if the pure delirium rate in the studies is only 10% of that reported, there would still be a considerable disparity between the delirium rates in the studies compared with the HES data.

5.7 Discussion

Accurate coding of clinical data relies on all of the following taking place: the recognition of the underlying diagnosis, recording of the diagnosis by a clinician in the medical notes and subsequent extraction of the correct diagnosis / diagnoses from the medical notes by clinical coders. It is possible that there is an attrition of delirium diagnoses at each of these three stages. Clinicians often fail to identify delirium in the hospital setting, with up to two thirds of cases of delirium remaining unrecognised (Inouye 1998). The 'terminological chaos' (Lindesay 1999) of delirium creates a situation in which a variety of terms are used to describe the diagnosis of delirium. If the correct diagnostic terminology for delirium is not used, clinical coders will be unable to extract accurate diagnostic data from the clinical record and hence there is the potential for considerable under-reporting of delirium at a national healthcare level.

Delirium is ubiquitous throughout the healthcare system, being particularly common in the critical care, hip fracture, vascular surgery, cardiac surgery and general medical patient populations. Delirium also appears to be common in the long-term care setting, with a point prevalence estimate of 15.9% when residents with dementia are included within the prospective cohort (Pitkala 2005). We note that this study was considered to be of low quality.

In many healthcare settings there are few studies and these studies are often of lower quality. There is therefore significant uncertainty present with regard to the true epidemiology of delirium in a large proportion of healthcare settings. In these healthcare settings further large prospective cohort studies of high methodological quality would help provide rigorous data informing the true epidemiology of delirium.

There is a paucity of prospective cohort studies of delirium in the UK healthcare environment, with the majority of epidemiological data originating from North America. There are potential differences between the structure and organisation of healthcare in the UK compared to North America that may limit between-system comparisons and there is consequent uncertainty regarding the true rates of delirium within the UK healthcare system.

There is a significant disparity between the expected rates of delirium from prospective epidemiological studies and the rates of delirium as recorded in the HES data set. National clinical coding is systematically failing to accurately record the considerable scale and consequent importance of delirium as a healthcare priority.

Table 5.6: Sensitivity analysis.

(low quality studies removed, amended data highlighted in bold with number of low quality studies removed)

Healthcare setting	No. of studies	Prevalence % (median, Range)	Incidence % (median, Range)	Occurrence Rate % (median, Range)	Total delirium % (median, range)
General Medicine	16	21.4 (18 - 32.6)	15.2 (12.5 - 17.9)	**22 (5.7 - 42) 4 removed**	**25 (15 - 42) 1 removed**
Stroke Medicine	2	12	No data available	24.3	24.3
HIV/AIDS Medicine	2	No data available	No data available	**12 (1 removed)**	**12 (1 removed_**
Medical ICU	7	36.6	24.4	**80 (48 - 83.3) 1 removed**	**70.9 (48 - 83.3)**
Surgical ICU	4	No data available	No data available	**44 (29.8 - 70) 1 removed**	**36.9 (29.8 - 44) 1 removed**
Trauma ICU	1	No data available	No data available	**59 (low quality)**	No data available
General ICU	3	No data available	No data available	31.8(19 - 45)	38.4 (31.8 - 45)
Emergency Department	4	9.8 (9.6 - 11.1)	No data available	No data available	9.8 (9.6 - 11.1)
General Surgery	5	No data available	No data available	**9 (9 - 11.4) 2 removed**	No data available
Orthopaedics (Acute Hip Fracture)	10	**23.1 (16.5 - 29.7) 1 removed**	**12.5 1 removed**	**28.3 (9.5 - 41) 4 removed**	**35 (29 - 41) 2 removed**
Orthopaedics (Elective)	3	No data available	No data available	**13.6 (12.5 - 14.7) 1 removed**	No data available
Orthopaedics (Spinal Surgery)	1	No data available	No data available	3.8	No data available
Cardiac Surgery	5	No data available	No data available	**21 (13.5 - 33.6) 2 removed**	No data available
Vascular Surgery	2	No data available	No data available	**29.1 (1 removed)**	No data available
Neurosurgery	1	No data available	No data available	14.9	14.9
Hepatobiliary	1	No data available	No data available	**17**	No data available
Urology	1	No data available	No data available	7	No data available
Gynaecology	1	No data available	No data available	**17.5 (low quality)**	No data available
Psychiatry	1	No data available	No data available	2.8	No data available
Long-term care	1	**15.9 (low quality)**	No data available	**15.9 (low quality)**	No data available

5.8 Health economic evidence

No relevant health economic papers were identified.

5.9 From evidence to recommendations

The GDG noted from the epidemiological review, that there is widespread occurrence of delirium throughout the healthcare system but it was poorly reported in the UK. The GDG wished to reinforce the importance of accurately recording delirium by making a recommendation on coding (recommendation 1.5.2). In addition, people recovering from delirium may not receive adequate follow up care because of poor communication between hospitals and GPs, and hospitals and long-term care facilities. The GDG emphasised in the recommendation that delirium should be recorded in both the hospital and primary care health records.

The GDG observed that healthcare professionals were often unaware of the possibility that delirium might or has occured. The GDG thought that the slogan, **"Think delirium"** summarised their rationale, and incorporated this into a prominent statement at the beginning of the list of recommendations (see chapter 4 and section 9.7 of the consequences review).

The GDG made a future research recommendation (FRR) about recording delirium. This was informed by the multicomponent review showing that staff education may increase the awareness of delirium. This future research recommendation can be found in section 10.25.3 and Appendix H.

5.10 Recommendations

Ensure that the diagnosis of delirium is documented both in the person's hospital record and in their primary care health record. **[1.5.2]**

6 Diagnosis and accuracy of diagnostic tests

CLINICAL QUESTIONS:

What are the symptoms that indicate a person may have delirium?

What is the diagnostic accuracy of practical diagnostic tests compared with the reference standard DSM IV, to identify delirium in people in hospital and long-term care settings?

What are the diagnostic criteria that must be fulfilled to identify that a person has delirium?

6.1 Clinical Introduction

Delirium is common but is frequently unrecognised by doctors and nurses despite the fact that it can be life-threatening and lead to serious preventable complications. Unfortunately there is no simple quick test for delirium comparable to the ECG or Troponin test in myocardial infarction. The reference standard for diagnosis is a careful clinical assessment using the DSM-IV criteria at the bedside but this takes time and needs clinical expertise. There are however many screening tests available and these are reviewed in this section. Clinical suspicion should be high in any patient with a sudden change of behaviour or mental state especially in older patients with dementia, severe illness or fracture neck of femur. Early identification of patients with delirium and patients at increased risk is an essential first step in improving the management and outcome for this serious condition. It is therefore important that clinicians and support staff who are involved in the care of people at risk of delirium become familiar with the clinical indicators and symptoms that suggest the onset of delirium.

6.1.1 Primary objective of the review

To determine the accuracy of various diagnostic tests in diagnosing delirium in patients in hospital and long-term care.

6.1.2 Inclusion criteria

Patients: Adult patients in hospital; studies were stratified by setting (hospitals, long-term care and ICU).

Prior tests: No prior tests were undertaken

The target condition: Delirium

6.1.3 The index test and who executes the test

The GDG identified the index tests and the personnel who should undertake it for hospital, long-term care and ICU settings;

- Hospital and long-term care:
    - Abbreviated Mental test (AMT); any personnel can do this;
    - Clock-drawing test; can be used by untrained nurses or volunteers;
    - Confusion Assessment Method [long version] (CAM); trained healthcare professionals;
    - Confusion Assessment Method [short version] (CAM); trained healthcare professionals;
    - DRS-R-98; trained healthcare professional;
    - Mini Mental State Examination (MMSE) or other cognitive assessment instrument;
- ICU:
    - CAM-ICU and RASS (together); trained healthcare professional.

The reference standard

DSM-IV or ICD-10 applied by trained specialists

Sensitivity analyses

Sensitivity analyses were carried out to address QUADAS quality items

Subgroup analyses

For this review, we stratified the data according to the setting (hospital, ICU, long-term care), and considered the following subgroups in order to investigate heterogeneity

- ethnicity
- whether English is the first language
- writing ability
- patients with and without dementia/cognitive impairment

6.2 Characteristics of included studies

Details on included and excluded papers together with study design are reported in table 6.1

Table 6.1: study inclusion, exclusion and design

Papers	Comments	Study
N= 34 evaluated for inclusion		
N= 15 excluded	Reasons for exclusion are reported in Appendix G.	
N= 1 identified in update searches	One study was not presented in the results section as it would not substantially add to the evidence.	Vreeswijk 2009
N= 20 reports of 18 studies were included	Study designs 19 reports of 17 studies cross-sectional design .	Andrew 2009; Ely 2001; Ely 2001b; Fabbri 2001; Gonzalez 2004; Hestermann 2009; Laurila 2002; Laurila 2003; Laurila 2004; Lin 2004; Monette 2001; Ni Chonchubhair 1995; O'Keeffe 2005; Pompei 1995; Radtke 2008; Rockwood 1994; Rolfson 1999b; Yates 2009; Zou 1998.
	1 case-control	Cole 2003

The Cole (2003) study, reported a secondary analysis of data collected in what the authors reported as RCT of management of delirium and a prospective study of prognosis of delirium which included non delirious patients [references were not provided for either study in the text]. This study appeared to be a case-control study; one set of patients were included if they had a score of 3 or more on the Short Portable Mental Status Questionnaire (SPMSQ) or if their nursing notes indicated symptoms of delirium and who met the DSM IIIR criteria for delirium. The other set of included patients were people free of delirium, selected following screening for delirium; the study reported that the selection of non delirious patients in the study took into account the patients’ age and initial cognitive impairment status (SPSMQ score <3).

One study (Laurila 2002) may have included some of the same patients as those included in the Laurila (2003) study. The study enrolment period or the time period when assessments were carried out was not reported in the Laurila (2002) study. However, as the setting was limited to hospitals only in the 2002 study and as the other study (Laurila 2004) included hospital and long-term care setting, the results are reported separately.

Information on study sizes and geographical location are described in table 6.2

Table 6.2: Study characteristics

Study	Size (N)	Geographical location
Andrew 2009	145	Canada
Cole 2003	322	Canada
Ely 2001	96	USA
Ely 2001b	38	USA
Fabbri 2001	100	Brazil
Gonzalez 2004	153	Spain
Hestermann 2009	39	Germany
Laurila 2002	81	Finland
Laurila 2003	425	Finland
Lin 2004	109	China
Monette 2001	110	Canada
Ni Chonchubhair 1995	100	UK
O'Keeffe 2005	165	Ireland
Pompei 1995	1168	USA
Radtke 2008	154	Germany
Rockwood 1994	434	Canada
Rolfson 1999b	71	Canada
Yates 2009	62	UK
Zou 1998	87	Canada

The studies were conducted in different settings:

- Fifteen studies were carried out in hospital (Andrew 2009; Cole 2003; Fabbri 2001; Gonzalez 2004; Hestermann 2009; Laurila 2002; Monette 2001; Ni Chonchubhair 1995; O'Keeffe 2005; Pompei 1995; Radtke 2008; Rockwood 1994; Rolfson 1999b; Yates 2009; Zou 1998);
    - The Andrew (2009) study included 73% [106/145] inpatients and the remainder were outpatients; 15/39 of the outpatients (10% overall) were seen at home
    - The Gonzalez (2004) study reported excluding patients in psychiatric wards.

- Three studies were conducted in an ICU setting (Ely 2001; Ely 2001b; Lin 2004);

- One study was conducted in both hospital and long-term care settings (Laurila 2003).

Two studies were carried out in the UK (Ni Chonchubhair 1995; Yates 2009) and the rest were conducted in: Ireland (O'Keeffe 2005); the USA (Ely 2001; Ely 2001b; Pompei 1995); Canada (Andrew 2009; Cole 2003; Monette 2001; Rockwood 1994; Rolfson 1999b; Zou 1998); Finland (Laurila 2002; Laurila 2003); Germany (Hestermann 2009; Radtke 2008); Spain (Gonzalez 2004); Brazil (Fabbri 2001); and China (Lin 2004).

6.2.1 Population

The inclusion and exclusion criteria for each of the studies are shown in Appendices D and G.

Rates of delirium ranged from 14% (Radtke 2008) to 64% (Zou 1998) in the hospital setting; 86% (Ely 2001; Ely 2001b) in the ICU setting; and 25% (Laurila 2003) in the mixed setting (hospital and nursing home wards).

Where reported, the mean age of the participants in the studies was mostly above 65 years but varied as follows:

- mean age above 65 years (Andrew 2009; Cole 2003; Fabbri 2001; Gonzalez 2004; Hestermann 2009; Inouye 2005; Laurila 2003; Lin 2004; Monette 2001; Ni Chonchubhair 1995; O'Keeffe 2005; Pompei 1995; Rolfson 1999b; Yates 2009; Zou 1998)

- mean age over 80 years (Andrew 2009; Cole 2003; Hestermann 2009; Laurila 2003; Zou 1998)

- mean age below 65 years (Ely 2001; Ely 2001b; Radtke 2008)

Eight studies had a lower limit to age for inclusion: the Monette (2001) study reported that patients were eligible for enrolment if their age was 66 years and over; five studies (Gonzalez 2004; Ni Chonchubhair 1995; O'Keeffe 2005; Pompei 1995; Zou 1999) included patients over 65 years; and two studies (Laurila 2002; Laurila 2003) excluded patients younger than 70 years.

The studies varied in the proportion of patients with dementia/cognitive impairment:

- Patients with dementia were excluded in one study (Lin 2004);

- The Ely (2001b) study reported patients with a history of severe dementia were excluded, however, patients with suspected dementia (29%) were identified following enrolment;

- One study (Ely 2001: 12.5%) reported that less than 20% of the patients had suspected dementia;

- Five studies (Andrew 2009: 40%; Cole 2003: 29%; Gonzalez 2004: 50%; O'Keeffe 1997: 22%; Pompei 1995: 21%) reported between 20 and 50% of the patients had dementia;
- Three studies (Hestermann 2009: 84.6% ; Laurila 2003: 64%; Monette 2001: 53%) reported over 50% of the patients had dementia;
- One study (Yates 2009) reported the mean MMSE scores for delirium and non delirium groups (4.64 versus 14.94; p=0.003); the scores indicate that the included patients in this study were likely to be severely cognitively impaired.
- Four studies did not report dementia status (Fabbri 2001; Ni Chonchubhair 1995; Radtke 2008; Zou 1998).
- One study (Rolfson 1999b) reported that patients were 'highly selected with a low proportion of dementia'. Patients were undergoing coronary artery bypass graft surgery.

The studies varied in their inclusion of non-English speaking people. None of the studies reported if English was the first language. Five studies (Ely 2001; Ely 2001b; Inouye 2005; Pompei 1995; Rolfson 2005) reported excluding patients who did not speak English; two studies (Cole 2003; Monette 2001) reported excluding patients who did not speak English or French and one study (Radtke 2008) conducted in Germany reported excluding patients who did not speak the local language. Four studies reported the validation of the translated CAM instrument into: Portuguese (Fabbri 2001); Chinese (Lin 2004); Spanish (Gonzalez 2004); German (Hestermann 2009). One study (Laurila 2002) reported using a previously validated, Finnish version of the CAM instrument. For the translation studies we have assumed English was not the first language.

Ethnicity was reported in six studies (Ely 2001; Ely 2001b; Fabbri 2001; Inouye 2005; O'Keeffe 2005; Pompei 1995); with three studies reporting the majority of the patients were white (Ely 2001; Ely 2001b; O'Keeffe 2005); European descent (Fabbri 2001), and one study (Pompei 1995) reporting that 29% of the patients were African-American.

One study (Fabbri 2001) reported that 32% of the patients included in the study were unable to read or write fluently.

6.2.2 Index tests

A range of index tests were described:

- Abbreviated Mental Test (AMT); serial test (comparison of day before surgery and 3 day postoperatively) (Ni Chonchubhair 1995);
    - A 10 item questionnaire (scale score range: 0 to 10, with a score less than 6 indicative of dementia);
- Confusion Assessment Method (CAM):

- CAM (short version: Laurila 2002; Monette 2001; Pompei 1995; Radtke 2008)
    - The CAM short version assesses on the following 3 criteria; acute onset and fluctuating course; inattention; *and* disorganised thinking or altered level of consciousness.
- CAM (long version: Cole 2003; Yates 2009; Zou 1998)
    - The CAM long version assesses on the following 10 criteria: acute onset, inattention, disorganised thinking, altered level of consciousness, disorientation, memory impairment, perceptual disturbances, psychomotor agitation, psychomotor retardation, and altered sleep-wake cycle
- CAM (type of version unclear: Rockwood 1994; Rolfson 1999b);
- CAM translations (Fabbri* 2001 [Portuguese]; Gonzalez* 2004 [Spanish]; Hestermann* 2009 [German]; Laurila* 2002 [Finnish]; (translations are indicated by an asterisk in the rest of this document)
    - Three studies reported a translation of the short version (Gonzalez* 2004; Hestermann* 2009; Laurila* 2002) and the other study (Fabbri 2002*) reported a translation of the long version.

- Confusion Assessment Method (ICU) (CAM-ICU) (Ely 2001; Ely 2001b);
    - The CAM-ICU assess on the presence or absence of the following features: acute onset or fluctuation course and inattention and either disorganised thinking or altered level of consciousness;
    - Both studies reported the Attention Screening Examinations (ASE) scores, with Ely (2001b) reporting that the ASE was used to assess the 'inattention' feature of CAM-ICU. The Ely (2001b) study reported that the Vigilance A Random Letter Test which is part of the ASE was performed selectively in visually impaired patients. The Ely (2001) study reported that patient's delirium status was assessed with RASS when they were alert.
    - CAM-ICU translations: (Lin* 2004: Chinese)
        - The study reported patients were followed up daily with the Glasgow Coma Scale and the RASS for assessment of acute onset of mental status changes or fluctuation course.
- Clock-drawing test (Rolfson 1999b);
    - The clock-drawing test is an instrument used for screening of cognitive disorders. The test can be administered in three formats: in the free-drawn method, the patient is asked to draw a clock from memory; in the pre-drawn method, the patient is presented

with a circular contour and is expected to draw in the numbers on the clock face; or in the third method the patient is asked only to set the hands at a fixed time on a pre-drawn clock, complete with contour and numbers.

  - The Rolfson (1999b) study did not report the clock-drawing test format. The study reported a score of 6 or less was considered abnormal (range: 1 to 10, with 10 being error-free).

- Mini Mental State Examination (MMSE) (Rolfson 1999b; O'Keeffe 2005);

  - The MMSE is a test that is used to screen for cognitive impairment. (range 0 to 30);

  - Score of 23 or less was considered to be indicative of cognitive impairment (Rolfson 1999b);

  - Serial change in MMSE score; change in score between day 1 and day 6 (O'Keeffe 2005).

    - The study reported using a version of the MMSE that was previously adapted and validated for use in an Irish population.

- Delirium Index (DI) (Cole 2003);

  - An instrument designed to be used in conjunction with the MMSE, for the measurement of severity of symptoms of delirium based solely on observation of the patients. Patients are assessed on the following seven domains: inattention, disorganised thinking, altered level of consciousness, disorientation, memory impairment, perceptual disturbances, and motor disturbances. Score range from 0 to 21, with 21 points indicating maximum severity.

- DRS-R-98 (Andrew 2009);

  - The revised version of the DRS, allows assessment for both diagnosis of delirium and severity of delirium. This 16-item scale includes 3 'diagnostic items' (temporal onset, fluctuation and physical disorder) and 13 'severity symptoms' (attention, orientation, memory [short and long-term], sleep-wake cycle disturbances, perceptual disturbances and hallucinations, delusions, liability of affect, language, thought process abnormalities,visuospatial ability and motor agitation or retardation). Scores range from 0 to 44, and patients with a score of at least or over 17.75 points were screened as positive for delirium.

- Chart assessment (Rolfson 1999b);

- Documentation of delirium or its symptoms in the health records by physicians and nurses

- A retrospective review of the records by non study physicians and nurses were conducted for terms [including 'delirium' ,'confusion', 'acute confusion', 'toxic psychosis' and 'metabolic encephalopathy'] and themes [features of delirium, for e.g. acute onset, altered metal status, hallucinations, memory impairment] that suggested the recognition of delirium; Results for this index test were not considered as the GDG considered retrospective chart review to be an inadequate method of delirium assessment.

Most studies reported that the patients received only one index test; with the following exceptions:

- Cole 2003: CAM; DI; DSM-III; DSM-III-R; ICD-10;
- Laurila* 2003: DSM-III-R; DSM-III; ICD-10;
- Rolfson 1999b: CAM; MMSE; clock-drawing test; Chart assessment.

Three other studies (Andrew 2009; Pompei 1995; Radkte 2008) reported patients received other index tests that were not considered within this review (Andrew 2009: Delirium Symptom Interview (DSI); Pompei 1995: Digit Span Test, Vigilance 'A' Test, Clinical Assessment of Confusion (CAC); Radkte 2008: Delirium Detection Score (DDS); Nursing Delirium Screening Scale (Nu-DESC))

6.2.3 Reference standard (and index tests with which they were compared)

Although the GDG specified that the reference standard was to be DSM-IV or ICD-10, a number of studies compared tests only with the reference standard of DSM-III-R or DSM-III. The GDG ruled that this was acceptable, especially for the purpose of comparing different index tests.

The reference standards were carried out in different ways:

- DSM-IV

    - Five studies (Ely 2001; Ely 2001b; Gonzalez* 2004; Hestermann* 2009; Lin* 2004) reported the DSM-IV criteria for delirium was applied following clinical interview, family and/or nurse interviews, medical records and/or mental status records.

    - Two studies (Ely 2001; Ely 2001b) reported patients were assessed as either normal, delirious, stupor or comatose using DSM-IV or standardised definition of stupor and coma.

    - Two studies (Radtke 2008; Yates 2009) reported that the presence of delirium was determined using the DSM-IV criteria and did not provide further information.

    - One study (Laurila* 2002) reported the criteria addressed in the DSM-IV were operationalised in one questionnaire which also addressed the criteria in other classification systems (DSM-III-R, DSM-III, ICD-10).

- ICD-10
    - One study (Laurila* 2002) reported the criteria addressed in the ICD-10 were operationalised in one questionnaire which also addressed the criteria in other classification systems (DSM-IV, DSM-III-R, DSM-III).
- DSM III R
    - In the Cole (2003) study, a nurse gave the CAM to patients with a SPMSQ score ≥3 or delirium symptoms in the nursing notes; then the 10 CAM symptoms of delirium appeared to be used to determine the reference standard.
    - One study (Laurila* 2002) reported the criteria addressed in the DSM-III-R were operationalised in one questionnaire which addressed in other classification systems (DSM-IV, DSM-III, ICD-10).
- CAM and Clinician interview
    - One study (O'Keeffe 2005) had an experienced consultant geriatrician interview the patients using the CAM (short version)
- Consensus diagnosis
    - In the Zou (1998) study, the study team (comprised of two geriatric psychiatrists, research fellow and a nurse clinician) arrived at a consensus diagnosis using a nominal group method based on the following: results reported by the nurse for the CAM, SPSMQ, chart review; one assessment by a psychiatrist based on chart review and clinical examination; and independent assessment by each member of the team indicating the presence or absence of the five DSM-IV criteria for delirium (both 'definite' cases, requiring five criteria and 'probable' cases, requiring four of the five were included).

Where reported, the reference standard was mainly carried out by geriatricians or psychiatrists, with the exception of three studies (Pompei 1995: assessed by geriatricians and a geriatric nurse specialist; Yates 2009: junior medical doctor; Zou 1998; consensus diagnosis included a nurse's CAM findings).

Two studies compared different diagnostic criteria. In each of these comparisons the patients were given the same questionnaire/interview and the criteria were deduced from the symptoms reported:

- DSM-III-R versus DSM-IV (Cole 2003; Laurila* 2003) ; the test was carried out by a:

  - geriatrician in the hospital setting, and a nurse's interview and notes were used to arrive at an assessment for the long-term care setting (Laurila* 2003)
  - nurse (Cole 2003).
- DSM III versus DSM-IV (Laurila* 2003) ; the test was carried out by :
  - geriatrician in the hospital setting, and a nurse's interview and notes were used to arrive at an assessment for the long-term care setting (Laurila* 2003)
- ICD-10 versus DSM-IV (Laurila* 2003); the test carried out by:
  - geriatrician in the hospital setting and a nurse's interview and notes were used to arrive at an assessment for the long-term care setting
- DSM-III versus DSM-III-R (Cole 2003); the test was carried out by :
  - nurse (Cole 2003).
- ICD-10 versus DSM-III-R (Cole 2003) the test was carried out by :
  - nurse (Cole 2003).

The following tests were compared with the different reference standards:

- <u>Reference standard DSM-IV</u>
  - CAM: short version (Gonzalez* 2004; Hestermann* 2009; Laurila* 2002; Radtke 2008); the test was carried out by a:
    - geriatrician (Fabbri* 2001; Laurila* 2002);
    - general physician or psychiatrist (Gonzalez* 2004);
    - psycho gerontologist and a resident (Hestermann* 2009);
    - trained assessor (Radtke 2008).
  - CAM: long version (Fabbri* 2001; Yates 2009)
    - Geriatrician (Fabbri* 2001)
    - One of two junior medical doctors (Yates 2009)

- CAM-ICU (Ely 2001; Ely 2001b; Lin* 2004); the test was carried out by:
    - two nurses (Ely 2001; Ely 2001b) and an intensivist (Ely 2001b).
    - a research assistant (Lin* 2004).
- DRS-R-98 (Andrew 2009);
    - Test was carried out by either a geriatrician or a resident.

- <u>Reference standard ICD I0</u>
    - CAM: short version (Laurila* 2002);
        - Test was carried out by a geriatrician

- <u>Reference standard DSM IIIR</u>
    - CAM : short version (Laurila* 2002; Pompei 1995); the test was carried out by:
        - a geriatrician (Laurila* 2002)
        - a research assistant (Pompei 1995)
    - CAM: long version (Cole 2003; Rockwood 1994; Rolfson 1999b); the test was carried out by:
        - a nurse (Cole 2003)
    - CAM: type of version unclear (Rockwood 1994; Rolfson 1999b)
        - the study physician (Rockwood 1994)
        - both physician (first 41 patients) and trained research nurses (second 30 patients) (Rolfson 1999b).
    - MMSE (Rolfson 1999b);
        - Unclear whether a physician or nurse carried out the assessment.
    - Clock-drawing test (Rolfson 1999b);
        - Unclear whether a physician or nurse carried out the assessment.

  - Delirium Index (DI) (Cole 2003)
    - Test carried out by a trained research assistant

- Reference standard DSM III
  - AMT (Ni Chonchubhair 1995);
    - For the reference standard, the study reported that a single experienced physician examined patients using the Delirium Assessment Scale and determined which patients had delirium according to the DSMIII criteria
    - Unclear who carried out the test.
  - CAM: short version (Laurila* 2002);
    - Test carried out by a geriatrician.

- Reference standard Consensus diagnosis;
  - CAM: long version (Zou 1998);
    - Test carried out by a nurse.

Additionally, two studies compared different index tests, using CAM (carried out by a geriatrician) as a reference standard. These studies are included for completeness, but should be considered indirect comparisons for studies of diagnostic test accuracy

- Reference standard CAM (short version)
  - CAM test carried out by one of three lay interviewers. The team of lay interviewers included a nurse without prior research experience, a nurse with some experience as a research interviewer and one research assistant without a nursing degree but with experience as a research interviewer (Monette 2001);
- Reference standard CAM (long version) and Clinician interview
  - MMSE test carried out by one of two trained registrars in geriatric and general internal medicine (O’Keeffe 2005);

6.2.4 Outcomes

Methods of reporting outcomes varied:

- One study reported raw data to enable calculation of diagnostic test accuracy, and 2 x 2 tables were constructed (Laurila* 2003);
- In ten studies the raw data were back-calculated from accuracy measures (Andrew 2009; Cole 2003; Ely 2001; Gonzalez* 2004; Lin* 2004; O'Keeffe 2005; Pompei 1995; Radtke 2008; Rockwood 1994; Yates 2009);
- In six studies both the raw data and accuracy measures were reported (Fabbri* 2001; Laurila* 2002; Monette 2001; Ni Chonchubhair 1995; Rolfson 1999b; Zou 1998);
- In one study (Ely 2001b), the raw data were obtained by an estimation process in order to reproduce the reported accuracy parameters.

In the Rockwood (1994) study limited raw data was reported. We estimated the number of patients who were delirious and non delirious by assuming the 52 patients (who were referred to the study physician) were roughly equally spread between the two groups.

One study (Laurila* 2004), provided insufficient raw data and we were unable to calculate accuracy measures.

6.3 Methodological quality of included studies

The methodological quality was assessed (Appendix E) using QUADAS criteria.

Most of the studies used a reference standard that was likely to classify the target condition correctly. Two studies (Monette 2001: CAM assessment by geriatrician; O'Keeffe 1997: CAM and clinical interview) used the CAM as the reference standard. In one study (Andrew 2009) it was unclear who performed the assessment.

Generally the studies reported the availability of additional clinical data, for example MMSE scores or other measures indicative of cognitive impairment or dementia, medical records or notes from interviews with family/carers were available when patients were assessed.

Overall, most studies briefly reported the execution of the index test and reference standard, with the exception of four studies which provided detailed information on the tests and/or the method of assessments (Ely 2001; Ely 2001b; Gonzalez* 2004: index test; Laurila* 2002). One study (Radtke 2008) reported that patients were assessed only once in the recovery room and length of stay ranged between 22 minutes to 147 minutes.

None of the studies reported intermediate or uninterpretable results. Withdrawals (18%: 35/200) in one study (O'Keeffe 2005) were due to deaths, early discharge or error. Two studies reported missing data (Andrew 2009: 1%, values were replaced with the mid-range score; Pompei 1995: 0.9% missing data and were excluded from the analysis);

In addition to the above quality issues, the following studies were found to be at risk of bias on the following criteria:

- Spectrum bias (Andrew 2009; Cole 2003; Monette 2001; Radtke 2008; Rolfson1999b)
    - Following first stage CAM assessment by the nurse, patients were selected from those classified as having probable delirium and no delirium; the CAM negative group had a higher proportion of cognitively impaired people (Monette 2001)
    - 30% of the patients were outpatients, of whom 10% were assessed at home (Andrew 2009)
    - Case control study in which two groups of patients with and without delirium were selected (Cole 2003)
    - Patients were in the recovery room following general anaesthesia. The GDG considered the ordinary version of CAM to be inappropriate for this environment (Radtke 2008)
    - Patients were undergoing CABG surgery and had a low proportion with dementia (Rolfson 1999b)
- Disease progression bias (Andrew 2009; Inouye 2005; Ni Chonchubhair 1995; O'Keeffe 2005; Rockwood 1994; Rolfson 1999b; Yates 2009; Zou 1998)
    - The authors reported that the index and reference tests were not necessarily done on the same day, which given the fluctuating course of delirium, is a limitation. (Andrew 2009);
    - The study reported that reference standard assessment was within the same day (O'Keeffe 2005)
    - The study reported that the time between assessments varied between 30 min and 8 hours (Zou 1998)
    - Time period was not reported so the studies were downgraded for this quality criterion (Ni Chonchubhair 1995; Rockwood 1994; Rolfson 1999b; Yates 2009).
- Partial verification bias (Cole 2003; Pompei 1995)

  - Reference standard appeared to be given only to patients with SPMSQ score ≥3 or delirium symptoms in notes (Cole 2003)
  - Only the patients with an acute change in mental status (61%:263/432) were referred to clinician for reference standard assessment (Pompei 1995)

- Review bias (Andrew 2009; Cole 2003; Laurila* 2003; Monette 2001; Rockwood 1994; Rolfson 1999b; Yates 2009; Zou 1998)
  - Two studies used the same data for both the reference standard and index test and it was very likely that there was review bias (Cole 2003; Laurila* 2003)
  - One study included the index test as part of the reference standard; results for DSM-IV as a separate reference standard were not reported (Zou 1998)
  - One study had the index and reference tests carried out by the same person (Rockwood 1994)
  - One study may have had the index and reference tests carried out by the same person/people (Yates 2009)
  - It was unclear whether the index test was interpreted without the knowledge of the reference standard, as the nurse [conducting the index test] observed the geriatrician [reference standard] (Monette 2001)
  - In the Rolfson (1999b) study the CAM assessments were administered by a physician [41/71 patients] and a nurse administered the CAM for the remaining patients; the same physician assessed the reference standard (but the other tests were not carried out by the same people)
  - For the rest of the above studies it was unclear whether the reference standard was interpreted with the knowledge of the result of the index test so studies were downgraded for this quality criterion

- Incorporation bias (Cole 2003; Laurila* 2003; Zou 1998)
  - The index test [CAM administered by the nurse] was part of the reference standard [consensus diagnosis] (Zou 1998)
  - The index tests and reference tests were based on the same data (Cole 2003; Laurila* 2003)

Overall, nine studies were considered as potentially or at risk of bias (Andrew 2009; Cole 2003 (all comparisons); Laurila* 2003 (all comparisons); Monette 2001; Pompei 1995; Rockwood 1994; Rolfson 1999b (for CAM only); Yates 2009; Zou 1998). These studies were considered in sensitivity analyses.

6.4 Results – hospital setting

The purpose of the tests examined is to identify delirium, possibly to be used as a screening tool. The GDG stated that they were most interested in a test that had high sensitivity and would 'rule in' patients with delirium. We examined the sensitivity, specificity, positive likelihood ratio and the pre and post test probabilities.

6.4.1 Comparison of diagnostic criteria (table 6.3)

One low quality, case control study (Cole 2003) compared different diagnostic criteria; raw data were calculated from the accuracy measures.

DSM-III-R versus DSM-IV

One low quality, case control study (Cole 2003) compared DSM-III-R with DSM-IV using the same symptoms to determine both test results, and considered the effect on sensitivity and specificity in relation to criterion A from the DSM-III-R and the DSM-IV (inattention versus clouding of consciousness). The test showed moderate sensitivity: 79%; specificity: 100% when *either* inattention *or* clouding of consciousness criterion was used. However, when the required criterion was *both* inattention and clouding of consciousness, the sensitivity showed a slight improvement [82%], however, the specificity was compromised [63%] and similar results were reported [sensitivity: 81%; specificity: 63%] when only the clouding of consciousness was the required criterion (figure 6.1, Appendix K).

DSM III versus DSM-III-R

One low quality, case control study (Cole 2003) compared DSM-III with DSM-III-R and considered the effect on sensitivity and specificity in relation to criterion A (inattention versus clouding of consciousness). The test showed high sensitivity [96%] and specificity [91%] when *either* inattention *or* clouding of consciousness criterion was used. However, when the required criterion was *both* inattention and clouding of consciousness, the sensitivity was compromised [52%], however, the specificity slightly improved [96%] and similar results were reported [sensitivity: 52%; specificity: 96%] when only the clouding of consciousness was the required criterion (figure 6.2, Appendix K).

ICD-10 versus DSM-III-R

One low quality, case control study (Cole 2003) compared ICD-10 with DSM-III-R and considered the effect on sensitivity and specificity in relation to criterion A (inattention versus clouding of consciousness). The test showed moderate sensitivity: 61%; specificity: 91% when *either* inattention *or* clouding of consciousness criterion was used. However, when the required criterion was *both* inattention and clouding of consciousness, the sensitivity was low [36%], however, the specificity slightly improved [96%] and similar results were reported [sensitivity: 36%; specificity: 96%] when only the clouding of consciousness was the required criterion (figure 6.3, Appendix K).

The DSM-III-R compared with DSM-IV showed moderate sensitivity and a high positive predictive value (PPV) (which is the proportion of patients with a positive test who have the target condition) indicating the DSM-III-R is inclusive. Of the two diagnostic tests (DSMIII and ICD-10) compared with DSM-III-R, the ICD-10 was least inclusive.

Table 6.3: diagnostic test accuracy statistics for different reference standards

Study name	Comments	Test operator	Sensitivity	Specificity	PPV	LR+	Pre-test probability	Post-test probability
Cole 2003	DSM-III-R vs DSM-IV; criterion A: either inattention or clouding of consciousness	Nurse	79.23	100	100	NA	65.84	100
Cole 2003	DSM III vs DSM-III-R; criterion A: either inattention or clouding of consciousness	Nurse	96.4	90.9	92.1	10.67	6.83	43.9
Cole 2003	ICD10 vs DSM-III-R; criterion A: either inattention or clouding of consciousness	Nurse	60.71	90.92	87.9	6.68	52.17	87.9

CAM (short version) versus different diagnostic criteria

One moderate quality study (Laurila* 2002) compared the CAM index test (short version) with different reference standards. The CAM test, which is based on the DSM-III-R criteria, showed a moderate sensitivity (80% to 85%) and specificity (63.4% to 83.7%) against the reference standards. The CAM had the most concordance with the DSM-IV [sensitivity: 81.3% and specificity: 83.7%] and was the least concordant with the ICD-10 [sensitivity: 80% and specificity: 63.4%]; table 6.4.

Table 6.4: diagnostic test accuracy statistics for CAM for different reference standards

CAM index test (short version)	Study name	Comm-ents	Test operator	Sensitivity	Specificity	PPV	LR+	Pre-test probability	Post-test probability
DSM-IV	Laurila* 2002	CAM vs DSM-IV	Geriatrician	81.3	83.7	76.0	5.0	39.5	76.5
ICD-10	Laurila* 2002	CAM vs ICD-10	Geriatrician	80.0	63.4	24.0	2.2	12.3	23.5
DSM III-R	Laurila* 2002	CAM vs DSMIII-R	Geriatrician	81.0	71.7	50.0	2.9	25.9	50.0
DSM III	Laurila* 2002	CAM vs DSMIII	Geriatrician	85.0	72.1	50.0	3.1	24.7	50.0

Subgroup analysis by dementia or no dementia

The Cole (2003) study reported separately the accuracy measures for different diagnostic criteria in patients with and without dementia. Dementia was diagnosed with the IQCODE.

DSM-III-R versus DSM-IV

The DSM-III-R instrument (compared with DSM-IV) shows a slightly higher sensitivity in people with dementia [80%] than in people without dementia [range: 75%] when the criterion A is interpreted as either clouding of consciousness or inattention. A forest plot of sensitivity and specificity is shown in figure 6.4 (Appendix K), but we note that the study used both tests to interpret the same symptoms.

DSM-III versus DSM-III-R

The DSM-III instrument (compared with DSM III-R) shows a high sensitivity and the ability of the test to rule in those with delirium is high and this is the case whether the patients have dementia [sensitivity: 97%] or not [sensitivity: 95%]; figure 6.5 (Appendix K). The reported results are for criterion A being interpreted as either clouding of consciousness or inattention.

ICD10 versus DSM-III-R

The ICD-10 instrument (compared with DSM III-R) showed a fairly low sensitivity and this is the case for patients with dementia [sensitivity: 59%] or for patients without dementia [sensitivity: 68%]; figure 6.6. The reported results are for criterion A being interpreted as either clouding of consciousness or inattention.

6.4.2 Diagnostic test accuracy (DSM-IV as the reference standard)

Seven studies compared index tests with DSM-IV as the reference standard: four investigated CAM short version (Gonzalez* 2004; Hestermann*2009; Laurila* 2002; Radtke 2008); two studies investigated CAM long version [Fabbri* 2001; Yates 2009 (low)]; and one study investigated the DRS-R-98 [(Andrew 2009 (low)].

A forest plot of sensitivity and specificity is shown in figure 6.7. The GDG agreed that the CAM long version, which assessed for 10 symptoms (acute onset, inattention, disorganised thinking, altered level of consciousness, disorientation, memory impairment, perceptual disturbances, psychomotor agitation, psychomotor retardation) and the CAM short version, which assessed for 3 symptoms (acute onset, inattention, disorganised thinking or altered level of consciousness) of delirium, should be treated separately and these are reported as subgroups. The diagnostic test accuracy statistics are summarised in table 6.5.

DRS-R-98

One low quality study (Andrew 2009) assessed the DRS-R-98 with DSM-IV showed a moderate specificity and fairly low sensitivity [sensitivity: 56%; specificity: 82%]. The study included patients with dementia (40%), had a high proportion of inpatients (73%), with high comorbidity [mean comorbidity count 7.1 (SD 2.7)). The study also examined a sub-sample of patients with underlying dementia, which had a sensitivity of 59% and a specificity of 67%. The study reported that the assessors of the index test had varying expertise and did not have extensive training in the use of the instrument; the study showed a moderate inter-rater reliability (k=0.76).

The number of patients identified with the DRS-R-98 instrument as delirious have a small likelihood of being delirious [likelihood ratio: 3.17]. However, the results are based on one low quality study so some uncertainty exists on DRS-R-98 utility as a screening instrument for delirium.

CAM

Of the six studies [Fabbri* 2001; Gonzalez* 2004; Hestermann* 2009; Laurila* 2002; Radtke 2008; Yates 2009 (low)] comparing CAM, we note that four of these (Fabbri* 2001; Gonzalez* 2004; Hestermann* 2009; Laurila* 2002) used a foreign language version of the CAM: Portuguese, Spanish, German, and Finnish respectively. The Gonzalez* (2004) study reported that in order to further assess the onset and course of the mental status changes and to evaluate thinking and attention, items from the Spanish version of the MMSE were included in the interview – so this study was considered as an adaptation study.

Two of the studies (Fabbri* 2001; Hestermann* 2009) reported that the instrument was translated and back translated and in the other two studies (Gonzalez* 2004; Laurila* 2002) the final version of the instrument was based on expert panel consensus.
In all of the studies, the CAM was rated by a physician, with the exception of the Yates (2009) study, where a trained assessor administered the instrument (CAM long version).

For the CAM short version, the sensitivity ranged from 43% to 90% and the specificity from 84% to 100%. The positive predictive value ranged from: 76% to 100% and likelihood ratio ranged form: 5.0 to 28.5.

There was heterogeneity, particularly for sensitivity and some variation in the specificity. Heterogeneity was considered in terms of the following factors: language and type of patients. As noted earlier, assessment was carried out with a foreign language version of the CAM in three studies (Gonzalez* 2004; Hestermann* 2009; Laurila* 2002). We note that the Radtke (2008) study, conducted in Germany, reported that patients who did not speak the local language were excluded; however, it was unclear if the CAM instrument was a version translated into the local language.

In terms of type of patients included in the study, we note the Radtke (2008) study was the only study which included patients with a mean age below 65 years (mean [range]: 54.5 years [25.4 to 80.8]) and the study included patients who were in the recovery following general anaesthesia. The GDG considered the ordinary version of CAM to be inappropriate for this environment.

The type of patients included, the setting and the inappropriate measure for this setting may account for the low sensitivity [43%] observed in the Radtke (2008) study.

For the CAM long version, the sensitivity ranged from 91% to 94% and the specificity was 96%. We note the Yates (2009) study was low quality.

The CAM instrument when compared with DSM-IV as the reference standard, was able to detect delirium and the likelihood of patients having delirium when CAM had identified patients as being delirious is high.

Table 6.5: diagnostic test accuracy statistics for DSM-IV as the reference standard

DSM-IV	Study name	Comments	Test operator	Sensitivity	Specificity	PPV	LR+	Pre-test probability
CAM Long vers-ion	Fabbri* 2001	CAM [geriatrician] vs DSMIV [psychiatrist]	Geriatrician	94.1	96.4	84.0	26.0	17.0
	Yates 2009	CAM vs DSM-IV	Study physician	90.90	96.10	83.00	23.2	17.7
CAM Short vers-ion	Gonzalez * 2004	CAM vs DSMIV	General Physician or Psychiatrist	90.0	100.0	100.0	NA	24.4
	Hesterman n|* 2009	CAM [rater 1= psychogeront ologist] vs DSM-IV[consensus]	Psychologist / Gerontologist and Resident	76.9	96.2	91.0	20	33.3
	Hesterman n* 2009	CAM [rater2= internal resident in geriatric medicine] vs DSM-IV[consensus]	Psychologist/Gerontologist and Resident	76.9	100.0	100.0	NA	33.3
	Laurila* 2002	CAM vs DSM-IV	Geriatrician	81.3	83.7	76.0	5.0	39.5
	Radtke 2008	CAM vs DSM-IV	Trained assessor (trained by psychiatrist)	42.9	98.5	82.0	28.5	13.6
DRS-R-98	Andrew 2009	Index: DRS-R98 Ref: 'clinically diagnosed delirium'=DSMIV	Geriatrician/ Resident	56.40	82.20	66.00	3.2	37.9

6.4.3 Subgroup analyses by dementia or no dementia

Subgroup analyses for DRS-R-98 compared with DSM-IV

One low quality study (Andrew 2009) reported subgroup analyses for patients with and without dementia for the DRS-R-98 test compared with DSM-IV as reference standard.

Dementia was diagnosed with DSM-IV and the number of patients with dementia and underlying dementia with superimposed delirium was 58. The study showed low sensitivity and specificity, 59% and 67%, respectively (figure 6.8, Appendix K). We note that this study was considered low quality.

Subgroup analyses for CAM (short version) compared with DSM-IV

One moderate quality study (Gonzalez* 2004) reported the diagnostic accuracy measures for the CAM test (short version) compared with DSM-IV as reference in people with and without dementia. Dementia was diagnosed on the basis of DSM-IV criteria, medical records, MMSE rating, and interviews with relatives. The study did not provide the number of patients diagnosed with delirium for the subgroups so we were unable to back-calculate the raw data.

The Spanish translation of the CAM (short version) showed a slightly lower sensitivity in people with dementia [sensitivity: 87%] compared to people without dementia [sensitivity: 93%]; the specificity was similar for both groups [100%].

6.4.4 ICD-10 as reference standard

One moderate quality study (Laurila* 2002) compared CAM (short version) with ICD-10 as a reference standard. We note that in this study, four reference standards [DSM-IV, DSM-III-R, DSM-III, and ICD-10] were operationalised in one questionnaire. The index test was a previously validated foreign language [Finnish] version of the CAM, which was developed by consensus.

The forest plot showing the specificity and sensitivity is shown in figure 6.9. The CAM (short version) showed moderate sensitivity [80%] with the ICD-10 classification, however, the specificity was fairly low [63%].

Although the positive predictive value is 24%, the negative predictive value is 96% which indicates that a negative result on the CAM test is able to exclude delirium. The low positive likelihood ratio of 2.18 indicating that a patient identified with delirium using the CAM instrument for assessment is 2.18 more likely to be delirious than non delirious.

As shown earlier, the ICD-10 diagnostic criteria (compared with DSM-III-R), performs poorly in relation to specificity and may have some limitations as a reference standard.

6.4.5 DSM-III-R as the reference standard

Two studies compared CAM short version with DSM-III-R (Laurila* 2002; Pompei 1995 (low); one study compared CAM long version with DSM-III-R (Cole 2003 (low); and type of version was unclear in two studies (Rockwood 1994 (low); Rolfson 1999b (partly low)). One study (Rolfson 1999b) also gave the patients other index tests compared with DSM-III-R [MMSE; clock-drawing test] – the study quality was considered to be moderate for these tests.

A forest plot of sensitivity and specificity is shown in figure 6.10. Results for the CAM short and long versions are reported as subgroups. The diagnostic test accuracy statistics are summarised in table 6.6.

The low quality Cole (2003) study also reported classification of delirium by number of symptoms for the CAM and DI; this is reported separately under the section below entitled ‘within group comparison’. In Figures 6.11 and 6.12 (Appendix K), for the Cole (2003) study, the values for more than 6 symptoms and more than 4 symptoms are used respectively. We note that the same data were used for the CAM and reference standard, but a separate test was carried out for the DI, so the CAM results are likely to be more biased.

CAM

Two studies compared CAM short version with DSM-III-R (Laurila* 2002; Pompei 1995 (low); one study compared CAM long version with DSM-III-R (Cole 2003 (low); and type of version was unclear in two studies (Rockwood 1994 (low); Rolfson 1999b (partly low)).

The Cole (2003) study used the CAM (long version) to determine 10 symptoms which were used for the reference standard. The study reported the sensitivity and specificity (for more than 6 symptoms) for patients with dementia or without dementia. The sensitivity and the specificity was 98% and 76% for patients with dementia and 95% and 83% for patients without dementia. We note this was a case control study; therefore the sensitivity and specificity are likely to be overestimated.

The two studies (Laurila* 2002; Pompei 1995 (low)) comparing CAM short version with DSM-III-R showed sensitivity ranging from 46% to 81% and specificity ranging from 72% to 92%. A sensitivity analysis was carried out excluding the low quality studies. Considering the remaining study (Laurila* 2002), which was of moderate quality, the CAM showed an 81% sensitivity and 72% specificity compared with DSM-III-R. The positive predictive accuracy was 50% and the negative predictive value was 91%, indicating that a negative result on the CAM instrument will accurately exclude delirium. The likelihood ratio is 2.86, which suggests a not particularly strong test.

In two studies (Rockwood 1994 (low); Rolfson 1999b (low)) the type of version used was unclear. The Rolfson (1999) study reported that the CAM and reference standard were carried out by the same physician for 41 patients and by different assessors for the next 30 patients: for the latter, assessment was by nurses, and these results are considered to be low quality. The results are reported separately for the two groups.

The Rockwood (1994) study reported the sensitivity [64%] and specificity [93%]. However, there was insufficient information and we were unable to calculate the raw data from the reported accuracy measures, although a rough estimate was obtained by assuming the 52 patients were roughly equally spread between delirium positive and delirium negative; the study is not included in the forest plot.

Clock- drawing and MMSE tests

Both the MMSE and the clock-drawing test index tests were administered on the day prior to surgery and on the fourth postoperative day in the Rolfson (1999) study; results were reported for the latter time. The MMSE showed a low sensitivity, 35%, a small positive likelihood ratio of 1.9. It was unclear in the study how many patients had impaired communication which would not allow the

MMSE to be administered (albeit patients with coma before day 4 were excluded).

The clock-drawing test showed a very low sensitivity of 9%, and a positive likelihood ratio of 4.2. It was unclear whether patients had been assessed with impaired writing ability at baseline as the administration of this index test in such population would be limited.

Test comparison

Overall, the CAM performed better than the MMSE or the clock-drawing tests; although this is based on different studies and there was variation in the index and reference test assessors.

Table 6.6: index test compared with DSM-III-R

DSM-III-R	Study name	Comments	Test operator	Sensitivity	Specificity	PPV	LR+	Pre-test probability	Post-test probability
CAM Long Version	Cole 2003	CAM >6 symptoms vs DSM IIIR for patients with dementia	Nurse	97.7	75.0	84.0	4.0	57.7	84.5
	Cole 2003	CAM >6 symptoms vs DSM IIIR for patients without dementia	Nurse	95.0	83.3	79.0	5.7	40.0	79.2
CAM Short Version	Laurila* 2002	CAM vs DSMIII-R	Geriatrician	81.0	71.7	50.0	2.9	25.9	50.0
	Pompei 1995	CAM vs DSMIIIR without 4 patients for whom no results	Research Assistant	45.9	92.1	49.0	5.8	14.3	49.1
CAM type of version unclear	Rockwood 1994	CAM vs DSMIIIR raw data estimated based on sensitivity and specificity	Study physician	63.0	93.0	88.2	8.75	46.15	88.2
	Rolfson 1999b	CAM nurse	Nurse	12.5	100.0	100.0	NA	26.7	100.0
	Rolfson 1999b	CAM [physician] vs DSM III-R [geriatrician]	Physician	69.6	100.0	100.0	NA	32.4	100.0
MMSE	Rolfson 1999b	MMSE vs DSM III-R	Nurse/physician	34.8	81.2	47.0	1.9	32.4	47.0
Clock Drawing	Rolfson 1999b	Clock-drawing test vs DSM III-R	Nurse/physician	8.7	97.9	67.0	4.2	32.4	66.7

Subgroup analyses

One low quality study (Pompei 1995) reported subgroup analyses for patients (21%: 96/438) with impaired cognitive status on admission. Cognitive status was assessed with the MMSE (range 0 to 30); with varying cut-off points adjusted for education level (score less than 21 was indicative of cognitive impairment for those with less than a high school experience; score less than 23 points was indicative of cognitive impairment for those with high school experience; and score less than 24 points was indicative of cognitive impairment for those with college education).

The study showed moderate/low sensitivity and specificity, 54% and 79%, respectively and a likelihood ratio of 2.6. The CAM's ability to screen patients with delirium when presented with underlying cognitive impairment was moderately compromised; however, we note that this study was of low quality.

The Cole (2003) study reported the sensitivity and specificity for patients with dementia [69%: n=222/322; sensitivity: 100.0%; specificity: 96.8%] and those without dementia [31%: n=100/322; sensitivity: 100.0%; specificity: 98.3%]. We note that this study was low quality and the same symptoms were used to determine the index test and reference standard results.

Within group comparisons

One study (Cole 2003) separately compared the CAM (long version) and the Delirium Index (DI) with the DSM-III-R to identify the sensitivity and specificity of number of symptoms of delirium, irrespective of the type of symptoms. We note that this was a low quality case control study and that the same data were used for the CAM and the reference standard, but a separate test was carried out for the DI. This makes a direct comparison between CAM and DI unreliable (figure 6.11, Appendix K)

As shown in figure 6.12, the ROC plot that explores the effect of varying thresholds on sensitivity and specificity in a single study, the presence of 6 or more number of symptoms of delirium on the CAM (long version) compared with the DSM-III-R criteria was considered the best threshold point. This cut-off point was similar for patients with and without dementia.

We note this is a poor quality study and the same symptoms were used to determine the index test and reference standard results.

On the Delirium Index instrument, the presence of 4 or more symptoms and 3 or more symptoms showed the best sensitivity and specificity in patients with and without dementia, respectively.

6.4.5.1 DSM III as the reference standard

Two studies (Laurila* 2002; Ni Chonchubhair 1995) reported an index test compared with DSM III as the reference standard. A forest plot of sensitivity and specificity is shown in figure 6.13, and the diagnostic test accuracy statistics are summarised in table 6.7.

6.4.5.2 AMT serial test

One study (Ni Chonchubhair 1995) compared the change in AMT scores using the Delirium Assessment Scale to determine delirium according to the DSM III criteria. A 2 point decrease between preoperative and postoperative AMT score showed high sensitivity and specificity, 93% and 84%, respectively. A 3 point decline in AMT scores showed a lower sensitivity [67%] and higher specificity [95%].

CAM

One study (Laurila* 2002) comparing CAM (short version) with DSM-III showed a moderate sensitivity and specificity [85% and 82%, respectively]. The ability of the instrument to exclude the condition is still high [94%]; but the positive likelihood ratio is low [3.05].

Table 6.7: index test compared with DSM-III-R

DSM-III	Study name	Comments	Test operator	Sensitivity	Specificity	PPV	LR+	Pre-test probability	Post-test probability
CAM short version	Laurila *2002		Geriatrician	85.00	72.10	50.00	3.05	24.70	50.00
AMT	Ni Chonchubhair 1995	Cut off at decline of 3 points or more	Not stated / unclear	66.70	95.30	71.00	14.17	15.00	71.40
	Ni Chonchubhair 1995	Cut off at decline of 2 points or more	Not stated / unclear	93.30	83.50	50.00	5.67	15.00	50.00

6.4.6 Consensus diagnosis as a reference standard

One low quality study (Zou 1998) reported separately the sensitivity and specificity for two index tests [nurse assessed CAM (long version) and psychiatrist assessment] compared with a reference standard (expert consensus diagnosis); the expert group comprised two geriatric psychiatrists, a research fellow and a nurse. The consensus diagnosis was comprised of the following: psychiatrist's findings from a chart review and clinical examination; each professional's independent assessment on the presence or absence of delirium based on the psychiatrist's application of the DSM-IV criteria and the nurse's findings from the CAM and chart review. The forest plot of the sensitivity and specificity is shown in figure 6.15. The nurse's CAM rating showed a higher sensitivity [89%] than the psychiatrist diagnosis [71%]. The authors attributed this partly to the fact the nurse had more opportunities to observe and reassess the patient, as opposed to the psychiatrist who assessed the patient only once.

The results from the study should be treated with caution as this was considered a low quality study.

6.4.7 CAM (short version) and expert interviewer as the reference standard; MMSE serial test

One study (O'Keeffe 2005) examined the change in the MMSE scale between day 1 and day 6 of hospitalisation, to identify the best determinant for detecting the development and resolution of delirium. The diagnosis of delirium was with the CAM (short version) instrument and clinician interview.

The study found, for the detection of delirium, a decline of 2 or more points was the best determinant. The sensitivity and specificity were 93% and 90% respectively (figure 6.16). There was some uncertainty with the raw data which were back calculated from the diagnostic accuracy measures. The diagnostic test accuracy statistics are summarised in table 6.8.

Table 6.8: index test compared with CAM (short version) and clinician interview

CAM + interview by experienced clinician	Study name	Comments	Test operator	Sensitivity	Specificity	PPV	LR+	Pre-test probability	Post-test probability
MMSE (serial change)	O'Keeffe 2005	Some uncertainty with the raw data that were back calculated from these measures	Trained assessor	92.90	90.10	46.00	8.9	8.48	46.40

6.4.8 Comparison of different assessors for CAM (short version)

One low quality study (Monette 2001) compared CAM (short version) assessment by a lay interviewer with a geriatrician; there was no reference standard in this study. The team of lay interviewers included a nurse without prior research experience, a nurse with some experience as a research interviewer or an experienced research assistant without a nursing degree but with experience as a research interviewer.

Subgroup analyses by dementia or no dementia

The low quality Monette (2001) study presented results by those with possible or suspected dementia or no dementia. High sensitivity was observed for the two subgroups, but the lower specificity [78%] observed in the possible dementia group was attributed to a suggested limitation in CAM's (short version) ability to exclude those with underlying cognitive impairment. However, we note that this is a low quality study, so that results should be treated with caution (figure 6.17, Appendix K). The diagnostic test accuracy statistics are summarised in table 6.9.

Table 6.9: CAM (lay person) compared with CAM (geriatrician)

Study name	Comments	Test operator	Sensitivity	Specificity	PPV	LR+	Pre-test probability	Post-test probability
Monette 2001	CAM for patients with possible or probable dementia	Trained assessor (trained by psychiatrist)	96.40	78.30	84.00	4.4357	54.90	84.40
Monette 2001	no dementia	Trained assessor (trained by psychiatrist)	94.70	95.00	90.00	18.947	38.80	92.30

6.5 Results: ICU setting

6.5.1 Diagnostic test accuracy (DSM-IV as the reference standard)

CAM-ICU

Three moderate to high quality studies (Ely 2001; Ely 2001b; Lin* 2004) compared CAM-ICU with DSM-IV.

A forest plot of sensitivity and specificity is shown in figure 6.18 (Appendix K), and diagnostic test accuracy statistics are summarised in table 6.10.

The remaining studies were of good quality and showed a high sensitivity [range: 91% to 96%] and specificity [93% to 100%]. The likelihood ratio ranged from 13.42 to 36.36, showing a high likelihood that a patient found to be delirious based on the CAM-ICU, is delirious.

Table 6.10: diagnostic test accuracy statistics for CAM-ICU

CAM-ICU	Study name	Comments	Test operator	Sensitivity	Specificity	PPV	LR+	Pre-test probability	Post-test probability
	Ely 2001	CAM-ICU [Nurse 2] vs DSM-IV	Nurse	93.00	100.00	100.00	NA	14.13	100
	Ely 2001b	CAM-ICU [Nurse 2] vs DSMIV	Nurse	96.00.	93.00	96.00	13.42	63.20	95.80
	Lin 2004	CAM-ICU [Chinese] [Assessor 1] vs DSMIV [psychiatrist]	Research Assistant	90.90	97.50	91.00	36.364	21.60	90.90

Subgroup analyses by dementia or no dementia

Two studies (Ely 2001; Ely 2001b) reported subgroup analyses by dementia status. The number of patients with suspected dementia was 12.5% [12/96] and 28.9% [11/38], respectively in the two studies. In both studies suspected dementia was defined as: the delirium expert rating of having dementia, a Blessed Dementia Rating Scale score of at least 3, or a rating by a surrogate of at least 3 of out of 5 as 'possibly having dementia'.

The diagnostic test accuracy statistics are summarised in table 6.11.

Both studies reported 100% sensitivity and 100% specificity for patients with suspected dementia. However, the 95% confidence interval around these values was 56% to 100% for both the sensitivity and specificity in the Ely (2001b) study for all three raters and 63% to 100% for sensitivity (nurse 1; nurse 2: 95% CI 66% to 100%) and 40% to 100% for the specificity (nurse 1; nurse 2: 95% CI 3% to 100%) in the Ely (2001) study. The number of patients within this subgroup analysis in both studies is small (Ely 2001: n=12; Ely 2001b: n=11) and the authors suggested that the criteria for identifying patients with suspected dementia was liberal.

Table 6.11: diagnostic test accuracy statistics for CAM-ICU - dementia subgroup

CAM-ICU	Study name	Comments	Test operator	Sensitivity	Specificity
	Ely 2001	CAM-ICU [Nurse 1] vs DSMIV; suspected dementia (n=12)	Nurse 1	100.00	100.00
	Ely 2001	CAM-ICU [Nurse 2] vs DSM-IV Suspected dementia (n=12)	Nurse 2	100.00	100.00
	Ely 2001	CAM-ICU [Nurse 1] vs DSMIV; not suspected dementia (n=84)	Nurse 1	98.00	100.00
	Ely 2001	CAM-ICU [Nurse 2] vs DSM-IV not suspected dementia (n=84)	Nurse 2	100.00	91.00
	Ely 2001b	CAM-ICU [Nurse 1] vs DSMIV] Suspected dementia (n=11)	Nurse 1	100.00	100.00
	Ely 2001b	CAM-ICU [Nurse 2] vs DSMIV] Suspected dementia (n=11)	Nurse 2	100.00	100.00
	Ely 2001b	CAM-ICU [Intensivist] vs DSMIV] Suspected dementia (n=11)	Intensivist	100.00	100.00

6.6 Results: mixed setting

6.6.1 Comparison of diagnostic criterion tools [DSM-IV as the reference standard].

One low quality study (Laurila* 2003) and one report of that study (Laurila* 2004) compared three sets of diagnostic criteria in the same patients, using the same data: DSM-III-R; DSM-III and ICD-10 with DSM-IV, in both hospital wards and nursing homes. The study operationalised the clinical and research criteria of the ICD-10 and the criteria from the DSM-IV, DSM-III-R, and DSM-III into one questionnaire. The Laurila* (2004) study reported a subgroup analysis (see section below titled '*Subgroup analyses*').

The forest plot of sensitivity and specificity is shown in figure 6.18 (Appendix K) and diagnostic test accuracy statistics are summarised in table 6.12.

The ICD-10 showed the lowest sensitivity [24%], whilst the DSM-III-R showed the highest sensitivity [78%]. All three tests showed high specificity. The study reported that the DSM-IV criteria were the most inclusive in the hospital [34.8% of the patients were considered to be delirious], and the DSM-III-R criteria were the most inclusive in the nursing homes [14.4% of the patients were considered to be delirious].

Table 6.12: diagnostic test accuracy statistics for diagnostic criterion tools; mixed setting

Study name	Comments	Test operator	Sensitivity	Specificity	PPV	LR+	Pre-test probability	Post-test probability
Laurila * 2003	ICD10 vs DSMIV	Geriatrician [hospital]/Nurse [LTC]	40.60	100.00	100.00	NA	24.90	100.00
Laurila* 2003	DSM IIIR vs DSMIV	Geriatrician [hospital]/Nurse [LTC]	79.57	97.18	89.00	28.20	24.94	90.3
Laurila * 2003	DSMIII vs DSMIV	Geriatrician [hospital]/Nurse [LTC]	75.50	100.00	100.00	NA	24.90	100.00

Subgroup analyses

One report (Laurila* 2004) of the low quality Laurila* (2003) study reported the number of patients with and without dementia diagnosed with delirium with three index tests. Dementia diagnosis was based on the consensus diagnosis of three geriatricians based on the following information: prior dementia diagnoses, Clinical Dementia Rating Scale, operationalised criteria according to the DSM-IV, nurses and/or caregivers' interviews and the results of the brain CT (computed tomography)/MRI (magnetic resonance imaging) and prior MMSE scores, where available. The number of patients diagnosed with and without dementia were as follows: ICD-10: 15% [38/255]: 2.9% [5/170]; DSM-III-R: 23% [58/255]: 13% [22/170]; DSM III: 23% [58/255]:13% [22/170] in comparison with DSM-IV (26% :[66/255]: 24% [40/170]) as the reference standard. However, there was insufficient information so we were unable to construct 2x2 tables and report on the sensitivity and specificity of these results.

6.7 Health economic evidence

No relevant health economic papers were identified.

6.8 Clinical evidence statements

Evidence statements relating to the CAM index test have only been presented for the short version. This is because the CAM short version is widely used in practice whilst the long version is mainly used for research purposes.

6.8.1 Hospital setting

- For the diagnosis of delirium there is moderate quality evidence to show that:
    - the CAM test (short version) has moderate sensitivity and specificity with the DSM-IV criteria for delirium, followed by the DSM-III and DSM-III-R, and is in least agreement with the ICD-10 criteria for delirium.
    - the CAM test (short version) compared with the DSM-IV has a moderate sensitivity and specificity as a method for assessing delirium.
    - the MMSE test compared with the DSM-III-R has a low sensitivity and specificity as a method for assessing for delirium.
    - the clock-drawing test compared with the DSM-III-R has a low sensitivity and specificity as a method for dignosing delirium.

- For the diagnosis of delirium there is low quality evidence to show that:
    - the DSM-III-R criteria for delirium shows a moderate sensitivity and specificity with the DSM-IV criteria for delirium; same symptoms were used to determine both test results.
    - the ICD-10 criteria for delirium are less inclusive than the DSM III criteria, when compared with the DSM-III-R criteria for delirium.

- the DRS-R-98 test compared with the DSM-IV has a fairly low to moderate sensitivity and specificity as a method for assessing delirium.
- the CAM test (short version) compared with the DSM-III-R has a low ability to identify patients with delirium with underlying cognitive impairment.
- for the CAM test (short version) and Delirium Index the best threshold points are the:
- presence of 6 or more symptoms of delirium on the CAM test (compared with the DSM-III-R) irrespective of dementia status. We note the study was of poor quality and the same symptoms were used to determine the index test and reference standard results.
- presence of 4 symptoms of delirium on the Delirium Index test (compared with the DSM-III-R) in patients with dementia.
- presence of 3 or more symptoms of delirium on the Delirium Index test (compared with the DSM-III-R) in patients without dementia.

6.8.2 ICU setting

- For the diagnosis of delirium, there is moderate to high quality evidence to show that the CAM-ICU test compared with the DSM-IV, has moderate sensitivity and specificity as an assessment method, irrespective of dementia status.

6.8.3 Mixed setting (hospital and long-term care)

- For the diagnosis of delirium there is moderate quality evidence to show that the:
    - DSM-III-R criteria is the most inclusive followed by the DSM-III criteria compared with the DSM-IV criteria for delirium.
    - ICD-10 criteria to be the least inclusive compared with the DSM-IV criteria for delirium.

6.9 From evidence to recommendations

The GDG identified two stages in the diagnostic process: an initial stage and a confirmation stage. The GDG agreed that the first stage should be a symptom-based approach and made a consensus recommendation based upon this. This was partly informed by the standard operational definition of delirium (the DSM criteria) and partly by GDG clinical experience. For the second stage, there was low to moderate quality evidence from the review of diagnostic test accuracy for different tests, comparing them with the reference standard of the DSM IV criteria. This review and the epidemiology review also compared different criteria over the years that have been developed as the standard operational definition for delirium.

1st stage of the diagnostic process (recommendations 1.2.1 and 1.4.1)

The initial stage is intended to alert any healthcare professional, including the non-specialist, to warning signs that the patient may have delirium.

The GDG debated the appropriate time to carry out the initial stage, and considered whether to complete the initial assessment at the person's first presentation to hospital or long-term care. This would mean that all patients presenting to the accident and emergency department would have to undergo the test and the GDG considered this impractical in this setting. They decided that only people who had already been determined to be at-risk of delirium at presentation (see recommendation 1.1.1) should be assessed for prevalent delirium (recommendation 1.2.1).

People 'in hospital' (i.e. admitted) or in long-term care should subsequently be observed at least daily for signs of incident delirium regardless of whether they are at-risk or not (recommendation 1.4.1).

The GDG considered using a simple validated diagnostic tool such as the clock drawing test and also the MMSE, but noted from the evidence that these tools had low sensitivity. The GDG was keen that the first stage of assessing delirium was based upon clinical signs and symptoms that could be easily identified by the non-specialist. The GDG noted that warning signs are recent changes or fluctuations in usual behaviour, and compiled a list of clinical indicators based on their clinical experience. The GDG considered these indicators still applied in the ICU, but noted that these patients were likely to pass rapidly to the 2nd stage assessment.

The GDG felt that healthcare professionals should be particularly vigilant in recognising hypoactive delirium, because those particular behaviours are easily missed in clinical practice. The behaviours indicating hypoactive delirium have been highlighted in the recommendation by the means of an asterisk.

Sometimes the patient, their family or carer notice and report changes in behaviour which would otherwise be unnoticed by the healthcare professional. The GDG decided to emphasise and include this in the recommendation.

2nd stage of the diagnostic process / confirmatory stage (recommendation 1.5.1)

For the second stage of the diagnostic process, the GDG recommended a clinical assessment should be carried out for delirium by a trained healthcare professional. They then considered whether this assessment should be based on the DSM-IV diagnostic criteria or a diagnostic test and concluded that it was important to give healthcare professionals the option of using either the DSM-IV or a diagnostic test. The review of diagnostic test accuracy showed that both the long and short versions of the CAM, CAM-ICU and the AMT, all had acceptable sensitivity. The GDG noted that the long version of the CAM was not used in clinical practice and serial tests (such as AMT and MMSE) may be considered for those under elective care, but have limited clinical utility in relation to patients with a high risk of delirium. The GDG decided the short version of CAM and CAM-ICU (for critical care patients) should be recommended as alternatives to DSM-IV as the basis for clinical assessment.

The GDG noted the evidence from one moderate quality study (Radtke 2008) that CAM had only 43% sensitivity for diagnosing delirium in a population that was in the recovery room following surgery. The GDG considered this to be an inappropriate test for this population and agreed to recommend using the CAM-ICU in the recovery room following surgery.

The GDG acknowledged that there is often difficulty differentiating between the diagnoses of delirium, delirium superimposed on dementia or dementia. The GDG considered that, when uncertainty existed, patients should be assessed and treated initially with an assumption of delirium. This prioritisation of delirium implicitly recognised and emphasised delirium as a serious acute illness that can be treated effectively.

Because a specific diagnostic test for delirium does not exist per se, the GDG wished to make a recommendation for future research (see below and Appendix H)

Future research recommendation:

The development and validation of a new test for delirium

6.10 Recommendations

Indicators of delirium: at presentation

At presentation, assess people at risk for recent (within hours or days) changes or fluctuations in usual behaviour. These may be reported by the person at risk, or a carer or relative. Be particularly vigilant for behavior indicating hypoactive delirium (marked *). These behaviour changes may affect:

- Cognitive function: for example, worsened concentration*, slow responses*, confusion.
- Perception: for example, visual or auditory hallucinations.
- Physical function: for example, reduced mobility*, reduced movement*, restlessness, agitation, changes in appetite*, sleep disturbance.
- Social behaviour: for example, lack of cooperation with reasonable requests, withdrawal*, or alterations in communication, mood and/or attitude.

If any of these behavior changes are present, a healthcare professional who is trained and competent in diagnosing delirium should carry out a clinical assessment to confirm the diagnosis. **[1.2.1]**

Indicators of delirium: daily observations

Observe, at least daily, all people in hospital or long-term care for recent (within hours or days) changes or fluctuations in usual behaviour (see recommendation 1.2.1). These may be reported by the person at risk, or a carer or relative.

If any of these behaviour changes is present, a healthcare professional who is trained and competent in the diagnosis of delirium should carry out a clinical assessment to confirm the diagnosis. **[1.4.1]**

Diagnosis (specialist clinical assessment)

If indicators of delirium are identified, carry out a clinical assessment based on the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) criteria or short Confusion Assessment Method (short CAM) to confirm the diagnosis. In critical care or in the recovery room after surgery, CAM-ICU should be used. A healthcare professional who is trained and competent in the diagnosis of delirium should carry out the assessment. If there is difficulty distinguishing between the diagnoses of delirium, dementia or delirium superimposed on dementia, treat for delirium first. **[1.5.1]**

7 Risk factors for delirium: non-pharmacological

CLINICAL QUESTION:

What are the risk factors for delirium?

What are the precipitating factors for delirium?

7.1 Clinical introduction

Delirium is a complex syndrome and patients appear to differ in their susceptibility to the condition. For example, some patients develop delirium with a urinary infection, while others do not. Understanding the underlying risk factors for delirium helps to explain this clinical variation. It also provides an opportunity to identify people who are at higher risk of delirium and, importantly, consider modifying key risk factors such that delirium incidence might be reduced.

7.2 Selection criteria

Selection criteria were as outlined in the general methods section (section 2.3.1) apart from the types of risk factor described below.

7.2.1 Types of risk factor

Any variable reported to be a risk factor for delirium was to be considered, including the following *a-priori* ones predicted by the GDG:

Patient Characteristics

- Age
- Sex
- Dementia
- Sensory impairment
- Severity of illness
- Depression
- Multiorgan failure
- Polypharmacy (having more than one drug)

- Dehydration
- Electrolyte disturbance
- Continence
- Constipation
- Hypoxia
- Immobility/ bedridden
- Infection
- Malnutrition
- Sleep deprivation

7.2.1.1 Environmental

- Setting
- Lighting
- Orientation
- Sensory overload

7.2.1.2 Procedural

- Type of anaesthesia
- Cardiac surgery
- Hip fractures
- Insertion of urinary catheter
- Any iatrogenic intervention
- Smoking cessation
- Physical restraint

7.3 Description of studies

Details of included and excluded papers together with study design are reported in table 7.1

Table 7.1: study inclusion, exclusion, and design

Papers	Comments	References
N=85 evaluated for inclusion		
N= 36 excluded	11 studies excluded because fewer than 20 patients developed delirium Reasons for exclusion for the remaining 25 studies are reported in Appendix G	(Clayer 2000: n=9; Duggleby 1994: n=16; Eriksson 2002: n=12; Hamann 2005: n=7; Kaneko 1997: n=6; Kawaguchi 2006: n=13; Koebrugge 2009: n=17; McAlpine 2008: n=18; Milstein 2000: n=10; Naughton 1995: n=18; Wakefield 1996: n=16);
N=11 identified in update searches	Studies were not presented in the results as they were of low quality	(Angles 2008; Chang 2008; Detroyer 2008; Galankis 2001; Gao 2008; Greene 2009; McManus 2009; Oh 2008; Robinson 2008; Van Rompaey 2009; Yang 2008).
N=38 included	Study design 32 Prospecitve cohort studies	(Andersson 2001; Böhner 2003; Bucerius 2004; Caeiro 2004; Edlund 2001; Ely 2007; Furlaneto 2006; Goldenberg 2006; Hofsté 1997; Inouye 1993; Inouye 2007; Kazmierski 2006; Korevaar 2005; Leung 2007; Levkoff 1992; Lin 2008; Margiotta 2006; McCusker 2001; Olin 2005; Ouimet 2007; Pandharipande 2006*; Pisani 2007; Pompei 1994; Ranhoff 2006; Rolfson 1999; Rudolph 2007; Santos 2004; Schor 1992; Sheng 2006; Veliz-Reissmüller 2007; Weed 1995; Zakriya 2002)
	3 retrospective cohort studies	(Levkoff 1988; Redelmeier 2008; Yildizeli 2005)
	3 cross-sectional design	(Ramirez-Bermudez 2006; Sandberg 2001; van Munster 2007). These studies were not reported further, because this is a poor study design and other data were available from the cohort studies.
	* Pandharipande (2006) study was identified from the pharmacological risk factors review (chapter 8) and details are given in section 8.3.	

None of the studies were carried out in the UK. Information on study sizes and geographical location are described in table 7.2.

Table 7.2: study characteristics. (*Studies denoted with italics indicate retrospective cohort studies).*

Study	Size (N)	Geographical location
Andersson 2001	457	Sweden
Böhner 2003	153	Germany
Bucerius 2004	16,184	Germany
Caeiro 2004	218	Portugal
Edlund 2001	101	Sweden
Ely 2007	53	USA
Furlaneto 2006	103	Brazil
Goldenberg 2006	77	USA
Hofsté 1997	321	Netherlands
Inouye 1993	107	USA
Inouye 2007	491	USA
Kazmierski 2006	260	Poland
Korevaar 2005	126	Netherlands
Leung 2007	203	USA
Levkoff 1988	1,285	USA
Lin 2008	151	Taiwan
Levkoff 1992	325	USA
Margiotta 2006	330	Italy
McCusker 2001	444	Canada
Olin 2005	61	Sweden
Ouimet 2007	764	Canada
Pisani 2007	304	USA
Pompei 1994	755	USA
Ranhoff 2006	401	Italy
Redelmeier 2008	284,158	Canada
Rolfson 1999	75	Canada
Rudolph 2007	1,218	USA
Santos 2004	220	Brazil
Schor 1992	291	USA
Sheng 2006	156	Australia
Veliz-Reissmüller 2007	107	Sweden
Weed 1995	138	USA
Yildzeli 2005	432	Turkey
Zakriya 2002	168	USA

All of the studies included hospital patients. The study by Pompei (1994) analysed data separately from two studies: n=432 from Chicago Hospital and n=323 from New Haven Hospital (data were not combined).

The study by Levkoff (1992) reported data separately for patients who were admitted to hospital from institutional settings (n=114, 35%), and those who were admitted from community settings (n=211), as well as combining the samples (reported for some risk factors). Nine other studies reported the patients' pre-hospital setting:

- Goldenberg (2006) had 79% of patients from the community and 21% from skilled nursing facilities

- Inouye (1993) reported that 3% of patients had been living in a nursing home
- Pisani (2007) had 18% patients from a nursing home
- Schor (1992) had 30% of patients from an institutional setting
- Andersson (2001) had 53% of patients living alone and 11% in sheltered accommodation
- Pompei (1994) Chicago hospital had 31% patients living alone and Pompei (1994) New Haven hospital had 41% living alone
- Ranhoff (2006) had 25% patients living alone
- Sheng (2006) had 90% patients living alone
- McCusker (2001) had 71% living alone, 18% from a foster home/senior residence, and 11% from a nursing home

Eighteen studies were carried out in patients admitted for surgery (Andersson 2001; Böhner 2003; Bucerius 2004; Edlund 2001; Furlaneto 2006; Goldenberg 2006; Hofsté 1997; Kazmierski 2006; Leung 2007; Olin 2005; Redelmeier 2008; Rolfson 1999; Rudolph 2007; Santos 2004; Veliz-Reissmüller 2007; Weed 1995; Yildizeli 2005; Zakriya 2002):

- Seven studies were conducted in patients undergoing cardiac operations generally (Veliz-Reissmüller 2007), with and without cardiopulmonary bypass (CPB) (Bucerius 2004), or with CPB only (Hofsté 1997), or undergoing coronary artery bypass graft (CABG) surgery (Rolfson 1999; Santos 2004), or open heart surgery (Kazmierski 2006), or aortic, carotid, and vascular surgery (Böhner 2003)
- Five studies were in patients who had surgery for hip fracture (Andersson 2001; Edlund 2001; Furlaneto 2006; Goldenberg 2006; Zakriya 2002)
- One study was in patients who had major elective or urgent thoracic surgery (Yildizeli 2005)
- One study was in patients who had abdominal surgery (Olin 2005)
- One study was in patients who had head and neck cancer surgery (Weed 1995)
- Two studies were in patients undergoing non-cardiac surgery (Leung 2007; Rudolph 2007)
- One study was in patients undergoing cardiac, thoracic, neurosurgical, vascular, musculoskeletal, lower urologic and gynaecologic, breast and

skin, external head and neck, and ophthalmologic surgery (Redelmeier 2008).

Four studies evaluated patients from both surgical and medical wards (Levkoff 1988; 1992; Pompei 1994; Schor 1992): in the study by Levkoff (1992) the principal diagnoses of patients admitted to hospital included circulatory, digestive, respiratory or genitourinary system diseases; endocrine, nutritional and metabolic diseases; fractures; cancer; diseases of the skin or other reasons not stated. Reasons for admission were not stated in the study by Pompei (1994). In the study by Schor (1992), 61% were admitted to medical wards, 21% to general surgery, and 8% to orthopaedic surgery.

Seven studies evaluated patients in medical wards only (Caeiro 2004 – stroke unit; Inouye 1993; Inouye 2007; Korevaar 2005; Margiotta 2006; McCusker 2001; Sheng 2006):

- Two studies included acute stroke patients (Caeiro 2004; Sheng 2006)
- One study included patients admitted to an internal medicine ward with diagnoses including infectious disease, malignancy, gastrointestinal bleeding, water and electrolyte disturbances and other reasons not stated (Korevaar 2005)
- Reasons for admission were not stated in four studies (Inouye 1993; Inouye 2007; Margiotta 2006; McCusker 2001).

Six studies evaluated patients in intensive care units (ICUs) (Ely 2007; Lin 2008; Ouimet 2007; Pandharipande 2006; Pisani 2007; Ranhoff 2006):

- Three studies included mechanically ventilated patients in ICU (Ely 2007; Lin 2008; Pandharipande 2006;)
- One study was in patients with admission diagnoses of respiratory, gastrointestinal haemorrhage, sepsis, neurological or other causes (Pisani 2007)
- One study included patients admitted to a sub-intensive care unit for older people; diagnoses included respiratory failure, cardiac diseases, stroke, gastrointestinal bleeding, cancer-related problems, acute renal failure or other diagnoses not stated (Ranhoff 2006)
- Reasons for admission were not stated in the study by Ouimet (2007)

7.3.1 Population

Details about the population are summarised in this section, focussing on the principal risk factors; further details are given in Appendix F.

The mean **age** ranged from 51.7 years (Yildizeli 2005) to 87.4 years (Levkoff institution 1992). Age ranges are given in table 7.3; two studies did not report on patient age (Böhner 2003; Levkoff 1988). The GDG concluded that two studies had a narrow age range that could be considered to be effectively constant (Olin 2005; Rolfson 1999).

Table 7.3: Patient ages (+/- indicates that the range was calculated from the mean +/- 1 standard deviation)

Study	Age range (years)	Study	Age range (years)
Andersson (2001)	65-96	Margiotta (2006)	65-100
Böhner (2003)	not stated	McCusker (2001)	76-90 (+/-)
Bucerius (2004)	54-75 (+/-)	Olin (2005)	70-80
Caeiro (2004)	24-86	Ouimet 2007)	48-78
Edlund (2001)	65-102	Pandharipande 2006	25-90
Ely (2007)	31-79	Pisani (2007)	66-83
Furlaneto (2006)	71-90	Pompei (1994) (Chicago)	68-83
Goldenberg 2006)	66-98	Pompei (1994) (Yale)	73-85 (+/-)
Hofsté (1997)	29-83	Ranhoff (2006)	60-94
Inouye (1993)	73-86 (+/-)	Redelmeier (2008)	67-80
Inouye (2007)	73-85 (+/-)	Rolfson (1999)	69-74
Kazmierski (2006)	25-81	Rudolph (2007)	63-75 (+/-)
Korevaar (2005)	71-87 (+/-)	Santos (2004)	66-78
Leung (2007)	66-78 (+/-)	Schor (1992)	73-88 (+/-)
Levkoff (1988)	not stated	Sheng (2006)	65-95
Levkoff (1992)	74-89 (+/-)	Veliz-Reissmüller (2007)	65-95
Levkoff institution (1992)	80-95 (+/-)	Weed (2005)	mean 64
Levkoff community (1992)	71-85 (+/-)	Yildizeli (2005)	18-86
Lin (2008)	64-86	Zakriya (2002)	50-98

The studies varied in the proportions of patients reported to have **cognitive impairment** at baseline. In addition, the GDG decided that, when this was not clearly stated, it was unlikely that patients undergoing elective cardiac surgery would have cognitive impairment at baseline. This gave the following subgroups:

- No studies were carried out in which all the patients had cognitive impairment

- Twenty-two studies reported that some patients had cognitive impairment or dementia at baseline (Caeiro 2004; Edlund 2001; Ely 2007; Furlaneto 2006; Goldenberg 2006; Hofsté 1997; Inouye 1993; Inouye 2007; Kazmierski 2006; Korevaar 2005; Leung 2007; Levkoff 1992; Margiotta 2006; McCusker 2001; Olin 2005; Pisani 2007; Pompei 1994; Rolfson 1999; Schor 1992; Sheng 2006; Veliz-Reissmüller 2007; Weed 2005)

- Inouye (1993) also excluded patients with severe underlying dementia

- Two studies stated that patients with cognitive impairment at baseline were excluded from their studies (Andersson 2001; Santos 2004) and four studies excluded patients with pre-existing dementia (Kazmierski 2006; Lin 2008; Rudolph 2007; Zakriya 2002).

- Rudolph (2007) included patients with mild cognitive impairment, but not dementia

- Kazmierski (2006) reported results for cognitive impairment as a risk factor

- One ICU study (Ranhoff 2006) reported scores on the MMSE at discharge from the hospital and used this together with measures of pre-admission activities of daily living (ADL) to determine pre-existing dementia (which the authors described as ‘probably demented’). This is, at best, an indirect measure of pre-existing dementia, but it was used in the multivariate analysis

- It was not stated if the patients had cognitive impairment at baseline in five studies (Böhner 2003; Bucerius 2004; Levkoff 1988; Ouimet 2007; Redelmeier 2008).

    - Three of these studies were carried out in elective heart surgery patients who would be unlikely to have cognitive impairment (Böhner 2003; Bucerius 2004; Redelmeier 2008)

    - However, we note that three elective cardiac surgery studies stated that some patients had cognitive impairment at baseline (e.g. Rolfson 1999; Veliz-Reissmüller 2007)

Of the studies that assessed cognitive impairment and/or dementia, 18 used the Mini Mental State Examination (MMSE) score, two used DSM-IV; four used Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE); and two used the Blessed dementia questionnaire; four studies did not report what scale was used (table 7.4). One study (Caeiro 2004) had less than 10% of patients with cognitive impairment, so that any results for cognitive impairment in this study were likely to be inaccurate. The GDG considered that the cut-off point of 28 on the MMSE scale, used in the Veliz-Reissmuller (2007) study, was unreliable and this study was not included in the analyses for cognitive impairment.

Sensory impairment was reported in twelve studies (Andersson 2001; Böhner 2003; Edlund 2001; Inouye 1993; 2007; Margiotta 2006; McCusker 2001; Pisani 2007; Ranhoff 2006; Schor 1992; Sheng 2006; Weed 2005). Four studies excluded patients with *severe* visual and/or hearing impairment (Levkoff 1992; Olin 2005; Santos 2004; Schor 1992); Hofsté (1997) and Rolfson (1999) excluded people who were blind or deaf, but the GDG did not consider this to be a modifiable risk factor for sensory impairment and noted that there would be other people who did have other degrees of sensory impairment. The studies did not generally give much information on how sensory impairment was assessed:

- Andersson (2001) and Pisani (2007): stated it was patient reported and proxy reported respectively

- Ranhoff (2006): patient/close relative was asked if they had vision problems affecting daily activity

- Inouye (1993) and Inouye (2007): Jaeger- and Snellen-type tests for standard vision – visual impairment was defined as corrected vision worse than 20/70 on both near and distant binocular tests. For hearing impairment, the Inouye (2007) study used a whisper test and Inouye (1993) used a Welch-Allyn audioscope and questions designed to screen for hearing loss – hearing impairment was defined if the patient heard fewer than three of eight tones on the audioscope (at 40 dB and frequencies of 500, 1000, 2000 and 4000 Hz) and a score of 4 or less (of 8) on the screening tests

- McCusker (2001): no details, but the study also included in the analysis whether or not the patient was wearing reading glasses

- Sheng (2006) in stroke patients recorded 'vision field loss'

Levels of sensory impairment are given in table 7.5.

Table 7.4: Cognitive impairment and/or dementia

Study	Cognitive impairment and/or dementia
Caeiro (2004)	Unstated scale: 3% had dementia/cognitive decline
Edlund (2001)	DSM-IV: 21 of 101 (21%) patients had dementia
Ely (2007)	IQCODE: 16% had a mean score of 4 or more
Furlaneto (2006)	MMSE: mean 12.07 (SD 9.04) in delirium group and 17.74 (SD 8.78) in control group; Blessed dementia questionnaire to caregiver: 45% had a score above 4
Goldenberg (2006)	MMSE: mean score 21.6 (range 2 to 30); DSM-IV: 53 of 77 (69%) had dementia
Hofsté (1997)	MMSE: 23% reported to have cognitive disorders
Inouye (1993)	MMSE: mean score 24.2 (5.0); 36% with a score below 24
Inouye (2007)	MMSE: mean 23.1 (SD 6.3); 39% with a score below 24; modified Blessed dementia questionnaire to family member: 20% had a score above 4
Kazmierski (2006)	MMSE: 53% in group with delirium and 16% in group without delirium (preoperatively) had a score equal to or below 24
Korevaar (2005)	MMSE: 53% had a score below 24; IQCODE: 43% had a mean score of 3.9 or more
Leung (2007)	MMSE: mean score 33 (SD 3.2)
Levkoff (1992)	Unstated scale: 24% had cognitive impairment
Margiotta (2006)	MMSE: mean score 16.9 (SD 6.8) in patients with delirium and 22.1 (SD 7.0) in patients without delirium
McCusker (2001)	IQCODE: 60% with a score of 3.5 or more
Olin (2005)	MMSE: mean score 28 (SD 3)
Pisani (2007)	IQCODE: 31% had a mean score of 3.3 or more
Pompei (1994)	MMSE: 37% had cognitive impairment
Ranhoff (2006)	MMSE on discharge: mean score was 19.1 (SD 11) prior to hospital admission; 30% had MMSE score less than 18 and/or Barthel Index less than 95 and/or IADL

Study	Cognitive impairment and/or dementia
	impairment on 1 or more tasks
Rolfson (1999)	MMSE: 9% in group with delirium and 12% in group without delirium using a cut-off of 24
Rudolph (2007)	MMSE: mean 27.8 (SD 1.6) at baseline
Santos (2004)	MMSE: no patients with cognitive impairment
Schor (1992)	Unstated scale: 42% had a history of cognitive impairment in delirium group and 10% in group without delirium
Sheng (2006)	MMSE: overall scores at one month were 23.4 (SD 6); 8% were reported to have dementia
Veliz-Reissmüller (2007)	MMSE: median score 29 (range 17-30) in group with delirium and 30 (range 27-30) in group without delirium; cut-off was 28
Weed (1995)	MMSE: mean score 26.3 in patients with delirium and 27.4 in patients without delirium
Zakriya 2002	Method of assessment not stated

Table 7.5: sensory impairment

Study	Visual impairment	Hearing impairment
Andersson 2001	31%	39%
Böhner 2003	61%	24%
Edlund 2001	23%	30%
Inouye 1993	6%	54%
Inouye 2007	38%	Not reported
McCusker 2001	20% with visual/hearing impairment; the authors also reported that 48% patients were wearing glasses, and 8% used a hearing aid	
Margiotta 2006	Some patients with sensory impairment (details not reported)	
Pisani 2007	11%	17%
Ranhoff 2006	29%	Not reported
Schor 1992	33%	21%
Sheng 2006	18%	Not reported
Weed 2005	5%	11%

Eight studies reported on the number of drugs (**polypharmacy**) taken by patients (Goldenberg 2006; Inouye 2007; Korevaar 2005; Olin 2005; Ranhoff 2006; Rolfson 1999; Veliz-Reissmüller 2007; Weed 1995). Where reported, the mean number of drugs ranged from 1.4 (Rolfson 1999) to 8.5 (Ranhoff 2006).

- Goldenberg (2006) reported that 87% of the patients had more than three medications at baseline (means were not reported)

- Inouye (2007) reported that 56% of the patients had more than three hospital medications in one day, and 29% had more than three psychoactive medications in one day

- Korevaar (2005) reported that the mean number of drugs used before admission was 4.4 (SD 3.2) in patients with delirium and 4.9 (SD 3.6) in patients without delirium

- Olin (2005) reported that the mean number of drugs taken was 3.0 (SD 3) in patients with delirium and 2.1 (SD 2) in patients without delirium

- Ranhoff (2006) reported that the mean number of drugs used was 8.5 (SD 3.4) in patients with prevalent delirium, 8.0 (SD3.2) in patients with incident delirium, and 7.3 (SD 3.1) in patients without delirium

- Rolfson (1999) reported that mean number of *selective drugs* used (dimenhydrinate, meperidine, or any benzodiazepine) was 1.4 in patients with delirium and 1.6 in the patients without delirium

- Veliz-Reissmüller (2007) reported that the mean number of drugs taken was 6.2 (SD 3.4) in the group with delirium and 6 (SD 3) in the group without delirium

- Weed (1995) reported that the mean number of medications was 3.4 in patients with delirium and 3.0 in patients without delirium.

The GDG considered a definition of polypharmacy and did not agree on a suitable cut-off point: either 3 or 5 drugs were suggested, depending on setting. The GDG ruled that, for studies in older patients undergoing cardiac surgery, polypharmacy was likely to be present in all patients (i.e., Böhner 2003; Bucerius 2004; Rolfson 1999; Santos 2004; Veliz-Reissmüller 2007). Similarly the GDG regarded studies in ICU as having the majority of patients with polypharmacy (i.e., Ely 2007; Lin 2008; Ouimet 2007; Pisani 2007; Ranhoff 2006).

Comorbidities were reported in most of the studies, with the exception of Inouye (1993); Inouye (2007) and Rolfson (1999). Generally, they included conditions related to heart disease (congestive heart failure, previous myocardial infarction, atrial fibrillation), angina, stroke, hypertension, diabetes, obesity, renal dysfunction, chronic obstructive pulmonary disease, asthma, hypothyroid, cancer, and depression. Two studies reported baseline Charlson Comorbidity Index data (Inouye 2007; McCusker 2001). In these studies, the mean scores were 2.7 (SD 2.1) and 2.7 (SD 2.0) respectively.

7.4 Methodological quality of included studies

The methodological quality of studies was assessed according to the type of study design. In evaluating the literature, RCTs and cohort studies were selected to be the best available evidence source for this review. Cross-sectional and case-control studies were not included in this review unless there was no other information. For details of quality assessment, see appendix E.

7.4.1 RCTs

No RCTs met the inclusion criteria.

7.4.2 Cohort studies

Representativeness and prospectiveness

None of the 35 cohort studies were considered to be truly representative of the population (i.e. adults in surgical and/or medical wards in hospital or people in long-term care). In all studies except the McCusker (2001) study, the non-exposed cohort was drawn from the same community as the exposed cohort. The McCusker (2001) was a secondary analysis of data from two related concurrent studies, an RCT in patients with delirium, and non-delirious patients were selected from patients screened for delirium but free of the condition.

All studies were prospective apart from three (Levkoff 1988; Redelmeier 2008; Yildizeli 2005), which were retrospective.

Missing data

Eight studies reported less than 20% loss to follow-up (Caeiro 2004; Edlund 2001; Inouye 2007; Leung 2007; Lin 2008; Rolfson 1999; Rudolph 2007; Veliz-Reissmüller 2007); the remaining studies reported that all the patients were followed up, with the exception of McCusker (2001) and Pandharipande 2006, in which it was not clearly reported.

One study reported an *a priori* sample size calculation (Rolfson 1999). In this study, a sample size of 81 was estimated assuming alpha=0.05, beta=0.20, and a desired margin of error of 0.10, with an anticipated proportion of delirium of 30%. The sample size of this study was 75.

Delirium at baseline

The studies varied in the number of patients with prevalent delirium (delirium at baseline): further details are given in Appendix D.

- Sixteen studies reported that none of the patients had delirium at baseline (Andersson 2001; Böhner 2003; Goldenberg 2006; Inouye 1993; Inouye 2007; Kazmierski 2006; Levkoff 1988; Lin 2008; Olin 2005; Rolfson 1999; Rudolph 2007; Santos 2004; Schor 1992; Veliz-Reissmüller 2007; Yildizeli 2005; Zakriya 2002)
    - eight of these studies excluded patients with delirium at baseline from their studies (Andersson 2001; Goldenberg 2006; Inouye 1993; Inouye 2007; Kazmierski 2006; Olin 2005; Rolfson 1999; Schor 1992; Zakriya 2002).
- Six studies reported that some patients had delirium at baseline (Edlund 2001; Furlaneto 2006; Levkoff 1992; Margiotta 2006; Pompei 1994; Ranhoff 2006).
    - Two studies excluded these patients from the analysis: (Edlund 2001: 61% of all patients; Levkoff 1992:10%)
    - Three studies (four cohorts) included these patients in the analysis together with patients with incident delirium:
        - Furlaneto (2006): 17% (17/103) prevalent, 13% (13/103) incident; 57% of all delirium was prevalent (17/30)
        - Pompei (1994) Chicago: 5% (21/463) prevalent, 9% (43/463) incident; 33% of all delirium was prevalent (21/64)
        - Pompei (1994) Yale: 15% (48/323) prevalent, 12% (38/323) incident; 56% of all delirium was prevalent (48/86)

    - Margiotta (2006): 9% (31/330) prevalent, 10% (32/330) incident; 49% was prevalent (31/63)

  - One study (Ranhoff 2006) reported that 16% (62/401) of patients had prevalent delirium, and 14% (55/410) had incident delirium; 53% of all delirium was prevalent. This study was carried out in a sub-ICU and prevalent delirium was diagnosed within 24 hours of admission to ICU. The GDG did not believe that incident and prevalent delirium could be distinguished in this population (because patients had come from other parts of the hospital) and all delirium was assumed to be incident.

- For 11 studies, it was unclear if the patients had delirium at baseline (Bucerius 2004; Caeiro 2004; Ely 2007; Hofsté 1997; Korevaar 2005; Leung 2007; Margiotta 2006; Ouimet 2007; Pisani 2007; Redelmeier 2008; Sheng 2006; Weed 1995).

  - In all of these studies the authors evaluated patients who 'developed' delirium, but they did not specifically state if any of the patients had existing delirium.

  - Two of these studies (Bucerius 2004; Hofsté 1997) included patients undergoing elective cardiac surgery and the GDG decided that this type of operation was unlikely to be carried out in patients with preoperative delirium.

  - Four studies (Ely 2007; Ouimet 2007; Pandharipande 2006; Pisani 2007) were carried out in ICU and the GDG considered that these patients were likely to have incident delirium only

- One study evaluated delirium severity (McCusker 2001); the authors reported that 73% of patients had prevalent delirium (although prevalent (versus incident) delirium was included as a risk factor in the multivariate analysis).

Method of delirium assessment

A number of validated instruments were used to evaluate delirium incidence or duration using DSM-IV or DSM-III-R criteria. The GDG considered that 27 studies had an adequate method of assessment; two had a partially adequate method (Levkoff 1992; Schor 1992); three had an inadequate method (Levkoff 1988; Redelmeier 2008; Yildizeli 2005) and one did not state the method (Weed 1995).

- Adequate method

  - Fifteen studies used the CAM (Ely 2007; Furlaneto 2006; Goldenberg 2006; Inouye 1993; Inouye 2007; Korevaar 2005; Leung 2007; Lin 2008; Margiotta 2006; Olin 2005; Pisani 2007; Ranhoff 2006; Rolfson 1999; Veliz-Reissmüller 2007; Zakriya 2002)

  - Two studies used the Organic Brain Syndrome (OBS) scale (Andersson 2001; Edlund 2001) (the study by Andersson 2001 used a modified version of this scale)
  - Two studies used the DRS (Böhner 2003; Caeiro 2004)
  - One study used the used the Intensive Care Delirium Screening Checklist (ICDSC) (Ouimet 2007)
  - One study used the CAM-ICU test with the Richmond Agitation Sedation Scale (RASS) (Pandharipande 2006)
  - One study used the Saskatoon Delirium Checklist (SDC) (Hofsté 1997)
  - Six studies assessed delirium based on clinical observations using DSM-IV, DSM-III-R or (Bucerius 2004; Kazmierski 2006; Pompei 1994; Rudolph 2007; Santos 2004; Sheng 2006).
  - Two studies (Levkoff 1992; Schor 1992) used the Delirium Symptom Interview (DSI) which assesses the domains of delirium specified in DSM III. The GDG considered this to be an adequate method.

- Inadequate
  - Three studies assessed delirium by retrospective chart review (Levkoff 1988; Redelmeier 2008; Yildizeli 2005)
  - The study by Weed (1995) did not report what diagnostic criteria were used to assess delirium, or what instrument was applied.

One study evaluated severity of delirium as an outcome measure (McCusker 2001). In this study, the authors developed in their group a Delirium Index (DI) based on the CAM criteria, which ranged from 0 to 21(maximum severity). This was compared with the Delirium Rating Scale which showed reasonably good correlation (Pearson correlation coefficient 0.84). However, the GDG regarded this as indirect evidence, and this was supported by the review of diagnostic test accuracy (chapter 6).

The GDG considered the three retrospective studies (Levkoff 1988; Redelmeier 2008; Yildizeli 2005) to be biased because the method of assessment was based on review of medical notes. The GDG agreed that the two studies (Levkoff 1992; Schor 1992), which used the DSM III (or methods based on DSM III) for assessment had an adequate method of assessment.

Confounders taken into account

Of the 35 cohort studies, 32 conducted multivariate analyses (Andersson 2001; Böhner 2003; Bucerius 2004; Caeiro 2004; Edlund 2001; Ely 2007; Furlaneto 2006; Goldenberg 2006; Hofsté 1997; Inouye 1993; Inouye 2007; Kazmierski 2006; Korevaar 2005; Leung 2007; Levkoff 1988; Levkoff 1992; Lin 2008; McCusker 2001; Ouimet 2007; Pandharipande 2006; Pisani 2007; Pompei 1994; Ranhoff 2006; Redelmeier 2008; Rolfson 1999; Rudolph 2007; Santos 2004; Schor 1992; Sheng 2006; Veliz-Reissmüller 2007; Yildizeli 2005; Zakriya 2002).

Three studies conducted only univariate analyses for the incidence of delirium: Margiotta 2006; Olin 2005; Weed 1995) and these studies were not considered further. Details of the factors included in the multivariate analysis are given in Appendix F.

We considered whether the cohort studies took account of particular confounders, either in the study design or the multivariate analysis. The GDG had identified, by consensus, four risk factors to be important: age, sensory impairment, polypharmacy and cognitive impairment. Following GDG discussion it was decided *post-hoc* to record whether the multivariate analyses included severity of illness or comorbidity, as well as polypharmacy.

Studies were summarised according to the number of key risk factors included in the multivariate analysis and the ratio of events to covariates (the GDG considered a ratio of 1 or less to be flawed and a ratio of 2 or 3 to be possibly confounded). We assumed that the key risk factors were the same for severity of delirium and duration of delirium. The following combinations were found:

- Confounders taken into account: all/most (4 or 3) of the important risk factors (RFs) taken into account in the multivariate analysis or held constant and a ratio of events to variables of 10 or more
    - Bucerius (2004) had a ratio of 39 (3 key RFs: age included in the analysis; cognitive impairment excluded because elective cardiac operations and polypharmacy constant because elective cardiac operations in older patients; missing key RF: sensory impairment)
    - Levkoff (1992) had a ratio of 23 (2-3 key RFs: age and cognitive impairment included in the analysis, and patients with severe sensory impairment were excluded; illness severity included. No systematic standardised method was used to detect cognitive impairment, with reliance on medical chart review)
    - McCusker (2001) had a ratio of 18 (3 key RFs: age, dementia, and sensory impairment included in the analysis; missing key RF: polypharmacy; comorbidity included)
    - Schor (1992) had a ratio of 10 (2-3 key RFs: age and cognitive impairment included in the analysis and patients with severe hearing or vision impairment excluded; missing key RF: polypharmacy; unstated scale for cognitive impairment)

- Possibly confounded: all/most of the important risk factors taken into account in the multivariate analysis but an insufficient ratio of events to variables
  - Ranhoff (2006) had a ratio of 7 (all 4 key RFs included in the analysis)
  - Böhner (2003) had a ratio of 7 (3 key RFs: age and cognitive impairment included in the analysis and polypharmacy constant because elective cardiac operations in older patients; missing key RF: sensory impairment)
  - Goldenberg (2006) had a ratio of 6 (3 key RFs included in the analysis – not sensory impairment)
  - Pandharipande (2006) had a ratio that ranged from 4 (66/17) to 7 (118/17) (3 key RFs: age, dementia, visual impairment)
    - The study reported the number with delirium for two subgroups: those who received antipsychotics (66/75 had delirium) and those who received anticholinergics (52/63); it is unclear if any patients had both drugs, therefore the number with delirium was considered to range from 66 to 118.
  - Veliz-Reissmüller (2007) had a ratio of 4 (3 key RFs: age and cognitive impairment included in the analysis and polypharmacy constant because elective cardiac operations in older patients; missing key RF: sensory impairment; inappropriate cut off point on MMSE scale for cognitive impairment)
  - Sheng (2006) had a ratio of 3 (3 key RFs included in the analysis – not polypharmacy)
  - 3 studies had ratio of 2:
    - Andersson (2001) (all 4 key RFs included in the analysis; comorbidity was also included)
    - Santos (2004) (3-4 key RFs: age and cognitive impairment included in the analysis; polypharmacy constant because elective cardiac operations in older patients; patients with severe sensory impairment excluded)
    - Inouye (1993) (3 key RFs included in the analysis; not polypharmacy; illness severity included)
- Possibly confounded: not enough of important risk factors taken into account in the multivariate analysis (2/4) but a sufficient ratio of events to covariates
  - Age and cognitive impairment

        - Rudolph (2007) had a ratio of 16 (1-2 RFs: age included in the analysis and patients with dementia (not mild cognitive impairment) were excluded)
    - Age and polypharmacy
        - Ouimet (2007) had a ratio of 19 (2 RFs: age included in the analysis and polypharmacy constant because patients in ICU; illness severity also included)
        - Redelmeier (2008) had a ratio of 200 (2 key RFs: age included in analysis and polypharmacy likely constant because surgical patients)
    - Cognitive impairment and polypharmacy
        - Lin (2008) had a ratio of 10 (2 RFs: patients with dementia excluded and polypharmacy constant because patients in ICU)
    - Cognitive impairment and sensory impairment
        - Inouye (2007) had a ratio of 10 (2 RFs: dementia and vision impairment included in analysis; Illness severity also included)

- Possibly confounded: not enough of important risk factors taken into account in the multivariate analysis (2/4) and not high enough ratio of events to covariates
    - Age and cognitive impairment
        - Hofsté (1997) had a ratio of 9 (2 key RFs: age included in analysis and cognitive impairment constant because elective cardiac operations)
        - Korevaar (2005) had a ratio of 4 (age and cognitive impairment included in the analysis)
        - Leung (2007) had a ratio of 3 (age and cognitive impairment included in the analysis)
        - Kazmierski (2006) had a ratio of 2 (2 key RFs included in analysis: age and cognitive impairment included in analysis)
    - Age and polypharmacy
        - Ely (2007) had a ratio of 8 (2 RFs: age included in the analysis and polypharmacy constant because patients in ICU; illness severity also included)
        - Rolfson (2003) had a ratio of 8 (age was constant due to narrow age range, and polypharmacy constant because elective cardiac operations in older patients)

        - Cognitive impairment and polypharmacy
        - Pisani (2007) had a ratio of 9 (cognitive impairment included in the analysis and polypharmacy constant because patients in ICU; illness severity also included)

- Probably confounded: not enough of important risk factors taken into account in the multivariate analysis (1/4), but did have a ratio of events to covariates of at least 10
    - Cognitive impairment
        - Furlaneto (*2006)* had a ratio of 15 (cognitive impairment included in the analysis)
        - Pompei (2002) had a ratio of 16 and 21(cognitive impairment included in the analysis; comorbidity also included)

- Probably confounded: not enough of the important risk factors taken into account in the multivariate analysis (1/4), and did not have high enough ratio of events to covariates
    - Age
        - Caeiro (2001) had a ratio of 7 (age included in the analysis)
        - Levkoff (1988) had a ratio of 6 (age included in the analysis)
        - Yildizeli (2005) had ratio of less than 1 (age included in the analysis)
    - Cognitive impairment
        - Zakriya (2008) had a ratio of 8 [patients with dementia were excluded but method of assessment not stated; illness severity also included (as American Society of Anesthesiologists, ASA grade)]

- Confounded: no important risk factors taken into account in the multivariate analysis (0/4) and did not have a high enough ratio of events to covariates
    - Edlund (2001) had a ratio 4 for incident delirium

The McCusker (2001) study reporting delirium severity used analyses at various times reflecting different states (repeated measures multivariate analyses, using the previous most recent severity score as a factor in the multivariate analysis). The GDG considered this to be an acceptable method.

Overall, the risk of bias was considered for each cohort study, and ratings were given of high, moderate and low quality, and biased/confounded.

- Six studies were judged to be biased and therefore not considered further:
    - Edlund (2001): no key risk factors
    - Furlaneto (2006): 57% prevalent delirium included
    - Levkoff (1988): inadequate method of delirium assessment; retrospective
    - Pompei 1994 (Yale): 56% prevalent delirium included
    - Redelmeier (2008): inadequate method of delirium assessment; retrospective
    - Yildizeli (2005): not enough patients for multivariate analysis (ratio less than 1); retrospective
- Twelve studies were given a low overall rating and were treated with caution (evaluated in sensitivity analyses) (Andersson 2001; Caeiro 2004; Inouye 1993; Kazmierski 2006; Korevaar 2005; Leung 2007; McCusker 2001; Pompei 1994 (Chicago); Santos 2004; Sheng 2006; Veliz-Reissmüller 2007; Zakriya 2008)
- Fifteen studies had a moderate rating; (Böhner 2003; Bucerius 2004; Goldenberg 2006; Ely 2007; Hofsté 1997; Inouye 2007; Levkoff 1992; Lin 2008; Ouimet 2007; Pandharipande 2006; Pisani 2007; Ranhoff 2006; Rolfson 1999; Rudolph 2007; Schor 1992)
- No studies had a high rating

7.4.3 Risk factors investigated by the cohort studies (multivariate analyses)

The following risk factors have been investigated in the included studies:

Patient characteristics

- Age (21 studies)
- Cognitive impairment and/or dementia (14 studies)
- Sensory impairment (7 studies)

- Polypharmacy (2 studies)
- Dehydration (5 studies)
- Severity of illness (5 studies)
- Comorbidity (4 studies)
- Sex (7 studies)
- Electrolyte disturbance (2 studies)
- Depression (6 studies)
- Infection (5 studies)
- Fracture on admission (1 study)
- Mobility (1 study)
- Continence (1 study)
- Constipation (no studies)
- Sleep deprivation (no studies)

Environmental

- Pre-hospital setting (3 studies)
- Hospital unit: ICU, surgery, medical, oncology, long-term care, mixed (1 study)
- Recent room change (1 study)
- Room type: private, semi-private, ward (1 study)
- Stimulation: based on the distance of the room from the nurses station (1 study)
- Same room (1 study)
- Single room (1 study)
- Surroundings not well lit (1 study)
- Surroundings sound too noisy/quiet (1 study)
- Radio/TV on (1 study)

- Clock/watch (1 study)
- Calendar (1 study)
- Personal possessions present (1 study)
- Wearing glasses (1 study)
- Using hearing aid (1 study)
- Family present (1 study)
- Isolation (because of infection risk) (1 study)

Procedural

- Type of surgery (5 studies)
- Iatrogenic interventions (2 studies)
- Physical restraint (2 studies)

7.4.4 Outcomes

The studies measured the following outcomes:

- Incidence of delirium
- Duration of delirium
- Severity of delirium

7.5 Results

The studies reported summary statistics with 95% confidence intervals as either odds, relative, hazard ratios or beta-coefficients. Beta coefficients were reported separately. These results were taken as reported from the primary papers.

7.5.1 Patient related risk factors

Setting

Pre-hospital setting as a risk factor for the incidence of delirium
Two studies included pre-hospital setting in their multivariate analysis (Andersson 2001, low; Schor 1992) and one study (Levkoff 1992) reported results separately for patients from long-term care and from the community, and also carried out a multivariate analysis in which pre-hospital long-term care was

included (the other factors were age, sex, pre-existing cognitive impairment and illness severity; and patients with severe sensory impairment were excluded).

The Andersson (2001) study (low rating) found no significant effect of sheltered housing relative to the person's own home, and Schor (1992) (moderate rating) found no significant effect of pre-hospital long-term care (the other risk factors were age, prior cognitive impairment, fracture on admission, sex, infection, pain (poorly controlled), neuroleptic use, and narcotic use). In neither case were data reported, although the Schor (1992) study reported the odds ratio adjusted for age and sex only - which is a low evidence rating - OR 2.54 (95%CI 1.38 to 4.67), and was statistically significant. The Levkoff (1992) study (moderate rating), however, found a statistically significant effect of long-term care on the incidence of delirium developing in hospital: OR 2.16 (95%CI 1.15 to 4.1).

The Levkoff (1992) study mostly analysed the data using separate analyses for the two pre-hospital groups of long-term care and the community, and as will be seen in subsequent risk factor analyses, there were large differences between the two groups. The GDG stated that dementia and comorbidity would likely be higher in people from long-term care settings.

<u>Setting as a risk factor for increased severity of delirium</u>
For severity of delirium, one large study (McCusker 2001: low; n=587 time dependent states) considered the effect of different hospital units, using a repeated measures multivariate analysis. At any given time, patients could be in long-term care, long-term care /medical, or in hospital wards (subdivided into general medical, oncology, surgery and ICU). Numbers of patients who had spent time in each unit were as follows:

- ICU (20/587 = 3%)
- Surgery (81/587 = 14%)
- General medical (281/587 = 48%)
- Oncology (20/587 = 3%)
- Long-term care (34/587 = 6%)
- Mixed long-term care/medical (151/587 = 26%)

Thus, we would expect some uncertainty around the results for ICU (3%), oncology (3%) and long-term care (6%). Results from the multivariate analysis (with medical ward as the reference) are reported in figure 7.1 (Appendix K) and show significant differences only for patients in ICU. However, this is likely to be of limited reliability because only a small proportion was in ICU. Furthermore, the GDG considered it likely that the ICU status was a proxy measure for polypharmacy and/or severity of illness, neither of which were included in the multivariate analyses.

<u>Summary of setting as a risk factor for delirium</u>
The evidence regarding the risk factor, 'long-term care setting prior to hospitalisation', is inconsistent for the incidence of delirium. The evidence is inconclusive for the effect of setting on the severity of delirium, although patients in ICU may be at higher risk than patients in medical wards.

Age

Seventeen studies presented data on age in their multivariate analyses, see table 7.6 (Andersson 2001 (low rating); Böhner 2003; Bucerius 2004; Caeiro 2004 (low); Ely 2007; Goldenberg 2006; Hofsté 1997; Kazmierski 2006 (low); Leung 2007 (low); Levkoff 1992; McCusker 2001 (low); Pandharipande 2006; Ranhoff 2006; Rudolph 2007; Santos 2004 (low); Schor 1992; Sheng 2006 (low)) (figures 7.2 and 7.3). Four other studies also included age as a risk factor in their multivariate analyses, but did not report any data (Korevaar 2005 (low rating); Inouye 1993 (low), Ouimet 2007 (moderate), Veliz-Reissmüller 2007(low). It was stated that age was not a significant risk factor in the studies by Ouimet 2007 and Inouye 1993.

One study carried out a 'Markov regression', which was a regression analysis that included the patient's cognitive state 24 hours previously. The study reported transitions to delirium and plotted graphically the probability of developing delirium versus age (Pandharipande 2006).

One of the studies investigated the duration of delirium (Ely 2007) (figure 7.4) and one investigated the severity of delirium (McCusker 2001; low) (figure 7.5); the rest evaluated incidence of delirium.

The standard error for the Böhner (2003) study was calculated from its p-value: confidence intervals were not reported for the odds ratio (but were for the beta coefficient).

Table 7.6: patient ages in 17 studies that conducted multivariate analyses

Study	Age range	Study	Age range
Bucerius	54-75 (+/- SD)	Caeiro	24-86
Rudolph	63-75	Schor	73-88 (+/-)
Santos	66-78	McCusker	76-90 (+/-)
Leung	66-78 (+/-)	Ranhoff	60-94
Ely	31-79	Sheng	65-95
Kazmierski	25-81	Levkoff institution	80-95 (+/-)
Hofste	29-83	Andersson	65-96
Bohner	NS	Goldenberg	66-98
Levkoff community	71-85 (+/-)	Pandharipande 2006	25-90 (graph)

+/- indicates that the range was calculated from the mean +/- one standard deviation

We note that, of these studies, nine were in patients admitted for surgery (Andersson 2001; Böhner 2003; Bucerius 2004; Goldenberg 2006; Hofsté 1997; Kazmierski 2006; Leung 2007; Rudolph 2007; Santos 2004), three were in patients admitted to ICUs (Ely 2007; Pandharipande 2006; Ranhoff 2006), three were conducted in patients from medical wards (Caeiro 2004; McCusker 2001; Sheng 2006), and the remaining two studies were in patients from both medical and surgical wards (Levkoff 1992; Schor 1992).

<u>Age as a risk factor for the incidence of delirium</u>

Fifteen studies investigated age as a risk factor for the incidence of delirium.

- Five studies evaluated age as a continuous variable (Andersson 2001, low; Leung 2007, low; Rudolph 2007; Santos 2004, low; Sheng 2006, low); the age range across all these studies was 63 to 96 years

- One study reported the probability of developing delirium as a function of age, between the ages of 25 and 90 years. Although the study reported the odds ratio for age as a continuous variable, this was not included in the analysis because of the non-linearity over the age range (Pandharipande 2006)

- Three studies evaluated age over 65 years versus age below 65 years (Böhner 2003; Caeiro 2004, low; Kazmierski 2006, low)

- One study evaluated age 70 years and over versus age below 60 years (Hofsté 1997)

- We note that the Hofsté (1997) study did not report the category 60 to 69 years in the multivariate analysis (and for the separate cognitive disorders analysis there are other categorical variables not reported). Therefore this study should be treated with caution for age as a risk factor.

- Four studies evaluated age over 80 versus age below 80 years (Goldenberg 2006 (age over 81); Levkoff 1992 community and institution; Ranhoff 2006; Schor 1992)

- The study by Bucerius (2004) evaluated three comparisons of categorical age variables (which we have inverted to allow for comparison with the other studies): over 70 versus under 50, over 70 versus 50-59 years, and over 70 versus under 60

The results are reported in Figures 7.2 and 7.3a (Appendix K), with a sensitivity analysis (excluding low quality studies) shown in figure 7.3b (Appendix K)

The sensitivity analysis in figure 7.3b showed no important differences compared with figures 7.2 and 7.3a, and so it was decided to use all the data.

For the age cut-off of 80 years, there was heterogeneity (assessed visually). However, the GDG noted that the mean age in the Levkoff (1992) institution group was 87.4 years and only 11.4% patients were younger than 80 years; suggesting that the age range may not have been large enough to allow conclusions to be derived .The Ranhoff (2006) study was the only one investigating the effect of age (on the incidence of delirium) that was conducted in an ICU setting; the GDG suggested that the effects of illness would be likely to overshadow the effects of age in this setting – the study had not included illness severity in the multivariate analysis, although it had taken account of polypharmacy. Following discussion, the GDG agreed that the effect of age over 80 years was best described by the other three studies.

The GDG wished to define a cut-off point for age as a risk factor and noted that the studies reported different age thresholds. Further information was provided by one moderate quality study (Pandharipande 2006), which reported the probability of developing delirium as a function of age. This probability showed a non-linear pattern across the age range 25 to 90 years. Between the ages of 25 and about 48 years there was a steady increase in the probability, then between 48 and 65 years the graph showed a plateau (same probability independent of age). Finally, above 65 years the probability increased rapidly. This study is the only one to demonstrate the importance of age 65 years as a cut off for age as a risk factor.

<u>Age as a risk factor: increased duration of delirium</u>

One small study (Ely 2007; n=47) investigated the effect of age as a continuous variable on the duration of delirium, for patients aged 31 to 79 years. We note that this study (with a moderate rating) was conducted in ICU in mechanically ventilated patients. There was no significant effect of age as a continuous variable on the duration of delirium (figure 7.4, Appendix K); OR 1.02 (95%CI 0.98 to 1.06).

<u>Age as a risk factor: increased severity of delirium</u>

One large study (McCusker 2001; low, n=444) investigated the effect of age as a continuous variable on the severity of delirium, for patients of mean age 83.3 years (SD 7.0). The effects of different risk factors are shown in figure 7.5 (Appendix K), reporting the beta coefficient representing the estimated difference in Delirium Index scores between the independent variable and the reference category. For age, as a continuous variable, there was no significant effect: beta coefficient 0.03 (95% CI -0.01 to 0.07).

Summary for age as a risk factor

Thus, the following summary can be given:

- For age as a continuous variable, the odds ratio for incidence of delirium ranged from 1.08 to 1.10. This means that for every year increase in age the odds of having delirium increases by a factor of 1.08 to 1.10.

Taking the 1.10 value, for a 10 year increase in age, the odds increases by $(1.10)^{10}$, which is 2.59. We note that the results are consistent over a range of studies, and are likely to be valid. The age range covered by the studies was 63 to 96 years.

- The odds ratio for delirium incidence for a cut-off point of age 65 years was 3.03 (95%CI 1.19 to 7.71) for the only study (Böhner 2003) that was not of low quality (this value was derived from the quoted beta coefficient of 1.11 (SE 0.468).

- Age was a significant risk factor for incidence of delirium for most (3/5) of the studies when a cut-off point of age 80 years was taken, with the OR ranging from 0.87 (95%CI 0.22 to 3.3) to 5.40 (95%CI 2.4 to 12.3). There appeared to be significant heterogeneity (assessed visually) amongst these studies, with two studies not showing a significant effect of age (Ranhoff 2006 and Levkoff 1992 institution (in patients who had come from a long-term care setting)), and three studies showing a similar significant odds ratio around 5.

    - The GDG noted that the mean age in the Levkoff (1992) institution group was 87.4 years and only 11.4% patients were younger than 80 years; suggesting that the age range was not large enough to allow conclusions to be derived.

    - The Ranhoff (2006) study was conducted in an ICU setting; the GDG suggested that the effects of illness would be likely to overshadow the effects of age in this setting, and noted that illness severity was not included in the multivariate analysis for this study, even though polypharmacy was.

- One moderate quality study (Pandharipande 2006) examined the variation across the age range 25 to 90 years, of the probability of developing delirium, which showed age 65 years to be a point above which the probability increased rapidly, and this was taken as the age cut-off.

- There was no significant effect of age as a continuous variable on the duration of delirium, over the range 31 to 79 years, in one small study (n=47) in mechanically ventilated patients in ICU; OR 1.02 (95%CI 0.98 to 1.06)

- There was no significant effect of age as a continuous variable on the severity of delirium, for patients of mean age 83.3 years (SD 7.0), in one large low quality study (n=444); beta coefficient 0.03 (95% CI -0.01 to 0.07).

Cognitive impairment and/or dementia

Fourteen studies evaluated cognitive impairment and/or dementia in their multivariate analyses (Böhner 2003; Goldenberg 2006; Inouye 1993, low; Inouye 2007; Kazmierski 2006, low; Korevaar 2005, low; Levkoff 1992; McCusker 2001, low; Pisani 2007; Pompei 1994, low; Ranhoff 2006; Schor 1992; Sheng 2006, low; Veliz-Reissmüller 2007, low) (figure 7.7). In the study by Pompei (1994), data from only one trial (the Chicago hospital) were reported because the Yale-New Haven hospital data was judged to be biased.

- Eight studies used an MMSE score:
  - below 18 cut off for patients at discharge (Ranhoff 2006)
  - below 21-24 cut off depending on education (Pompei 1994)
  - below 24 (Goldenberg 2006; Inouye 1993; Inouye 2007; Kazmierski 2006)
  - below 25 (Böhner 2003)
  - below 28 (Veliz-Reissmüller 2007)
- Three studies used IQCODE (Pisani 2007: above 3.3; McCusker 2001: above 3.5; Korevaar 2005: above 3.9) IQCODE
- Two studies did not state the assessment method (Schor 1992; Sheng 2006)
- One study (Levkoff 1992) stated that no systematic standardised method was used to detect cognitive impairment, with reliance on medical chart review, which would have led to underreporting

Of these studies, the GDG did not consider the definition of cognitive impairment to be reliable in the Veliz-Reissmüller (2007) and Levkoff (1992) studies, so these were not included in the analysis. Due to the low percentage (8%) of patients with dementia in the study by Sheng (2006) (table 7.7), the results from this study were also omitted from the analysis. The Ranhoff (2006) study was considered in sensitivity analyses because cognitive impairment was assessed at discharge, in association with activities of daily life measurements.

We note that of the remaining studies, three were in patients admitted for surgery (Böhner 2003; Goldenberg 2006; Kazmierski 2006), two were in patients admitted to ICUs (Pisani 2007; Ranhoff 2006), and the other studies were in patients from both medical/surgical wards (Inouye 2007; Korevaar 2005; McCusker 2001; Pompei 1994; Schor 1992).

Nine studies evaluated incidence of delirium (Böhner 2003; Goldenberg 2006; Inouye 1993; Kazmierski 2006; Korevaar 2005;Pisani 2007; Pompei 1994; Ranhoff 2006; Schor 1992) (figure 7.5, Appendix K). One of the studies investigated the severity of delirium (McCusker 2001) (figure 7.5, Appendix K); and one investigated persistent delirium (Inouye 2007) (figure 7.7, Appendix K) and one the rest.

We note that the Inouye (1993) study excluded people with severe underlying dementia.

The standard error for the Böhner (2003) study was calculated from its p-value: confidence intervals were not reported.

Table 7.7: cognitive impairment and/or dementia in 11 studies that conducted multivariate analyses

Study	Cognitive impairment / dementia	Study	Cognitive impairment / dementia
Goldenberg	69%	Pisani	31%
Inouye 1993	36%	Pompei-Chicago	37%
Inouye 2007	39%	Ranhoff	30%
Kazmierski	53% & 16%	Schor	19%
Korevaar	43%	Sheng	8%
McCusker 2001	60%	Bohner	Not reported

There was some heterogeneity (assessed visually) in figure 7.6a which was removed when only the higher quality studies were analysed (figure 7.6b, Appendix K), so the sensitivity analysis was considered more reliable. There was a large significant effect of cognitive impairment on the risk of delirium. The odds ratio ranged from 6.3 (95%CI 2.9 to 13.8) to 11.5 (95%CI 6.1 to 20.1) with an apparent outlier (Böhner 2003) at OR 28.0 (p value 0.001; beta coefficient 3.33 (SE 0.927).

Cognitive impairment and/or dementia as a risk factor for the incidence of persistent delirium

One moderate quality study investigated the effect of cognitive impairment on the incidence of persistent delirium (Inouye 2007) in 491 patients. We note that these results are from a subpopulation of patients with delirium (n=443). Cognitive impairment was a significant risk factor for persistent delirium (figure 7.7, Appendix K); OR 2.3 (95%CI 1.4 to 3.7).

Cognitive impairment and/or dementia as a risk factor for increased severity of delirium

One large low quality study (McCusker 2001; n=444) investigated the effect of dementia (IQCODE score at least 3.5). Figure 7.5 (Appendix K) shows a significant effect; the beta coefficient for the mean difference in delirium severity score is 1.13 (95% CI 0.58 to 1.68).

Summary for cognitive impairment/dementia as a risk factor

- Restricting the analysis to the studies that were of higher quality, there was a large significant effect of cognitive impairment on the risk of delirium. The odds ratio ranged from 6.3 (95%CI 2.9 to 13.8) to 11.5 (95%CI 6.1 to 20.1) with an apparent outlier (Böhner 2003) at OR 28.0 (p value 0.001; beta coefficient 3.33 (SE 0.927)).

- For persistent delirium, the odds ratio was 2.30 (95% CI 1.41 to 3.74). We note that these results are from a subpopulation of patients with delirium.

- There was a statistically significant effect of cognitive impairment on the severity of delirium; the beta coefficient for the mean difference in severity of delirium was 1.13 (95% CI 0.58 to 1.68) in one large low quality study.

Sensory impairment

Seven studies included sensory impairment in their multivariate analyses (Andersson 2001, low; Inouye 1993, low; Inouye 2007; McCusker 2001, low; Ranhoff 2006; Sheng 2006, low; Schor 1992).

<u>Sensory impairment as a risk factor for incidence of delirium</u>
Two studies presented data on vision impairment in their multivariate analyses (Andersson 2001 – low; Inouye 1993 - low). One other study also evaluated impaired vision as a risk factor in multivariate analysis, but did not report the non-significant results (Sheng 2006, low), and another study (Schor 1992) carried out an analysis adjusted for age and sex for each of vision and hearing loss. Since this Schor (1992) analysis included only age as a key risk factor, we gave it a low quality rating. Results for this study were included in Figure 7.8 (Appendix K) for vision impairment; hearing impairment had a non significant adjusted odds ratio of 1.62 (95% CI 0.85 to 3.06).

- In Andersson (2001) (low rating; n=457 patients), 31% of the surgical patients had vision impairment and 39% had hearing impairment.

- In Inouye (1993) (low rating; n=107), 6% of patients in the medical wards had vision impairment and 54% hearing impairment.

- In Ranhoff (2006) (moderate rating; n=401), 29% of the ICU patients had vision impairment (hearing impairment was not reported).

- In Schor (1992) (low rating for this risk factor; n=291), 33% of patients (in medical and surgical wards) had vision impairment and 21% hearing impairment

- In Sheng (2006) (low rating; n=156), 18% of the patients (in medical wards) had vision impairment (hearing impairment was not reported)

The proportion of only 6% in the Inouye (1993) study is considered likely to lead to inaccuracy. In both the Andersson (2001) and Inouye (1993) studies, the authors reported results for impaired vision only; hearing impairment was included in their multivariate analyses, but the non-significant results were not reported.

Figure 7.8 (Appendix K) shows a significant effect of vision impairment on the incidence of delirium. In the absence of the low quality studies, the remaining large study (Ranhoff 2006; n= 401) showed a small effect for patients in ICU: OR 1.70 (1.01 to 2.85). We note that this study did not define what was meant by vision impairment.

<u>Sensory impairment as a risk factor for incidence of persistent delirium</u>
One large, moderate rated study included vision impairment as a risk factor (Inouye 2007) in 443 patients; 38% of patients in the medical wards had vision impairment (hearing impairment was not reported). There was a significant effect (figure 7.9, Appendix K), OR 2.1 (95% CI 1.3 to 3.2).

<u>Sensory impairment as a risk factor for increased severity of delirium</u>
One large low quality study (McCusker 2001; n=444) investigated the effect of sensory impairment; 20% of the patients in the medical wards were reported to have vision/hearing impairment. Figure 7.5 (Appendix K) shows there was no significant effect; the beta coefficient for the mean difference in delirium severity score is 0 (95% CI -0.63 to 0.63).

<u>Summary for sensory impairment as a risk factor</u>

- Restricting the analysis for delirium incidence to the study that was of higher quality (Ranhoff 2006), this large ICU study showed a small effect of vision impairment: OR 1.70 (1.01 to 2.85). We note that this study did not define what was meant by vision impairment.

- For persistent delirium, there was a significant effect in a study that defined vision impairment carefully; OR 2.1 (95% CI 1.3 to 3.3). We note that these results are from a subpopulation of patients with delirium.

- The beta coefficient for the mean difference in severity of delirium for vision impairment was not significant in one large low quality study: 0.0 (95% CI -0.63 to 0.63)

- There was very limited evidence that hearing impairment was not an important risk factor for delirium incidence from low quality studies

Polypharmacy

<u>Polypharmacy as a risk factor for incidence of delirium</u>
Two studies presented data on the number of drugs as a risk factor for the incidence of delirium in their multivariate analyses (Goldenberg 2006; Ranhoff 2006). In neither case was illness severity or comorbidity included in the multivariate analyses.

In the study by Goldenberg (2006), the use of more than three medications (other than vitamins) was defined to represent multiple medication use, with 87% polypharmacy use in this sample. In the study by Ranhoff (2006), the authors evaluated the maximum concurrent number of drugs (including laxatives) as the following dichotomous variable: 7 or more drugs versus fewer than 7. The mean number of drugs used was 8.5 (SD 3.4) in patients with prevalent delirium, 8.0 (SD3.2) in patients with incident delirium, and 7.3 (SD 3.1) in patients without delirium. These studies both had moderate ratings. We note that the small study (n=77) by Goldenberg (2006) was in patients admitted for surgery, whereas the large study (n=401) by Ranhoff (2006) was conducted in ICU patients, a setting in which patients are likely to receive multiple medications. Figure 7.10 (Appendix K) shows a significant effect of polypharmacy on the incidence of delirium for both studies, but the confidence interval is very wide for the study with a cut-off point of 3 drugs.

<u>Summary for polypharmacy as a risk factor</u>

There was little evidence on polypharmacy as a risk factor.

The odds ratio was 33.60 (95% CI 1.9 to 591.6) in the study by Goldenberg (2006), and 1.9 (95% CI 1.1 to 3.2) in the study by Ranhoff (2006).

We note that 87% of the patients in the study by Goldenberg (2006) had taken more than 3 medications.

The GDG stated that more than 7 drugs in an ICU setting was not a useful clinical risk factor to assess.

Dehydration

<u>Dehydration as a risk factor for incidence of delirium</u>

A widely accepted laboratory measure of dehydration is the disproportionate rise in blood urea nitrogen (BUN) to creatinine. This was measured in two studies (Inouye 1993, low; Pisani 2007, moderate). Three other studies (Kazmierski 2006, low; Korevaar 2005, low; Santos 2004, low) recorded the blood urea content only; this measure is not considered to have high specificity for dehydration.

Three studies presented data on dehydration as a risk factor for the incidence of delirium in their multivariate analyses (figure 7.11, Appendix K). All of these studies had low quality ratings. We note that the study by Santos (2004) was in patients admitted for surgery, and the studies by Inouye (1993) and Korevaar (2005) were in medical wards.

- In the study by Inouye (1993), a baseline blood urea nitrogen/creatinine ratio of 18 or more was used as an index of dehydration; 67% in the group with delirium were dehydrated compared with 39% in the group without delirium (data calculated) [OR 2.02 (95% CI 0.72 to 5.64)]

- In the study by Korevaar (2005), the mean baseline urea nitrogen (mmol/l) concentration was 15.9 mmol/l (SD 13.6) in patients with delirium after acute admission compared with 10.6 mmol/l (SD 6.2) in patients without delirium [1.10 (95% CI 1.02 to 1.18)]

- In the study by Santos (2004), the pre-operative blood urea level ranged from 15-127 mg/dl; it was on average, 50.63 mg/dl (SD 23.26) in patients with delirium, and 41.85 (SD 14.39) in patients without delirium[OR 1.03 (95% CI 1.01 to 1.05)

In addition, two studies included dehydration as a risk factor in their multivariate analyses, but did not report the non-significant results (Kazmierski 2006, low; Pisani 2007, moderate).

- In the study by Kazmierski (2006), 5/30 (17%) of delirious patients had a pre-operative serum urea concentration greater than 50 mg/dl compared to 6/230 (7%) in patients without delirium; 8% overall

- In the study by Pisani (2007), 148/214 (69%) patients with delirium, and 54/90 (60%) patients without delirium, had a ratio of serum urea nitrogen to creatinine greater than 18 (measured in the first 48 hrs of ICU admission).

<u>Summary of dehydration as a risk factor</u>

- The GDG stated that a urea/creatinine ratio of 18 is difficult to interpret and depends on the units used (e.g. mmol/l), and a high urea level is not specific for dehydration

- One low quality study recorded the outcome representative of dehydration (urea/creatinine ratio) and the confidence interval was too wide to determine if dehydration was a risk factor for delirium.

Severity of illness

<u>Illness severity as a risk factor for incidence of delirium</u>
Three studies presented data on illness severity as a risk factor for the incidence of delirium in their multivariate analyses: Inouye (1993) (low), Levkoff (1992) (moderate) and Ouimet (2007) (moderate) (figure 7.12, Appendix K); the Ouimet (2007) study was conducted in ICU. A further ICU study included illness severity as a risk factor in their multivariate analysis, but the non-significant results were not reported (Pisani 2007, moderate). In none of the studies were polypharmacy or comorbidity included in the multivariate analyses.

- In the study by Inouye (1993), a composite score defined by a nurse rating of ‘severe’ or an Acute Physiology and Chronic Health Evaluation (APACHE) II score of more than 16 was considered to represent severe illness. In this study, 44% in group with delirium and 10% in group without delirium had ‘severe illness’ (data calculated). This study was conducted in a medical ward.

- The Ouimet (2007) study in ICU also used the APACHE II score (0 to 71 maximum possible) as a continuous variable; the mean score at baseline was 16.5 (SD 8.2), range 0 to 59

- The APACHE II score was also used in the Pisani (2007) study; the mean score was 24.7 (SD 6.1) in patients with delirium compared to 20.0 (SD 5.6) in patients without delirium

- In the study by Levkoff (1992), an illness severity score was calculated by summing the severity scores assigned to 15 medical conditions; they ranged from 1 for conditions that were not likely to have an impact on the process of care, to 4 for conditions that were imminently life threatening (baseline data were not reported). This study was conducted in both medical and surgical wards. The GDG noted that this was an unvalidated scale, and treated these results with caution.

For the two studies using validated scales (Inouye 1993, low and Ouimet 2007), there was a significant effect of illness severity on the incidence of delirium. The results from the Levkoff 1992 study were considered to be paradoxical by the GDG, and they noted that this study used an unvalidated scale, The GDG decided to remove this study and the low quality one (Inouye 1993) in a sensitivity analysis (not shown). The remaining very large study (n=764), Ouimet 2007, showed a significant effect of illness severity as a continuous variable: OR 1.25 (95%CI 1.23 to 1.27) per 5 point increase in APACHE II score, or 1.049 (95%CI 1.028 to 1.070) per point increase, which is a fairly large effect. The former means that for every 5 points on the APACHE II scale, the odds of delirium increases by 1.25. We note that this remaining study was conducted in ICU patients.

<u>Illness severity as a risk factor for increased duration of delirium</u>
One small, moderate quality study conducted in mechanically ventilated patients in ICU (Ely 2007; n=53) examined the effect of illness severity on the duration of delirium. Illness severity was determined using the APACHE II score, and this had mean scores of 26.8 (SD 8.0) to 27.8 (SD 5.3).

Results are shown in figure 7.13(Appendix K), and there is no significant effect of illness severity as a continuous factor on the duration of delirium [OR 0.97 (95% CI 0.90 to 1.07)].

<u>Summary of illness severity as a risk factor</u>
The following summary can be given:

- Illness severity had a significant effect on the incidence of delirium in one large study conducted in ICU; for APACHE II scores as a continuous variable, the odds ratio was 1.25 (95% CI 1.23 to 1.27) per 5 point increase, or 1.049 (95%CI 1.028 to 1.070) per point increase.

- One low quality non-ICU study showed severity of illness to be a risk factor for the incidence of delirium; patients were assessed to have severe illness if they had an APACHE II score of more than 16

- Illness severity did not show a significant effect on the duration of delirium in one small study in mechanically ventilated patients in ICU

Comorbidity

Comorbidity as a risk factor for incidence of delirium

Two studies presented data on comorbidity (Andersson 2001; Pompei 1994); both had a low quality rating.

- In Andersson (2001), 10% of patients with 'acute confusional state' had four or more diseases compared to 1% of patients without acute confusional state.

- In Pompei (1994), we considered the number of Major Diagnostic Categories (MDCs) to be indicative of comorbidity. MDCs related to a major body system (e.g. circulatory or respiratory system), or conditions that impact on more than one body system (e.g. sepsis or major trauma) (a patient with hypertension, ischaemic heat disease, and aortic vascular sclerosis would have three diagnoses but only one MDC). The mean number of MDCs in patients with delirium was 4.2 (SD 1.6) and 2.9 (SD 1.5) in patients without delirium.

We note that the study by Andersson (2001) was in patients admitted for surgery, and the study by Pompei (1994) was conducted in patients from both surgical and medical wards. In neither study was polypharmacy or illness severity taken into consideration in the analysis.
There was a significant effect of comorbidity on delirium incidence (figure 7.14, Appendix K).

Comorbidity as a risk factor for incidence of persistent delirium

One large, moderate quality study analysed comorbidity as a risk factor (Inouye 2007) in 443 patients. The study was conducted in patients in medical wards, of whom 29% had a Charlson Comorbidity score of 4 or more, with a mean baseline score of 2.7 (SD 2.1); the study did not include illness severity or polypharmacy in the analysis. There was a significant effect of comorbidity on the incidence of persistent delirium (figure 7.15(Appendix K)): OR 1.7 (95%CI 1.1 to 2.6).

Comorbidity as a risk factor for increased severity of delirium

One large, low quality study (McCusker 2001; n=444) investigated the effect of comorbidity on the severity of delirium; the study did not include illness severity or polypharmacy in the analysis. The study was conducted in patients in medical wards, for whom the mean baseline Charlson Comorbidity score was 2.7 (SD 2.0).
Figure 7.5 (Appendix K) shows no significant effect; the beta coefficient for the mean difference in delirium severity score is 0.09 (95% CI -0.03 to 0.21).

Summary of comorbidity as a risk factor

- Both studies that evaluated incidence of delirium had a low rating, and their results should be treated with caution, but both showed a significant effect of comorbidity on delirium incidence

- For persistent delirium, there was a significant effect of comorbidity (as measured by the Charlson comorbidity index) in a large moderate quality study; OR 1.7 (95% CI 1.1 to 2.6). We note that these results are from a subpopulation of patients with delirium

- In one large, low quality study, the beta coefficient for the mean difference in severity of delirium for comorbidity (as measured by the Charlson comorbidity index) was not significant: 0.09 (95% CI -0.03 to 0.21)

Sex (gender)

Sex as a risk factor for incidence of delirium
Three studies presented data on sex as a risk factor for the incidence of delirium in their multivariate analyses (Hofsté 1997; Levkoff 1992; Schor 1992) (figure 7.16, Appendix K). Proportion of male patients in each study is shown in table 7.8. All studies had a moderate quality rating (Hofsté 1997). In addition, four studies included sex as a risk factor in multivariate analyses, but the non-significant results were not reported (Andersson 2001 (low); Inouye 1993 (low); Kazmierski 2006 (low); Rudolph 2007 (moderate).

The studies were conducted in surgical patients (Andersson 2001; Kazmierski 2006; Hofsté 1997; Rudolph 2007), and medical/surgical patients (Inouye 1993; Levkoff 1992; Schor 1992).

Table 7.8: percentage of males in studies that conducted multivariate analyses

Study	Male	Study	Male
Schor	33%	Inouye	46%
Andersson	34%	Rudolph	53%
Levkoff - community	29%	Hofsté	73%
Levkoff- institution	35%	Kazmierski	76%

Summary of sex as a risk factor

- The odds ratio for male sex ranged from 0.4 (95% CI 0.2 to 0.8) to 4.9 (95%CI 1.6 to 15.3).

- There was heterogeneity (visually assessed) amongst these studies with one study showing a significant effect of the risk factor, male sex, one study showing a protective effect of male sex and one study showing a non-significant effect (Levkoff 1992) (community and institutional settings combined).

- The evidence was unable to show if sex is a clinically important risk factor.

Electrolyte disturbance

One low quality study presented data on electrolyte disturbance as a risk factor for the incidence of delirium in surgical patients in their multivariate analysis (Zakriya 2008) (figure 7.17, Appendix K). In addition, one study included electrolyte disturbance as a risk factor in multivariate analysis, but the non-significant results were not reported (Korevaar 2005). Both studies had a low quality rating.
The study by Zakriya (2008) considered abnormal serum sodium (Na+) (below 135 or above 148 mEq/l) to be indicative of electrolyte disturbance. Overall, 22% of the patients had abnormal serum sodium (data not reported for patients with and without delirium).

Due to the low rating of this study, the results should be treated with caution.

<u>Summary</u>
There was low quality evidence to suggest that electrolyte disturbance is a risk factor for delirium, but the absence of other important risk factors in the analysis made this uncertain; OR 2.40 (95% CI 1.09 to 5.27)]

Depression

<u>Depression as a risk factor for incidence of delirium</u>
Four studies presented data on depression as a risk factor for the incidence of delirium in their multivariate analyses (Böhner 2003; Inouye 1993; Kazmierski 2006; Pompei 1994) (figure 7.18, Appendix K). The study by Böhner (2003) had a moderate rating; the three other studies had low ratings. Two further studies included depression as a risk factor in multivariate analyses, but the non-significant results were not reported (Leung 2007 (low); Pisani 2007 (moderate).

We note that these studies were conducted in all settings: surgical patients (Böhner 2003; Kazmierski 2006; Leung 2007) medical/surgical wards (Inouye 1993; Pompei 1994) and ICU patients (Pisani 2007).

- In the study by Böhner (2003), a score of more than 8 using the Hamilton Depression Scale was indicative of depression; patients with delirium had a mean score of 8.16 (5.50) and patients without delirium had a mean score of 5.32 (5.52)

- In the study by Inouye (1993), depressive symptoms were considered present if the Geriatric Depression Score was 8 or more; 63% in the group with delirium and 44% in the group without delirium were depressed at baseline (data calculated).

- The method of defining depression was not reported in the study by Kazmierski (2006); 13% in the group with delirium, and 5% in the group without delirium had major depression.

- In the study by Pompei (1994), a score of 5 or more using the short form of the Yesavage Geriatric Depression scale was considered indicative of depression; of the Chicago sample, 41% with delirium and 17% without delirium were depressed

- In the study by Leung (2007), the authors evaluated depression using the Geriatric Depression Score: 12% had a score of 6 or higher

- The study by Pisani (2007) reported that 33% of the patients with delirium had a history of depression compared with 16% of patients without delirium (the scale used to measure depression was not reported).

The standard error for the Böhner (2003) study was calculated from its p-value: confidence intervals were not reported for the odds ratio.

The GDG noted that the scales used to measure depression were not diagnostic tools for that condition, and the cut-off points were not necessarily appropriate. The GDG also noted that in these studies, only Inouye (1993) also included illness severity in the multivariate analysis, and there was likely to be some confounding by physical illness. Thus, although there appeared to be a significant effect of depression as a risk factor for delirium, the GDG was not confident in this result. Considering only the higher quality study (Böhner 2003), the effect was just non-significant; OR 2.43 (95%CI 0.93 to 6.35) or beta coefficient 0.89 (SE 0.483; p=0.066).

Summary of depression as a risk factor

Although there appeared to be a significant effect of depression on the incidence of delirium, the majority of the studies were low quality, and there was likely to be some confounding. Restricting the analysis for delirium incidence to the study that was of higher quality (Böhner 2003), this moderate sized study showed an almost significant effect of depression OR 2.43 (95%CI 0.93 to 6.35) or beta coefficient 0.89 (SE 0.483). The GDG considered that even this result could be confounded by physical illness and was not confident in its validity.

Infection

Infection as a risk factor for incidence of delirium

Three studies presented data on infection as a risk factor for the incidence of delirium in their multivariate analyses (Lin 2008; Santos 2004; Schor 1992). Two studies had a moderate rating (Lin 2008; Schor 1992), and one had a low rating (Santos 2004). One other study included infection as a risk factor in the multivariate analysis, but the non-significant results were not reported (Sheng 2006 (low).

We note that these studies were conducted in all settings: surgical patients (Santos 2004), medical/surgical wards (Schor 1992; Sheng 2006) and ICU patients (Lin 2008).

The study by Lin (2008) reported that 80% of patients with delirium had sepsis (defined by the American College of Chest Physicians and the Society of Critical Care Medicine) and 57% without delirium had sepsis. The study by Santos (2004) reported that 19% patients with delirium and 3% of patients without delirium had post-operative pneumonia. The study by Schor (1992) reported that 37% with delirium and 17% without delirium had symptomatic infection. The study by Sheng reported that 15% of the patients with delirium had urinary tract infection compared to 4% of patients without delirium.

Figure 7.19 (Appendix K) shows that infection is a significant risk factor for delirium, although the confidence intervals are wide. A sensitivity analysis without the low quality study (Santos 2004) makes little difference.

Infection as a risk factor for increased duration of delirium

One small, moderate quality study in mechanically ventilated patients in ICU patients evaluated infection as a risk factor for the duration of delirium (Ely 2007). The study reported that, overall, 15% had sepsis and 23% had pneumonia. Figure 7.20 (Appendix K) shows no significant effect of infection on the duration of delirium, although the CI is wide [OR 1.73 (95% CI 0.57 to 5.28)] in this small study.

Summary of infection as a risk factor

- Three moderate quality and one low quality studies showed a similar trend, indicating that infection is a risk factor for delirium, despite the different types of infection evaluated; the odds ratio ranged from 2.96 (95%CI 1.42 to 6.16) to 6.36 (95%CI 1.24 to 32.71).

- Evidence from one small study mechanically ventilated patients in ICU showed no significant relationship between infection and duration of delirium.

Fracture on admission

One moderate quality study in 291 patients (Schor 1992) included fracture on admission as a risk factor for delirium. The study did not report what type of fractures were found, but there were 8.3% of patients with a fracture (8.3% of patients were also admitted to orthopaedic surgery). This is a relatively small percentage so there is likely to be some inaccuracy in the results. There was a significant effect of fractures on admission on the incidence of delirium (figure 7.21, Appendix K); OR 6.57 (95%CI 2.23 to 19.33).

This conclusion was supported by a second study (Andersson 2001, low quality), which showed that emergency hip fracture surgery was a significant risk factor for delirium incidence, compared with elective surgery for knee arthritis or hip arthritis (see procedural risk factors, section 7.5.3); OR 4.74 (95%CI 1.76 to 12.80) (see figure 7.21, Appendix K).

Summary of fracture as a risk factor

In summary, there was a significant effect of fractures on admission on the incidence of delirium in a single study, but there is some uncertainty associated with the effect; OR 6.57 (95%CI 2.23 to 19.33). The conclusion was supported by evidence from a low quality study comparing emergency hip fracture surgery with elective surgery for knee or hip arthritis.

Immobility

One low quality study included immobility (ability to walk without aid before admission) as a risk factor for the incidence of delirium in multivariate analysis, but the non-significant results were not reported (Andersson 2001). This study had a low rating. The study reported that 29% of patients with delirium were able to walk without an aid before admission compared to 46% of patients without delirium.

Summary

There is a lack of evidence on immobility as a risk factor for the incidence of delirium.

Incontinence

One low quality study included urinary and faecal incontinence as risk factors for the incidence of delirium in multivariate analysis, but the non-significant results were not reported (Sheng 2006 (low)). In this study 31% of patients with delirium and 13% of patients without delirium had urinary incontinence, and 23% with delirium and 8% without delirium had faecal incontinence.

Summary

There is a lack of evidence on continence as a risk factor for the incidence of delirium.

7.5.2 Environmental risk factors

One low quality study presented various environmental factors in their multivariate analysis of delirium severity (McCusker 2001). This study reporting delirium severity used analyses at various times reflecting different states (repeated measures multivariate analyses, using the previous most recent severity score as a factor in the multivariate analysis). The proportions of each of these states as a function of the number of different states for that variable are given below.

Some of the measures are subjective: for example, the research assistant decided whether the patient's surroundings were too noisy or whether the room was well lit. Other risk factors were more objective: e.g. whether or not various orientation aids were present and whether physical restraints were used. The study reported that the inter-rater reliability was assessed for these environmental observations in 29 patients and 75-100% agreement was found.

- Recent room change (173/617 = 28%)
- Stimulation: based on the distance of the room from the nurses station: high (105/573 = 18%), moderate (243/573 = 42%), low (225/573 = 39%)
- In same room (403/590 = 68%)
- Single room (124/509 = 24%)
- Surroundings' not well lit (61/504 = 12%)
- Surroundings' too noisy/quiet versus normal (159/421 = 38%)
- Radio/TV on (72/513 = 14%)
- Clock/watch absent versus present (294/585 = 50%)
- Calendar absent versus present (430/498 = 86%)
- Personal possessions absent versus present (421/538 = 78%)
- Not wearing glasses (375/587 = 64%)
- Not using hearing aid (433/470 = 92%)
- Family absent when carrying out assessment versus present (426/558 = 76%)
- In isolation because of screening for infection control (52/490 = 11%)

The results of the multivariate analyses are reported in figures 7.22 to 7.24(Appendix K). Most environmental risk factors showed no significant effect on the severity of delirium, but there was reported to be a significant effect for the following:

- Greater number of room changes
- Absence of a clock or watch
- Not wearing reading glasses

The GDG noted that in the UK, however, the number of moves is often influenced by management, rather than clinical reasons, and commented that it was unclear why the patients had been moved in this study.

The study also carried out exploratory analyses and noted two statistically significant interactions:

- The number of room changes was affected by the level of stimulation: a higher number of room changes had a strong impact on the severity of delirium only if the patient was in a room with high stimulation
- Moderate stimulation had a greater impact on patients in a unit with mixed medical and long-term care patients than in a medical ward

However, the authors stated that a large number of interactions were tested so that these results should be interpreted with caution.

<u>Summary of environmental risk factors for the severity of delirium</u>

- In one large, low quality study, the beta coefficient for the mean difference in severity of delirium was significant for the following factors:
    - The number of room changes: beta coefficient 0.37 (95% CI 0.04 to 0.70)
    - The absence of a clock or watch: beta coefficient 0.41 (95% CI 0.04 to 0.78)
    - Not wearing reading glasses: beta coefficient 0.82 (95% CI 0.45 to 1.19)
- In one large, low quality study, the beta coefficient for the mean difference in severity of delirium did not appear to be significant for the following factors: level of stimulation, single room, surroundings not well lit, surroundings too noisy or quiet, radio/TV on, calendar absent, absence of personal possessions, not using a hearing aid, family member present.
- We note that this study also controlled for age, dementia, baseline delirium severity; age, dementia, comorbidity, and visual or hearing impairment.

7.5.3 Procedural risk factors

Type of surgery

Five studies evaluated surgery as a risk factor for the incidence of delirium in their multivariate analyses (Andersson 2001; Bucerius 2004; Rolfson 1999; Rudolph 2007; Veliz-Reissmüler 2007) (figure 7.25, Appendix K). Two studies had a low rating (Andersson 2001; Veliz-Reissmüler 2007); the remaining studies had a moderate rating. Three of these studies evaluated cardiac surgery. None of the studies included illness severity in their multivariate analyses, although the Andersson (2001) study included comorbidity.

- The study by Bucerius (2004) compared patients who underwent beating heart surgery (no cardiopulmonary bypass) with those who underwent bypass (conventional) surgery. [OR 0.47 (95% CI 0.32 to 0.69)

- The study by Veliz-Reissmüller (2007) compared patients who underwent valve operation plus coronary bypass grafting (CABG) with CABG only. [OR 3.25 (95% CI 0.98 to 15.50).

- The study by Rolfson (1999) evaluated the duration of cardiopulmonary bypass (minutes) [OR 1.02 (95% CI 1.00 to 1.04).

- The GDG suggested that differences in the type of operation may be a proxy for illness severity.

Figure 7.26 (Appendix K) presents the results for two studies: one low quality study evaluated the risk of delirium in emergency hip fracture surgery patients versus patients admitted for elective surgery for knee arthritis or hip arthritis (Andersson 2001). The GDG concluded that this risk factor was connected with the underlying condition (i.e. hip fracture), rather than the type of surgery. [OR 4.74 (95% CI 1.76 to 12.78)]

One moderate quality study compared vascular surgery with all other surgery (abdominal, orthopaedic, genitourinary, thoracic and other) (Rudolph 2007), and showed that vascular surgery puts the patient at greater risk of delirium than other forms of surgery. [OR 2.70 (95% CI 1.72 to 4.24)]

The GDG stated that vascular surgery may be a proxy for other factors, such as undiagnosed vascular dementia or cerebral damage.

<u>Summary of surgical procedural factors as risk factors for delirium incidence</u>

- One moderate quality study showed a significant protective effect on the incidence of delirium for beating heart surgery compared with conventional bypass surgery.

- One moderate quality study showed that vascular surgery was a significant risk factor for delirium incidence, compared with other types of (non-cardiac) surgery.

- One moderate quality study showed a borderline significant effect of cardiopulmonary bypass time as a risk factor.

- None of the studies included illness severity in their multivariate analyses and the GDG concluded that the effects were likely to be a proxy for illness severity.

Iatrogenic interventions and medical restraint

Iatrogenic interventions

Two studies evaluated iatrogenic interventions as risk factors for the incidence of delirium in their multivariate analysis (Andersson 2001, low; Ranhoff 2006)

Both studies evaluated if a fitted bladder catheter was a risk factor. In the study by Ranhoff (2006), 81% of patients started to have prevalent delirium, and 80% of patients with incident delirium, used a bladder catheter (data were not reported for Andersson 2001). The study by Andersson (2001) did not report the non-significant results for the use of bladder catheter for emergency surgery patients in their multivariate analysis.

The study by Andersson (2001) was conducted in surgical patients and had a low rating while the study by Ranhoff (2006) was conducted in ICU patients and had a moderate rating.

- Due to the low rating of the Andersson (2001) study, the results for this study should be treated with caution.

- The GDG noted that the risk factor examined in the Ranhoff (2006) study was in-situ bladder catheter in ICU, rather than a bladder catheter being introduced, but they found the clinical interpretation of this study difficult. [OR 2.70 (95% CI 1.44 to 5.05)] (figure 7.27, Appendix K).

Medical restraint

One low quality study presented data on medical restraint in their multivariate analysis for the severity of delirium (McCusker 2001; figure 7.29, Appendix K). Medical restraint was stated to include intravenous and oxygen tubing, and occurred in 320/658 (49%) patient states. This was a significant risk factor; beta coefficient 0.41 (95% CI 0.04 to 0.78).

Physical restraint

Two studies presented data on physical restraint in their multivariate analyses (Inouye 2007; McCusker 2001) (figures 7.28 and 7.29, Appendix K). The Inouye (2007) study was of moderate rating, but the McCusker (2001) study was considered to be of low quality; both were conducted in medical wards. In the Inouye (2007) study, restraint use during delirium occurred in 15% of the patients. In the McCusker (2001) study, physical restraint was examined as a risk factor for delirium severity and occurred in 303/658 (44%) patient states; more detailed information was not reported.

Both studies reported a significant effect of physical restraint on delirium persistence (OR 3.20 (95%CI 1.93 to 5.29) and the severity of delirium (beta coefficient 1.24 (95% CI 0.91 to 1.57)).

- For persistent delirium, the odds ratio was 3.2 (95% CI 1.9 to 5.2). We note that these results are from a subpopulation of patients with delirium.

- The beta coefficient for the mean difference in severity of delirium was 0.21 (95% CI 0.08 to 1.54).

Summary

- There was moderate quality evidence that a bladder catheter used in ICU patients was a risk factor for the incidence of delirium, but the GDG was uncertain how to interpret this information

- There was low quality evidence that medical restraint was a risk factor for the severity of delirium

- There was low quality evidence that physical restraint was a risk factor for the severity of delirium and moderate evidence that it was a risk factor for persistent delirium

7.5.4 Overall summary

Results for risk factors for delirium incidence are summarised in table 7.9, ordered by size of effect. Values are represented by the midpoint or highest quality study where there was more than on similar value. The corresponding values for persistent delirium are shown on table 7.10. Results for severity and duration of delirium are shown in tables 7.11 and 7.12. For severity of delirium, key results taken into consideration informing the GDG discussions are presented.

Table 7.9: summary of results: risk factors for delirum incidence

Incidence of delirium- GDG had confidence in the results for the following risk factors	
Risk Factor	Summary statistic [OR (95% CI)]
Vision impairment	1.70 (1.01 to 2.85)
Infection	2.96 (1.42 to 6.15)
Age over 65	3.03 (1.19 to 7.71)
Illness severity (APACHE)	3.49 (1.48 to 8.23)
Age over 80	5.22 (2.61 to 10.44)
Cognitive impairment	6.30 (2.89 to 13.74)
Fracture on admission	6.57 (2.23 to 19.33)
Incidence of delirium- GDG had less confidence in the results for the following risk factors	
Vascular surgery	2.70 (1.72 to 4.24)
Comorbidity >3 diseases	15.94 (4.60 to 55.27)
Incidence of delirium- GDG noted uncertainty in the results for the following risk factors	
Sex	1.36 (0.64 to 2.89)
Polypharmacy >3drugs	33.60 (1.90 to 591.6)

Incidence of delirium- GDG had confidence in the results for the following risk factors	
Polypharmacy >7drugs	1.90 (1.11 to 3.24)
Dehydration BUN/creatinine	2.02 (0.72 to 5.64)
Electrolyte disturbance	2.40 (1.09 to 5.27)
Depression	2.43 (0.93 to 6.35)
Bladder catheter	2.70 (1.44 to 5.05)

Table 7.10: summary of results: risk factors for persistent delirum

Persistent delirium	
Risk Factor	Summary statistic [OR (95% CI)]
Charlson comorbidity >3	1.70 (1.11 to 2.61)
Vision impairment	2.10 (1.34 to 3.29)
Cognitive impairment	2.30 (1.41 to 3.74)
Physical restraint	3.20 (1.93 to 5.29)

Table 7.11: summary of results: risk factors for severity of delirum

Severity of delirium	
Risk Factor	Summary statistic [B coefficient (95% CI)]
Setting	Patients in ICU may be at higher risk than patients in medical wards: 4.37 (3.17 to 5.57)
Age (continuous variable)	0.03 (-0.01 to 0.07)
Environmental risk factors	
Number of room changes	0.37 (0.04 to 0.70)
Clock/watch absent	MD 0.41 (0.04 to 0.78)
Calendar absent	MD -0.13 (-0,72 to 0.46)
Not wearing glasses	MD 0.82 (0.45 to 1.19)
Family absent	MD -0.48 (-0.99 to 0.03)

Table 7.12: summary of results: risk factors for duration of of delirum

Duration of delirium	
Risk Factor	Summary statistic [OR (95% CI)]
Age (continuous variable)	1.02 (0.98 to0 1.06)

The dichotomous results for the risk factors for delirium incidence are summarised on a forest plot, ordered by size of effect (figure 7.30). This is intended to give a visual summary and the values are represented by the midpoint or highest quality study where there was more than one similar value. These values have not been incorporated in the economic model. The corresponding values for persistent delirium are shown on figure 7.31.

Figure 7.30: risk factors for incidence of delirium

Study or Subgroup	log[Odds Ratio]	SE	Odds Ratio IV, Fixed, 95% CI
19.1.1 GDG confidence			
Vision impairment	0.530628	0.264309	1.70 [1.01, 2.85]
Infection	1.085189	0.373927	2.96 [1.42, 6.16]
Age over 65	1.108563	0.476525	3.03 [1.19, 7.71]
Illness severity (APACHE)	1.24990174	0.43769	3.49 [1.48, 8.23]
Age over 80	1.6524974	0.353647	5.22 [2.61, 10.44]
Cognitive impairment	1.84054963	0.397948	6.30 [2.89, 13.74]
Fracture on admission	1.88251383	0.550933	6.57 [2.23, 19.34]
19.1.2 GDG weak confidence			
Vascular surgery	0.99325177	0.230729	2.70 [1.72, 4.24]
Comorbidity >3 disease	2.76883167	0.634413	15.94 [4.60, 55.27]
19.1.3 GDG uncertainty			
Sex	0.307485	0.385459	1.36 [0.64, 2.89]
Polypharmacy >7drugs	0.641854	0.272408	1.90 [1.11, 3.24]
Dehydration BUN/creat	0.70309751	0.5239	2.02 [0.72, 5.64]
Electrolyte disturbance	0.875469	0.401122	2.40 [1.09, 5.27]
Depression	0.887891	0.49003	2.43 [0.93, 6.35]
Bladder catheter	0.993252	0.319582	2.70 [1.44, 5.05]
Polypharmacy >3drugs	3.51452607	1.464535	33.60 [1.90, 592.86]

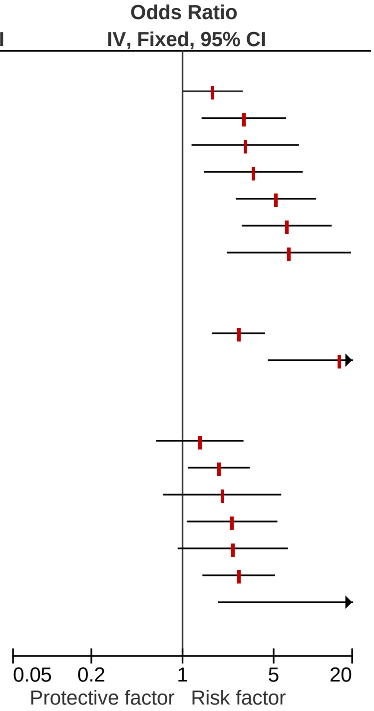
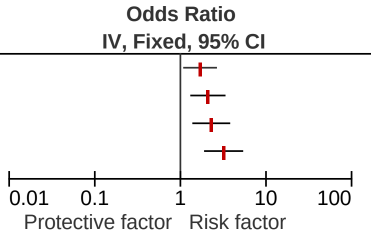


Figure 7.31: risk factors for persistent delirium

Study or Subgroup	log[Odds Ratio]	SE	Odds Ratio IV, Fixed, 95% CI
Comorbidity Charlson >3	0.530628	0.219439	1.70 [1.11, 2.61]
Vision impairment	0.741937	0.229792	2.10 [1.34, 3.29]
Cognitive impairment	0.832909	0.247924	2.30 [1.41, 3.74]
Physical restraint	1.163151	0.256838	3.20 [1.93, 5.29]



7.6 Health economic evidence

No relevant health economic papers were identified.

7.7 Clinical evidence statements

The quality of the studies included in this review has been summarised within each risk factor section and has also been taken into account in figure 7.30.

7.8 From evidence to recommendations

Risk factor assessment (recommendations 1.1.1 and 1.1.2)

The GDG noted that there was moderate or low quality evidence from the risk factors review for each of 20 risk factors for the incidence of delirium, and limited evidence for the duration, severity and persistence of delirium. The GDG separated the evidence into three categories: those risk factors for which the GDG had some confidence in the evidence, those for which it had slight confidence and those for which there was inconsistency or uncertainty. The risk factors in which the GDG had some confidence were:

- Age as a continuous variable
- Age over 65 years
- Age over 80 years
- Cognitive impairment /dementia
- Vision impairment
- Illness severity using the APACHE II as a continuous variable
- Fracture on admission
- Infection
- Physical restraint

These risk factors can be considered of two types, those that can be modified (e.g. infection) and those that are not modifiable (e.g. age). The magnitude of the independent modifiable risk factors ranged from an odds ratio of around 1.7 (vision impairment) to around 3.0 (infection). The magnitude of the independent non-modifiable risk factors ranged from about 3.0 (age over 65 years) to about 6.6 (fracture).

The risk factors in which the GDG had less confidence in were:

- Comorbidity
- Vascular surgery

The GDG noted that the following risk factors had inconsistent or uncertain results:

- Depression
- Hearing impairment
- Polypharmacy
- Dehydration
- Sex
- Electrolyte disturbance
- Immobility
- Incontinence
- Bladder catheter

The GDG wished to define an at-risk group of people, who would be targeted to receive the multicomponent preventative intervention (section 10.25.1).

The GDG recognised that the multicomponent interventions addressed modifiable risk factors only. There was no expectation that the incidence of delirium would be reduced for people who did not have any modifiable risk factors. In defining the at-risk group, the GDG considered which risk factors were important. People who had non-modifiable risk factors for delirium had a higher baseline risk, and the additional presence of a modifiable risk factor would raise the risk of developing delirium. For example, one person with no risk factors might have a baseline risk of 5%, and another person aged 75 years with a hip fracture might have a 50% risk of delirium. If the two people also had an infection (e.g. with a relative risk of 2), the risks of delirium would be 10% and 100% for the two cases.

Taking these factors into consideration, and that the clinical and cost-effectiveness evidence for the multicomponent intervention only applied to people at intermediate and higher risk of delirium (as defined in the studies),, the GDG concluded that the intervention(s) should not be offered to everyone in hospital or long-term care, and that non-modifiable risk factors should be used to define the 'at-risk' group. The modifiable risk factors would then be addressed by the multicomponent intervention.

The GDG noted that the Inouye (1993) study defined the intermediate risk group as 1or 2 risk factors from: visual impairment, severe illness, cognitive impairment, high blood urea nitrogen to creatinine ratio; all these risk factors had relative risks of at least 2. On this basis, the GDG decided to define at-risk patients as those with at least one risk factor with a relative risk of at least 2, but restricted this to non-modifiable risk factors about which they were confident from the evidence.

The GDG, decided to exclude visual impairment, infection and physical restraint in the definition of the at-risk group; infection and visual impairment are covered by the multicomponent intervention. The evidence pertaining to physical restraint as a risk factor for the severity and persistent delirium was low and moderate quality. The GDG noted that restraint is sometimes used in patients with delirium to prevent them causing harm to themselves, for example, self-extubation in ICU. In addition, restraint can indirectly result from medical interventions, for example, by intravenous infusions reducing people's ability to mobilise. The GDG decided against including restraint as a risk factor as part of the multicomponent intervention.

In formulating the recommendations, the GDG considered the following non-modifiable risk factors:

- Age: a cut-off point of 65 years; this decision was based the evidence from one moderate quality study from the risk factor review (Pandharipande 2006), which demonstrated 65 years as a clear cut off point. From the evidence on age as a continuous variable, the GDG were confident that increasing age (above age 65 years) increases the risk of delirium.

- Cognitive impairment/dementia: The GDG acknowledged that, because the studies included in the risk factor review investigated either cognitive impairment or dementia, both factors should be considered within the risk factor assessment.The GDG emphasised that either a known history of cognitive impairment should be ascertained, or that suspected cognitive impairment should be confirmed with a validated measure.

- The GDG agreed that where dementia was suspected, healthcare professionals should cross-refer to the NICE dementia guideline. A foot note in recommendation 1.1.1 was added to this effect.

- Current hip fracture: there was moderate quality evidence for 'fracture on admission' as a risk factor (fracture type unspecified) and low quality evidence for emergency hip fracture surgery in comparison with elective surgery for hip or knee arthritis. The GDG consensus was that the risk factor should be 'current hip fracture'.

- Illness severity: the GDG debated the appropriate illness severity measure and agreed to cross refer to the NICE guideline on acutely ill patients in hospital.

No studies conducted solely in long-term care settings were found. The GDG agreed that the same risk factors would be applicable regardless of the setting.

The GDG discussed when people should be assessed for risk factors, and agreed that this should be conducted when the person presents to hospital or long-term care.

The GDG recognised that during the course of a hospital stay or long-term care, there might be a change in the risk factors for delirium in the group previously defined as not at risk, particularly in terms of illness severity. The GDG therefore added recommendation 1.1.2 covering risk factors developing subsequently to the initial presentation.

There is insufficient evidence available for assessing the role of immune markers such as cytokines as risk factors for delirium. The GDG wished to make a recommendation for future research in this area (see below and Apppendix H):

Future research recommendation:

Is the presence of immune system markers, particularly cytokines, a risk factor for the development of delirium?

Interventions to prevent delirium: care environment (recommendation 1.3.1)

For environmental risk factors there was low quality evidence pertaining to the severity of delirium and no evidence relating to the incidence of delirium. The GDG extrapolated the evidence to incidence of delirium and added to it from their experience, referring to some of the multicomponent prevention studies. There was no economic evidence to underpin this recommendation.

Frequent changes in surroundings, of both room and people, may contribute to feelings of disorientation and delirium, and with frequent changes of staff, information may be lost. The GDG noted the evidence that the number of room changes was a risk factor for delirium severity, but also recognised that trying to reduce the number of room moves may conflict with service provision and operational factors for example assessment wards, single sex wards and that delirium in itself may trigger for a patient being moved to a side ward. This led to the GDG making a recommendation about the care environment (recommendation 1.3.1). The GDG also noted that the evidence underpinning this recommendation came from a North American study (and hence North American healthcare system) which is why it showed that changes in 'rooms' was a risk factor for delirium. The GDG therefore translated this to make it applicable to the UK NHS setting (i.e. changes in wards and rooms) and worded recommendation 1.3.1 to reflect this.

The GDG felt that maintaining a suitable care environment was also important for people diagnosed with delirium as well as those at risk of delirium. This information was added to recommendation 1.6.2 addressing the initial management of delirium (see section 12.7).

Factors related to orientation can help towards minimising the risk due to cognitive impairment. The GDG therefore felt that it was important for people at risk of delirium to be provided with calendars and clocks that are easily visible. It was also noted that some wards (particularly ICU) have no natural light and it is difficult for patients to ascertain whether it is day or night. This highlighted that it was important to consider providing a 24 hour clock for people in critical care. These factors were included in the recommendation addressing disorientation as part of the multicomponent intervention package.

7.9 Recommendations

Risk factor assessment

When people first present to hospital or long-term care, assess them for the following risk factors. If any of these risk factors is present, the person is at risk of delirium:

- Age 65 years or older.
- Cognitive impairment (past or present) and/or dementia[11]. If cognitive impairment is suspected, confirm it using a standardised and validated cognitive impairment measure.
- Current hip fracture.
- Severe illness (a clinical condition that is deteriorating or is at risk of deterioration)[12].

[1.1.1]

Observe people at every opportunity for any changes in the risk factors for delirium. **[1.1.2]**

Interventions to prevent delirium

Ensure that people at risk of delirium are cared for by a team of healthcare professionals who are familiar to the person at risk. Avoid moving people within and between wards or rooms unless absolutely necessary. **[1.3.1]**:

11 If dementia is suspected, refer to further information on the diagnosis, treatment and care of people with dementia in, 'Dementia: supporting people with dementia and their carers in health and social care' (NICE clinical guideline 42).

12 For further information on recognising and responding to acute illness in adults in hospital see 'Acutely ill patients in hospital' (NICE clinical guideline 50).

8 Risk factors for delirium: pharmacological agents

CLINICAL QUESTION: What are the risk factors for delirium?

8.1 Clinical introduction

Delirium often occurs in individuals who are already on medications either for longstanding conditions or acute illness. Some medications seem to be associated with higher incidence of delirium. It appears that many classes of drugs are implicated in the development of delirium. By identifying those drugs responsible, clinicians would not necessarily avoid their use altogether but potentially consider alternatives or be more judicious in their use. Also by identifying pharmacological risk factors, staff or carers looking after the individual would be more vigilant for the signs of the development of delirium. It is not known whether it is the individual's drugs that pose a risk, or the combinations of the different types of drugs.

The knowledge of the propensity of different drugs or groups of drugs to contribute to the development of delirium will help clinicians to reduce the individual's risk at many stages in the patient's journey e.g. admission to a new in-hospital care setting, on admission to long-term care or on routine review by the person's General Practitioner.

8.2 Selection criteria

Selection criteria were as outlined in the general methods section (section 2.3.1) apart from the types of risk factor.

8.2.1 Types of study design

The study designs for pharmacological agents as risk factors were to be RCTs (because they are interventions) or cohort studies. If neither of these designs were available for a particular risk factor, case control studies could also be included.

8.2.2 Types of pharmacological risk factor

Any pharmacological agent used that the GDG considered *a-priori* as reported to be a risk factor for delirium was to be considered.

8.2.3 Types of comparison

The following comparisons were to be included:

- Intervention versus placebo / no intervention
- Intervention 1 + intervention 2 versus intervention 2 alone
- Drug A versus drug B (both drugs in same class)
- Drug class A versus drug class B
- Drug class A (dose 1) versus Drug class A (dose 2)

It was decided to combine the two types of comparison: (i) intervention versus placebo / no intervention and (ii) intervention 1 + intervention 2 versus intervention 2 alone, and examine this assumption using sensitivity analyses.

8.2.4 Type of outcome measure

The types of outcome measure were to be:

- Incidence of delirium [also recording when incidence was measured]
- Severity of delirium
- Duration of delirium

8.2.5 Stratification and subgroup analyses

We planned to stratify the studies by class of drug.

The following subgroups were to be considered:

- Type of pharmacological agent
- Dose

8.3 Description of studies

Details of included and excluded papers together with study design are reported in table 8.1

One study (Pisani 2007) had more than one report (Pisani 2007; Pisani 2009); hereafter, these studies are referred to by the first named report, but separately in the methodological quality assessment and results sections.

Table 8.1: study inclusion, exclusion and design

Papers	Comments	Study
N= 28 evaluated for inclusion		
N= 6 excluded	Reasons for exclusion are reported in Appendix G.	
N= 3 identified in update searches	All 3 studies were excluded from the results section on the basis of low quality	Oh 2008; Shiba 2009; Van Rompaey 2009
N= 22 reports of 21 studies included	Study designs 9 RCTs	Beaussier 2006; Christe 2000; Herrick 1996; Kim 1996; Leung 2006; Nitschke 1996; Papaioannou 2005; Scott 2001; Williams-Russo 1992
	9 reports of 8 prospective cohort studies	Agostini 2001; Dubois 2001; Foy 1995; Han 2001; Morrison 2003; Pandharipande 2006; Pandharipande 2008; Pisani 2007; Pisani 2009
	3 retrospective cohort studies	Centorrino 2003; Holroyd 1994; Shulman 2005
	1 case control studies	Marcantonio 1994

8.3.1 Study Design

Information on study sizes, geographical location and funding are described in table 8.2

Table 8.2: study characteristics

Study	Size (N)	Geographical location	Funding
Agostini 2001	426	USA	Not Stated
Beaussier 2006	59	France	Not Stated
Centorrino 2003	139	USA	Not Stated
Christe 2000	65	Switzerland	Pharmaceutical
Dubois 2001	216	Canada	Not Stated
Foy 1995	418	Australia	Not Stated
Han 2001	278	Canada	Not Stated
Herrick 1996	136	Canada	Non-Pharmaceutical Funding
Holroyd 1994	114	USA	Not Stated
Kim 1996	127	USA	Pharmaceutical/Non – pharmaceutical funding
Leung 2006	228	USA	Non-Pharmaceutical Funding
Marcantonio 1994	245	USA	Non-Pharmaceutical Funding
Morrison 2003	541	USA	Non-Pharmaceutical Funding
Nitschke 1996	92	USA	Non-Pharmaceutical Funding
Pandharipande 2006	275	USA	Non-Pharmaceutical Funding
Pandharipande 2008	100	USA	Non-Pharmaceutical Funding
Papaioannou 2005	50	Greece	Non-Pharmaceutical Funding
Pisani 2007	304	USA	Non-Pharmaceutical Funding
Scott 2001	420	UK	Not Stated
Shulman 2005	10,230	Canada	Non-Pharmaceutical Funding
Williams-Russo 1992	60	USA	Non-Pharmaceutical Funding

The Leung (2006) study also carried out a multivariate analysis on the study population for risk factors other than those randomised, and is treated as a prospective cohort study for the other risk factors. The Han (2001) study reported that patients were those diagnosed with delirium enrolled in what the authors reported as ‘an RCT of a delirium geriatric service or in an observational cohort study of outcomes of delirium’ [references not provided for either study in the text].

8.3.2 Population

The mean age (table 8.3) where reported, ranged from 40.8 (Centorrino 2003) to 83 years (Han 2001). The age ranges varied, and are shown in table 8.1.

Table 8.3: patient ages.

Study	Age (range) years	Study	Age (range) years
Agostini (2001)	80 (73.2 to 86) ±	Morrison (2003)	range not reported
Beaussier (2006)	77.5 (72 to 83) ±	Marcantonio (1994)	73 (65 to 81)
Centorrino (2003)	40.8 (26.7 to 54.9) ±	Nitschke (1996)	66.6 (65 to 69)
Christe (2000)	Median 84 (63 to 98)	Pandharipande (2006)	55.5 (38.5 to 72.5) ±
Dubois (2001)	64.8 (49.3 to 79.7) ±	Pandharipande (2008)	median: 48 (IQR 36 to 60)
Han (2001)	83.4 (76.1 to 90.7) ±	Papaioannou (2005)	median : 68
Foy (1995)	70.2 (59 to 88)	Pisani (2007)	74.6 (67 to 81) ±
Herrick (1996)	72 (65 to 85)	Pisani (2009)	75 (67 to 83) ±
Holroyd (1994)	74.1 (65 to 92)	Scott (2001)	60.8 (49.6 to 68.1) ±
Kim (1996)	66 (24 to 86)	Shulman (2005)	74.7 (67.8 to 81.5)

Leung (2006)	74 (65 to 95)	Williams-Russo (1992)	68 (48 to 84)
Morrison (2003)	9% of the patients had a mean age less than 70 years, 26% were between the ages of 70 to 79 years and 65% were 80 years or older.		

Unless otherwise specified, all data are presented as mean (range); ± indicates that the range was estimated from the mean ± 1 standard deviation; IQR = interquartile range.

The studies varied in the proportions of patients reported to have **cognitive impairment** at baseline. In addition, the GDG decided that, when this was not clearly stated, it was unlikely that patients undergoing elective cardiac surgery would have cognitive impairment at baseline. This gave the following subgroups:

- Three studies reported patients with cognitive impairment/dementia were excluded
    - one study (Christie 2000) reported that patients with moderate to severe cognitive impairment were excluded at baseline;
    - one study (Pandharipande 2006) reported patients with severe dementia and psychosis were excluded;
    - one study (Shulman 2005) reported that patients with a past diagnosis of dementia were excluded *a priori*.
- Fourteen studies reported that some patients had cognitive impairment at baseline (Agostini 2001; Beaussier 2006; Christe 2000; Foy 1995; Han 2001; Herrick 1996; Holroyd 1994; Kim 1996; Leung 2006; Marcantonio 1994; Morrison 2003; Nitschke 1996; Papaioannou 2005; Pisani 2007).

Information on cognitive impairment status was not reported in the remaining studies (Centorrino 2003; Dubois 2001; Pandharipande 2008; Scott 2001). The Scott (2001) study included patients undergoing CABG and the GDG advised that these patients were unlikely to have cognitive impairment at baseline.

Cognitive impairment/dementia was assessed using different scales:

- Nine studies assessed cognitive impairment based on the MMSE score (Agostini 2001; Beaussier 2006; Christe 2000; Foy 1995; Herrick 1996; Kim 1996; Holroyd 1994; Nitschke 1996; Papaioannou 2005);
- Two studies reported excluding patients with a preoperative MMSE score of 23 or below (Foy 1995; Papaioannou 2005);
- Two studies (Herrick 1996; Nitschke 1996) reported the cognitive impairment change scores;
- Two studies (Leung 2006; Marcantonio 1994) used the Telephone Interview For Cognitive Status (TICS);

- One study (Williams-Russo 1992) used the Mattis Dementia Rating Scale;
- One study (Pandharipande 2006) used the Blessed Dementia Rating Scale;
- Two studies used the IQCODE (Han 200; Pisani 2007: short version);
- One study (Morrison 2003) based its assessment on the diagnosis or history of memory impairment or a dementing illness or if one or more errors were made in answering a four item screening test (assessing orientation [place and time]; circumstances of the fracture [place, time, circumstance]; immediate recall of the nature and purpose of the research study; recall of the name or position of the person administering informed consent);
- One study did not state what scale was used to assess cognitive impairment (Shulman 2005).

Six studies reported the mean MMSE score (range 0 to 30) and cognitive impairment status was deduced from the scores. In one study the mean MMSE score indicated that some patients had no cognitive impairment (Beaussier 2006) and in five studies some patients had some cognitive impairment (Agostini 2001; Christie 2000; Kim 1996; Holroyd 1994; Papaioannou 2005).

- The mean Blessed Dementia Rating Scale (range: 0 to 17, with 17 indicating worst; score of 4 or higher representing threshold for dementia) reported in one study (Pandharipande 2006) indicated patients had low prevalence of dementia
- In two studies (Leung 2006; Marcantonio 1994) the mean TICS score was reported (range 0 to 41; cutoff score not reported in either study) indicating that some of the patients may be cognitively impaired.
- One study (Williams-Russo 1992) reported the mean Delirium rating scale (DRS) score (range: 36 item; 5 subscales; score less than 123 points is the cut off for dementia) and range and reported two patients would be classified as mildly demented pre-operatively.
- One study (Pisani 2007) reported the 31% [94/304] of the patients scored above 3.3 in the IQCODE (range: 1 to 5; with 1 indicating much improved compared to 10 years ago and 5 indicating much worse compared to 10 years ago).

Sensory impairment at baseline was reported in four studies (Han 2001; Pandharipande 2006; Pisani 2007; Shulman 2005) and not reported in the remaining studies. Levels of sensory impairment are given in table 8.4. The studies did not generally give much information on how sensory impairment was assessed:

- sensory impairment was patient reported (Pisani 2007)
- assessed clinically at enrolment for presence or absence (Han 2001)
- not reported (Pandharipande 2006; Shulman 2005)

One study (Papaioannou 2005) reported excluding patients with severe auditory or visual disturbances.

Table 8.4: levels of sensory impairment

Study	Visual impairment	Hearing impairment
Han 2001	19.8%	
Pandharipande 2006	58%	16%
Pisani 2007	10.5%	17%
Shulman 2005	1.6%	10.6%

Fourteen reports of 13 studies reported medications taken; some patients were taking several drugs; table 8.5.

Table 8.5: mean number and/or types of mediations

Study	Mean number of medications/ Types of medications
Agostini (2001)	5.4 (SD 3.1) and 5.6 (SD 3.2) medications for the diphenhydramine-exposed and non-exposed groups, respectively. Type of medications not stated
Centorrino (2003)	At least one centrally active drug: benzodiazepine, antipsychotic, antidepressants, anticonvulsant, lithium or a combination (97%)
Christie (2000)	Benzodiazepines (49%), antidepressants (15%), neuroleptics (11%), opioids (11%);
Dubois (2001)	Benzodiazepines, lorazepam, propofol, opioids (fentanyl, meperidine), steroids, antipsychotics (haloperidol or other), corticosteroids
Han 2001	Atypical antipsychotics, anticholinergics, benzodiazepine (not all types of medications listed)
Holroyd (1994)	Treatment with psychotropic medication (various tricyclics (58.8%), antipsychotics (27.2%) serotonin reuptake inhibitors (13.2%), anticholinergic medication (8.8%), methylphenidate (8.8%), buproprion (8.8%), carbamazepine (8.8%), MAOIs (5.1%), thyroid augmentation (3.5%), valproate (3.5%), verapamil (1.8%)
Morrison (2003)	Benzodiazepines or other sedatives and hypnotics, opioids (including meperidine)
Pandharipande (2006)	Opioids (morphine or fentanyl), sedatives (lorazepam, propofol or midazolam),antipsychotics (haloperidol or olanzapine), anticholinergics (atropine, diphenhydramine, bupropion hydrochloride, metoclopramide, prochlorperazine, promethazine)
Pandharipande (2008)	Sedatives, opioids, anticholinergics, antipsychotics, general anaesthesia, histamine blockers, antiarrhythmics, NSAIDs, steroids, antidepressants
Pisani (2007)	History of benzodiazepines or narcotics as an outpatient (25%); and narcotics before ICU admission (20%)
Pisani (2009)	Benzodiazepine or opioids use on admission (25%); during study: benzodiazepine or opioid use (81%), medium to high potency anticholinergic

Study	Mean number of medications/ Types of medications
	medication use (32%), haloperidol use at any point during the ICU stay (32%), steroid use at any point during ICU stay (52%)
Scott (2001)	All patients received 250 ml of 20% mannitol and 8 mmol of magnesium sulphate
Shulman (2005)	13.66 (SD 8.04) ; number of drugs taken in year prior to first treatment for drug of interest
Williams-Russo (1992)	Medications for psychiatric illness (4%)

One study (Kim 1996) examining the role of H2 antagonists on delirium reported patients taking an H2 antagonist preoperatively were excluded. In two studies (Foy 1995; Pisani 2007) evaluating the use of benzodiazepines, use of benzodiazepines within the month prior to admission was confirmed in 26% of the patients in one study (Foy 1995) and use of benzodiazepines or narcotics was confirmed in 25% of the patients in another study (Pisani 2007).

The studies were conducted in the following settings:

- Four studies in medical wards (Agostini 2001; Centorrino 2003; Foy 1995; Han 2001);
- Four studies in the ICU (Dubois 2001; Pandharipande 2006; Pandharipande 2008; Pisani 2007);
- Eleven studies were in a surgical setting (Beaussier 2005; Christie 2000; Herrick 1996; Kim 1996; Leung 2006; Marcantonio 1994; Morrison 2003; Nitschke 1996; Papaioannou 2005; Scott 2001;Williams-Russo 1992);
- One study (Holroyd 1994) evaluated outpatients;
- One study (Shulman 2005) did not clearly describe the setting.

Type of surgery ranged from cardiac surgery (Kim 1996; Scott 2001); colon resection surgery (Beaussier 2006; Nitschke 1996), gastrointestinal endoscopy (Christe 2000) orthopaedic surgery (Herrick 1996; Morrison 2003) general or orthopaedic surgery (Marcantonio 1994) and mixed types of surgery (Leung 2006:spine/orthopaedic, gynaecological and others; Papaioannou 2005; gynaecological, orthopaedic, urological, and vascular).

Eight studies reported some patients were admitted with multiple diagnoses:

- cardiopulmonary diseases (Agostini 2001; Christie 2000);
- hypertension, chronic obstructive pulmonary disease, (Dubois 2001);

- Central nervous system (CNS) and mental disorders, circulatory, respiratory (Foy 1995);
- respiratory, gastrointestinal haemorrhage, sepsis, neurologic, diabetes mellitus, metabolic abnormalities, acute renal failure and cardiac causes (Pisani 2007);
- diabetes mellitus, cardiovascular or respiratory diseases (Papaioannou 2005);
- sepsis/acute respiratory distress syndrome, pneumonia, myocardial infarction/congestive failure, chronic obstructive pulmonary disease (COPD), GI bleeding, drug overdose, hepatic or renal failure, malignancy, other (Pandharipande 2006);
- haemorrhage, airway or facial trauma, chest trauma, colonic or gastric trauma, gastric surgery, neurosurgical trauma, hepatobiliary-pancreatic surgery, orthopaedic surgery, septic shock or acute respiratory distress syndrome (ARDS), other (Pandharipande 2008).

Comorbidities were not reported in the remaining studies.

8.3.3 Pharmacological risk factors

The reviewed studies included the following pharmacological agents (benzodiazepines, antipsychotics, anticholinergics, H2-receptor antagonists, mood stabilising drugs, non-steroidal anti-inflammatory drugs, opioids, anaesthesia/analgesia, benzodiazepines or opioids) as risk factors for delirium. Studies included RCTs or prospective cohort studies (multivariate analyses). Other designs or methods of analysis were included only if there were no other data. Where reported, the indication for the drug is given and noted if the drug was possibly given to treat delirium.

Benzodiazepines

- Midazolam
    - one RCT (Christe 2000) used midazolam as a sedative for endoscopy
    - two cohort studies (Pandharipande 2006; Pandharipande 2008); both used midazolam as a sedative to reduce anxiety in mechanically ventilated patients

- Lorazepam: two cohort studies (Pandharipande 2006; Pandharipande 2008) used lorazepam as a sedative to reduce anxiety in mechanically ventilated patients

- Benzodiazepines (short acting: oxazepam, lorazepam, triazolam, midazolam, and temazepam) given postoperatively (reason not stated): one case control study (Marcantonio 1994)

- Benzodiazepines (long acting: chlordiazepoxide, diazepam, flurazepam) given postoperatively (reason not stated): one case control study (Marcantonio 1994)

- Benzodiazepines (not specified): three prospective cohort studies (Foy 1995, prescribed pre-hospital usually for insomnia; Leung 2006, given postoperatively [reason not stated]; Pisani 2007, given before ICU admission [reason not stated])

The Pandharipande (2008) study reported that patients may have received sedative medications as consequence of delirium. The GDG considered this study likely to be confounded and this study is not considered further.

Similarly, for other studies set in the ICU (Pandharipande 2006; Pisani 2007), the GDG recognised that results may be possibly confounded as patients may have received other sedatives. The overall quality of these studies will be downgraded.

Antipsychotics:

- Clozapine: one retrospective cohort study (Centorrino 2003)

- Haloperidol: one cohort study (Pisani 2009), haloperidol indication unclear, but 70% of patients had agitation on the first day they received haloperidol

Anticholinergics

- Antihistamines with anticholinergic activity:

  - Diphenhydramine given 24h postoperatively: one prospective cohort study (Agostini 2001) and one case control study (Marcantonio 1994)

  - Benztropine: one retrospective cohort study (Shulman 2005)

- All medications with anticholinergic activity:

  - All drugs with anticholinergic activity given 24h postoperatively (antihistamines, tricyclic antidepressants, antiemetics, some neuroleptics): one case control study (Marcantonio 1994)

  - Anticholinergics (including antipsychotics and benzodiazepines), purpose not stated, but 43% haloperidol: one cohort study (Han 2001)

  - The GDG judged this classification of 'all anticholinergics' to be too vague, so this risk factor was not considered further.

H2-receptor antagonists

- Cimetidine (high dose intravenous) versus ranitidine: one RCT (Kim 1996)

  - The GDG noted that the IV form of cimetidine is rarely used in the UK any more, although low dose oral cimetidine can be bought over the counter. However, this study using a high dose intravenous route did not approximate to the over the counter medicine. Therefore this study was not considered further.

- H2 blockers (type and dose not specified): one cohort study (Pandharipande 2008)

Mood stabilising drugs

- Lithium: two retrospective cohort studies (Holroyd 1994; Shulman 2005)

  - Lithium (dose not reported) for mean duration of 7.5 years (SD 2.1) (Holroyd 1994) and mean follow up duration of 8.2 months (new users) (Shulman 2005)

  - Valproate: one study; mean follow up duration of 7.5 months (new users) (Shulman 2005)

Non Steroidal Anti-inflammatory Drugs (NSAIDs)

- Ketorolac tromethamine: one RCT (Nitschke 1996)

Opioids

- Morphine: one RCT (Preoperative intrathecal morphine in addition to postoperative patient controlled analgesia (PCA) morphine (Beaussier 2006)

- Morphine: two cohort studies (Pandharipande 2006; Pandharipande 2008)

- Opioids via PCA: two RCTs (Herrick 1996; Nitschke 1996) and one prospective cohort study (Leung 2006)

- Opioids general: two cohort studies (Dubois 2003: morphine, fentanyl or other; Morrison 2003)

- Meperidine via epidural and via PCA: one case control study (Marcantonio 1994)

- Meperidine : one cohort study (Morrison 2003)

- Fentanyl: one case control study (Marcantonio 1994)

- Fentanyl: one cohort study (Pandharipande 2008)

- Oxycodone: one case control study (Marcantonio 1994)

The Pandharipande (2008) study reported that patients may have received analgesic medications (fentanyl and morphine) as consequence of delirium. The GDG considered this study likely to be confounded and this study is not considered further.

Similarly, for other studies set in the ICU (Pandharipande 2006; Dubois 2003) the GDG recognised that results may be possibly confounded as patients may have received other analgesics. The overall quality of these studies will be downgraded.

Anaesthesia/Analgesia

- Thoracic epidural anaesthesia versus opioid analgesia: one RCT (Scott 2001)
    - Bupivacaine plus clonidine perioperatively versus patient controlled analgesia morphine pump postoperatively; all patients had general anaesthesia

- Continuous epidural bupivacaine plus fentanyl (Williams- Russo 1992)

- Nitrous oxide with oxygen versus oxygen: one RCT (Leung 2006)
- General anaesthesia versus regional anaesthesia: one RCT (Papaioannou 2005)
- Anaesthetics (unspecified): one cohort study (Pandharipande 2008)

More than one drug class

- Benzodiazepine or opioids : one cohort study (Pisani 2009)

8.3.4 Comparisons

For the cohort studies the reference for most of these drugs was the absence of the drug, apart from the following:

- Leung (2006): PCA opioids relative to oral opioids
- Shulman (2005): benztropine and valproate relative to lithium
- Morrison (2003): low dose (below 10 mg) and moderate dose (10 to 30 mg) relative to high dose opioid (above 30 mg/day morphine equivalent)

For the RCTs, the following comparisons were carried out:

Benzodiazepine comparisons

- Benzodiazepines versus placebo/no treatment
    - Midazolam (30 μg/kg IV) versus placebo (saline 0.9% IV) (Christe 2000).

Opioid comparisons

- Opioid versus placebo
    - Intrathecal morphine injected via the 4-5 interspace versus placebo (subcutaneous saline 3 ml injected at the L4-L5 level); both groups also had PCA morphine(300 μg of preservative-free morphine [100 μg /ml] (Beaussier 2006)
- Opioid 1 versus opioid 2
    - PCA fentanyl (10 μg/dose) versus PCA morphine (1mg/dose) (Herrick 1996)

- Opioid route of administration 1 versus route 2
    - PCA morphine versus IM morphine (Nitschke 1996)
    - The doses, intervals and lockout levels for PCA morphine were determined individually based on patients' weight, age and serum creatinine level. Dosing interval: every 4 hours for IM morphine

Analgesia comparisons

- Type of analgesia 1 versus type 2
    - Thoracic epidural anaesthesia perioperatively versus PCA morphine postoperatively (Scott 2001)
        - Thoracic epidural anaesthesia intra- and postoperatively: initial bolus of 5 ml bupivacaine 0.5% followed by another 5 ml bolus after 5 minutes and after surgery a top up bolus up to a maximum of 4 ml of 0.25% when needed. Control group: PCA morphine pump using a 1 mg bolus postoperatively.
        - All patients also received standardised general anaesthesia and analgesia (alfentanil)
    - Postoperative continuous epidural bupivicaine (4 mg/ml) plus fentanyl (10 mcg/ml) versus continuous IV fentanyl (10 mcg/ml) (Williams Russo 1992)
    - IM morphine (opioid) versus IM ketorolac tromethamine (NSAID) (Nitschke 1996)

Anaesthesia

- Anaesthesia versus placebo
    - Nitrous oxide plus oxygen versus oxygen (Leung 2006)
- Type of anaesthesia 1 versus type 2
    - General anaesthesia versus regional anaesthesia: one study (Papaioannou 2005)
    - Further details on drugs and doses not reported

8.3.5 Outcomes

All studies but one reported the incidence of delirium as an outcome; one study reported the duration of delirium (Pisani 2009).

8.4 Methodological quality of included studies

The methodological quality of studies was assessed according to the type of study design. In evaluating the literature, RCTs and prospective cohort studies were selected to be the best available evidence source for this review. One case control study was also included in this review because there was no other information for some risk factors.

8.4.1 RCTs

The quality assessment for the eight included trials is shown in Appendix E.

An adequate method of randomisation was reported in five studies (computer generated: Beaussier 2006; Leung 2006; Papaioannou 2005; table of random numbers: Christe 2000; drawing lots: Scott 2001). The remaining three studies (Herrick 1996; Nitschke 1996; Williams-Russo 1992) did not state the method of randomisation.

An adequate method of allocation concealment was reported in three studies in which an independent member of staff performed the randomisation (Beaussier 2006; Scott 2001) or this was carried out in the hospital pharmacy (Christe 2000). A partially adequate method of allocation concealment was reported in two studies (sealed envelope: Leung 2006; Nitschke 1996) and was not reported or unclear in the remaining studies.

Two studies (Leung 2006; Nitschke 1996) reported that the outcome assessors were blinded to the interventions, one study (Scott 2001) reported blinding was not maintained and blinding was not clearly stated in the remaining studies.

Four studies (Beaussier 2006; Christe 2000; Leung 2006; Scott 2001) described an *a-priori* power calculation. In one study (Leung 2006) the sample size was calculated for the primary outcome, the incidence of delirium. In order to detect a 50% reduction in delirium for the patients not receiving N_20, 114 patients were needed at 80% power, p=0.05.

The remaining studies reported sample size calculations for other outcomes. Further details are in Appendix E.

One study (Christe 2000) reported delirium as an adverse event following sedation with midazolam or placebo (saline) for an upper gastrointestinal endoscopy.

Five studies reported loss to follow up of less than 20% (Beaussier 2006; Christe 2000; Nitschke 1996; Papaioannou 2005; Scott 2001)

Two studies (Leung 2006; Papaioannou 2006) reported an intention to treat analysis, two studies (Beaussier 2006; Scott 2001) carried out an available case analysis and analysis details were not reported or unclear in the remaining studies.

The Papaioannou (2006) study reported conducting both an intention to treat analysis and a per protocol analysis to examine the effect of type of anaesthesia on the MMSE score. It was unclear whether an intention to treat or per protocol analysis was conducted for analysing the incidence of delirium.

All studies included in the review demonstrated baseline comparability of the groups on characteristics such as age, gender, duration of surgery, weight, and type of surgery.

The method of assessment of delirium was:

- **adequate** in three studies (CAM: Beaussier 2006; Leung 2006; DSMIII: Papaioannou 2005);
- **inadequate** in five studies (Christe 2000: a 3 point decline in MMSE scores and medical chart review; Herrick 1996: medical chart review; Nitschke 1996: MMSE; Scott 2001: the GDG agreed that 'confusion' was an inadequate definition of delirium; Williams-Russo 1992: physician and nursing reports)

The overall risk of bias was assessed for the RCTs. Five studies were considered to have potential for bias and were not considered further: four used an inadequate method of assessment of delirium (Christe 2000; Herrick 1996; Nitschke 1996; Williams-Russo 1992) and one (Scott 2001) reported an inadequate *definition* of delirium. The remaining study (Papaioannou 2005) did not describe allocation concealment blinding of outcome assessors was not stated. This study was therefore considered at increased risk of bias.

8.4.2 Cohort studies

There were seven reports of six prospective cohort studies (Agostini 2001; Dubois 2001; Foy 1995; Morrison 2003; Pandharipande 2006; Pisani 2007; Pisani 2009); three were retrospective cohort studies (Centorrino 2003; Holroyd 1994; Shulman 2005) and one was an RCT that was analysed as a cohort study for the benzodiazepine risk factor (Leung 2006). In the Centorriono (2003) study, in patients with more than one admission within the study period, one entry was randomly selected for analysis without knowledge of delirium.

None of the cohort studies were considered to be truly representative of the population (i.e. adults in surgical and/or medical wards in hospital or long-term care).

In all studies, the non-exposed cohorts were drawn from the same community as the exposed cohort.

Levels of missing data were as follows:

- Three studies (Dubois 2001; Pisani 2007; Shulman 2005) reported less than 20% missing data, that is, acceptable levels of missing data;
- The remaining studies did not report on missing data.

One study (Shulman 2005) reported patients with inconsistent data (0.1% [11/10230]) were excluded; the Pisani (2007) study reported imputing missing values (missing: 0.3% for visual impairment to 26% bilirubin).

One study (Foy 1995), reported an *a priori* sample size calculation and calculated that 400 patients would give a power of 98% to detect a relative risk of 4 for the development of cognitive impairment in the benzodiazepine group. Of the 964 patients screened, 568 patients met the eligibility criteria and 418 patients were available for analysis. The study reported separate results for the development of cognitive impairment and delirium.

The studies varied in the number of patients with prevalent delirium (delirium at baseline): further details are given in Appendix D.

- Four studies reported that none of the patients had delirium at baseline (Agostini 2001; Foy 1995; Morrison 2003 (patients with delirium not enrolled); Shulman 2005)
- Two studies reported that some of the patients had delirium at baseline (Dubois 2001: 4% [9/216]; Pandharipande 2006: at least 33% with delirium [66 +/198])
- One study reported these patients were excluded (Dubois 2001);
- Three reports of two studies reported the number of patients who developed delirium following admission (Morrison 2003: 16% [87/541]; Pisani 2007: 70.4% [214/304] within first 48h of ICU admission; Pisani 2009: 79% [239/304] during the ICU stay)

One study (Pandharipande 2006) reported the number of patients who experienced delirium during ICU admission who were administered antipsychotics [88%: 66/75] and anticholinergic drugs [83%: 52/63]. Information on delirium status is missing for 30% (60/198) of the patients.

The method of delirium assessment used was:

- **Adequate** in four studies:
    - Assessed with CAM-ICU and the Richmond Agitation Sedation Scale (Pandharipande 2006)
    - Assessed with CAM-ICU on weekdays and medical chart review at weekends (Pisani 2007)
    - Assessed with CAM on weekdays and medical chart reviewed at weekends (for key words: for example, 'delirious/delirium' 'agitated/agitation' to supplement the CAM observations); delirium was diagnosed if either the CAM or the medical record chart criteria were met (Morrison 2003)
    - MMSE scores and nurse assessed checklists to assess orientation, overall cognitive function, level of alertness and personal care and staff description of nocturnal events to assess criteria according to DSM IIIR criteria (Foy 1995);
- **Partially inadequate** in two studies:
    - Assessed by intensivist and confirmed by a formal psychiatric assessment (Dubois 2001)
    - Multivariate analysis only for 'cognitive decline', which consisted of commonly accepted delirium symptoms in addition to standardised, validated instruments including CAM for delirium and MMSE (Agostini 2001)
- **Inadequate** in two studies:
    - Assessed from medical charts, and from a 3 point severity scale [mild, moderate, severe]. (Centorrino 2003)
    - Information on delirium (classified as a side effect) was extracted by the author in a chart using a structured instrument (no further information on the instrument) (Holroyd 1994).

The method of assessment was not reported in one study (Shulman 2005).

<u>Confounders taken into account</u>

We considered whether the cohort studies took account of particular confounders, either in the study design or the multivariate analysis. The GDG had identified, by consensus, three risk factors to be important: age, sensory impairment, and cognitive impairment.

Studies were summarised according to the number of key risk factors included in the multivariate analysis and the ratio of events to covariates (the GDG considered a ratio of 1 or less to be flawed and a ratio of 2 or 3 to be possibly confounded). We assumed that the key risk factors were the same for severity of delirium and duration of delirium.

Eight reports of seven studies conducted multivariate analyses (Agostini 2001; Dubois 2001; Foy 1995; Morrison 2003; Pandharipande 2006; Pisani 2007; Pisani 2009; Shulman 2005). Two studies conducted only univariate analyses (Centorrino 2003; Holroyd 1994) and these are not considered further. Further details of the factors included in the multivariate analysis are given in Appendix F.

- One study had all/most (3 or 2) of the important risk factors taken into account in the multivariate analysis or they were held constant and had a ratio of events to variables of 10 or more:
    - Shulman (2005): valproate vs lithium: ratio: 12 [72/6]; benztropine vs lithium: 16 [93/5]; key factors were taken into account: age, hearing and visual impairment; patients with dementia were excluded so treated as a constant
- Two studies had all/most (3 or 2) of the important risk factors taken into account in the multivariate analysis or they were held constant but had insufficient ratio of events to variables:
    - Morrison (2003): ratio: 5 [87/16]; key risk factors taken into account: age, cognitive impairment.
    - Pandharipande (2006): ratio ranging from: 4 [66/17] to 7[118/17]; key risk factors taken into account: age, visual and hearing impairment, dementia
    - The study reported the number of patients who experienced delirium for two subgroups: those who received antipsychotics (66/75) and those who received anticholinergics (52/63); it is unclear whether any of the patients were prescribed both drugs. We estimated the incidence of delirium, with incidence ranging from 33% (66/198: the minimum number who had delirium) to 60% (118/198; assuming that patients received either antipsychotics or anticholinergics).
- Six reports of seven studies were possibly confounded: not enough of the important risk factors were taken into account in the multivariate analysis:
    - Agostini 2001) ratio: 31 [122/4] had one key risk factor (age) in the analysis and patients with profound dementia were excluded.

  - Foy (1995) ratio: 2[21/12]; one key risk factor was taken into account: age
  - Leung (2006) ratio:18 [90/5] had one key risk factor taken into account: age
  - Pisani (2007) ratio: 9 [214/23] had one key factor taken into account: dementia (IQCODE score greater than 3.3)
  - Pisani (2009) ratio: 30 [304/10]; key risk factor taken into account: dementia (IQCODE score greater than 3.3)
  - Dubois (2001 ratio: 5 [38/7] had no risk factors taken into account

Overall quality for the cohort studies

- Two cohort studies were considered to be biased and were not considered further:
  - Retrospective study and the method of assessment for delirium was not reported (Shulman 2005);
  - None of the key risk factors were taken into account (Dubois 2001)
- Five reports of four cohort studies were given a low overall quality and treated with caution (evaluated in sensitivity analysis):
  - Only one key risk factor was taken into account (Agostini 2001; Foy 1995; Leung 2006; Pisani 2007; Pisani 2009); and Foy (1995) also had a ratio of 2.
- Two studies (Morrison 2003; Pandharipande 2006) were given a moderate quality rating.
- We note that two studies (Pandharipande 2006; Pisani 2007) were both set in ICU and may possibly be confounded as patients may have received other analgesics and/or opiods (Pandharipande 2006) and therefore were further downgraded in the overall quality.

8.4.3 Case control studies

The nested case control study (Marcantonio 1994) was not considered to be truly representative of the population (i.e. adults in surgical and/or medical wards in hospital or long-term care). The Marcantonio (1994) study was in a surgical setting and the non-exposed cohort was drawn from the same community as the exposed cohort.

The study did not report on missing data or on an *a priori* sample size calculation. The study reported 9% (117/1341) of the patients had delirium at baseline (Marcantonio 1994).

The method of delirium assessment was adequate (CAM).

Confounders taken into account

We considered whether the case control study took account of particular confounders, either in the study design or the multivariate analysis. Cases and controls were matched for: age; poor cognitive function; poor physical function; self reported alcohol abuse; markedly abnormal preoperative serum sodium, potassium or glucose levels; aortic aneurism surgery; and noncardiac thoracic surgery. Thus matching was carried out on two of the key risk factors (age and cognitive impairment). A matched analysis was carried out with drugs being analysed by a logistic regression method so that the effect of each was obtained independently.

Overall, the case control study was both considered to be of low quality because of its design and was considered only if there were no other data.

8.5 Results

We consider below the effects of different risk factors on the incidence, duration and severity of delirium. Results from RCTs and prospective cohort studies are reported mainly and case control studies where there is no other evidence.

8.5.1 Benzodiazepines as a risk factor for the incidence of delirium

Two low quality prospective cohort studies (Leung 2006; Pisani 2007), one moderate quality prospective cohort study (Pandharipande 2006) and one case control study (Marcantonio 1994) reported the effect of benzodiazepines on the incidence of delirium.

Benzodiazepine dose as a continuous variable

Midazolam

One low quality cohort study (Pandharipande 2006) evaluated the use of midazolam (sedative for mechanically ventilated patients to reduce anxiety) as a risk factor for delirium. The analysis considered the transition from normal, delirious or comatose states during the previous 24h to either normal or delirious states in the following 24h. The Pandharipande (2006) study reported that there were small numbers of patients receiving midazolam.

The Pandharipande (2006) study reported the effect of dose (in mg) of midazolam in the previous 24 hours, as a continuous variable (analysed by dose in mg), on the incidence of delirium [OR 1.7 (95% CI 0.9 to 3.2); figure 8.1, Appendix K]. The odds of increased risk of delirium increased by 1.70 per unit of midazolam.

There was no significant effect of midazolam on the incidence of delirium.

Lorazepam

One low quality cohort study (Pandharipande 2006) evaluated the use of lorazepam (as a sedative for mechanically ventilated patients to reduce anxiety) as a risk factor for delirium. The multivariate analysis considered the transition from normal, delirious or comatose during the previous 24h to either normal or delirious status in the following 24h. The number of patients who received lorazepam was not reported.

The Pandharipande (2006) study reported the effect of dose (in mg) of lorazepam in the previous 24 hours, as a continuous variable, on the incidence of delirium (figure 8.2, Appendix K).

The study reported that administration of lorazepam in the previous 24h resulted in a 20% increased risk in transition to delirium in the range 0 to 40 mg [OR 1.2 (95% CI 1.06 to 1.4)]. The study also reported that the incremental risk was large at low doses and the risk of delirium versus dose reached a plateau at 20 mg. It is unclear how this affected the multivariate analysis.

Benzodiazepines as dichotomous variables

Three low quality cohort studies (Foy 1995; Leung 2006; Pisani 2007) and one case control study (Marcantonio 1994) evaluated the use of benzodiazepines as a dichotomous risk factor (use of drug versus no drug) for delirium. The Foy (1995) study evaluated as a risk factor the use of benzodiazepines within 5 days of admission, the Marcantonio (1994) study and the Leung (2006) study evaluated postoperative use on day 1 and days 1 or 2 respectively and Pisani (2007) evaluated use before admission to the ICU.

The Marcantonio (1994) study reported exposure to long-acting agents, including chlordiazepoxide, diazepam and flurazepam, compared with short-acting agents, including oxazepam, lorezepam, triazolam, midazolam and temazepam. Type of benzodiazepines in the Foy (1995) study were diazepam, oxazepam, temazepam, nitrazepam, bromazepam, flunitrazepam, and clorazepate, usually these were prescribed for insomnia. Type of benzodiazepine was not specified in two studies (Leung 2006; Pisani 2007). Indications for benzodiazepine use were not reported. The GDG decided that the studies in which benzodiazepines were given postoperatively were likely to be confounded: it was anticipated that a new prescription of a benzodiazepine would be given for agitation. Therefore, these studies were not considered further.

In the remaining study (Foy 1995), the incidence of delirium was 5% (21/418) and exposure to benzodiazepines was indicated by self-report in 23% (96/418) of the patients.

The odds ratio was 1.0 (95% CI 0.3 to 3.0) indicating use of benzodiazepines 5 days before admission was not a significant risk factor for delirium (figure 8.3, Appendix K).

8.5.2 Antipsychotics

Haloperidol as a risk factor for increased duration of delirium

One low quality cohort study (Pisani 2009) evaluated use of haloperidol as a risk factor for increased duration of delirium in ICU. The study reported that haloperidol was a significant risk factor for the increased duration of delirium (OR 1.35 (95% 1.21 to 1.50) (figure 8.4, Appendix K). The study stated that the haloperidol indication was unclear, but 70% of patients had agitation on the first day they received haloperidol. The GDG considered this study likely to be confounded.

8.5.3 Anticholinergics

Two studies examined specific drugs with anticholinergic activity as a risk factor for the incidence of delirium: one prospective cohort study (Agostini 2001) and one case control study (Marcantonio 1994) evaluated diphenhydramine. The GDG advised that diphenhydramine should be classified as an antihistamine with anticholinergic activity.

One low quality prospective cohort study (Agostini 2001) with 426 patients reported a multivariate analysis (controlling for age, gender and baseline delirium risk) for the risk of cognitive decline in diphenhydramine-exposed group. Cognitive decline was assessment was based on CAM rating for delirium, MMSE scores and presence of delirium symptoms. The number of patients meeting the CAM delirium criteria and decline in MMSE score (≥3 points) was 13% (9/71) in patients receiving the 25mg dose, 17% (7/43) in patients receiving 50mg dose, and 8% (25/312) in patients who did not receive diphenhydramine. 67% of the patients (59/114) were administered the drug for one day and 1 patient received the drug for six consecutive days. Mean number of doses per patient was 2.1 (SD 1.6), and the maximum cumulative daily dose given was 100 mg. Indications for use of diphenhydramine included sleep (68%) and agitation (0.4%).

The Marcantonio 1994 (study) reported results for diphenhydramine administered to 7.3% of the patients (18/245). Of the 22 patients receiving all anticholinergics, 68% (15/22) received a low-dose (defined as one therapeutic dose or less; for example, 25mg for diphenhydarmine). The remaining patients

(7/22) were administered a higher dose, given in either single or multiple doses. Indications for the use of diphenhydramine were not reported.

The odds ratio ranged from 1.80 (95% CI 0.7 to 4.5) to 2.30 (95% CI 1.43 to 3.69) for antihistamines (with anticholinergic activity); figure 8.5(Appendix K). We note that both studies had a potential for bias.

8.5.4 H2 receptor antagonists (H2 blockers)

One cohort study (Pandharipande 2006) evaluated whether exposure to histamine blockers (type not specified) in the previous 24 hours was a risk factor for transitioning to delirium. The number of patients who received H2 blockers was not reported. There was no significant effect of H2 blockers as a risk factor for delirium [OR 1.45 (95% CI 0.80 to 2.62); figure 8.6, Appendix K].

8.5.5 Opiate analgesics

Five studies evaluated opioid analgesics as a risk factor for delirium: three evaluated the effects of individual opioids (cohort studies: Morrison 2003; Pandharipande 2006; case control: Marcantonio 1994); one considered the class of opioids (cohort study: Morrison 2003); one RCT examined the added effect of morphine (Beaussier 2006); one cohort study (Leung 2006) compared patient controlled analgesia (PCA) postoperative opioid analgesia versus oral administration. The case control study (Marcantonio 1994) examined the effect of different types of opioid (meperidine, morphine, fentanyl and oxycodone); because there are higher quality studies reporting the effects of meperidine, morphine and fentanyl, only the results for oxycodone are presented.

Effect of individual opioids

Two prospective cohort studies (Morrison 2003; Pandharipande 2006) and one case control study (Marcantonio 1994) evaluated the effect of exposure to individual opioids on the incidence of delirium. The Pandharipande (2006) study reported the effect of dose of the individual opioid in the previous 24 hours, as a continuous variable, on the incidence of delirium. The Pandharipande (2006) study accounted for the delirium status for only 69% of the patients. The study reported the number of patients who experienced delirium for two subgroups: those who received antipsychotics (66/75) and those who received anticholinergics (52/63); it is unclear whether any of the patients were prescribed both drugs. We estimated the incidence of delirium, with incidence ranging from 33% (66/198: the minimum number who had delirium) to 60% (118/198; assuming that patients received either antipsychotics or anticholinergics).

Opioids as continuous variables (analysed by dose in mcg or mg)

Fentanyl

One moderate quality cohort study (Pandharipande 2006) evaluated the effects of administration of fentanyl (every unit dose in mcg) in the previous 24h on delirium status. Details on doses and number of patients who were administered the drugs were not reported.

The study showed no significant effect of fentanyl as a risk factor for the incidence of delirium The confidence interval is wide [OR 1.2 (95% CI 1.0 to 1.5) (figure 8.7a, Appendix K).

Morphine

One moderate quality cohort study (Pandharipande 2006) evaluated the effect of morphine on the incidence of delirium. Details on doses and number of patients who were administered the drugs were not reported. Exposure of morphine (every unit dose in mg) in the previous 24h on delirium status was reported (OR 1.1 (95% CI 0.9 to 1.2). The confidence interval is wide.

Although this is not a significant effect (OR 1.10), this means that for every increment of a unit dose (in mg) of morphine, the odds of having delirium could increases by a factor of 1.10. Therefore for a 10 mg dose increase, the odds increases by $(1.10)^{10}$, which is 2.59, with the odds ratio ranging from $(0.95)^{10}$ to $(1.27)^{10}$, which is 0.60 to 10.9.

The Pandharipande (2006) study showed no significant effect of morphine on the incidence of delirium (figure 8.7a, Appendix K).

Opioids as dichotomous variable (use of opioids versus no opioids)

Meperidine

One moderate quality study (Morrison 2003) evaluated meperidine use as a risk factor for the development of delirium following admission for hip fracture. 21% of the delirious patients (27/129) received meperidine following admission. Meperidine is a significant risk factor: RR 2.4 (95% CI 1.3 to 4.5); figure 8.7b (Appendix K).

Oxycodone

One case control study (Marcantonio 1994) examined the effect of oxycodone administered during a 24 hour period on the incidence of delirium; 10% of the patients with delirium (9/91) received oxycodone. Details on dose were not reported, nor were indications for the use of oxycodone. There was no significant effect on the incidence of delirium of oxycodone: RR 0.70 (95% CI 0.3 to 1.6); figure 8.7b (Appendix K). We note this study was considered to be of low quality because of the study design.

Effect of all opioids: dose effect

The Morrison (2003) study evaluated the effect on delirium incidence of three different dose ranges (less than 10 mg; 10 mg to 30 mg; above 30 mg) different total daily doses of parenteral morphine sulphate equivalents; doses of all opioids, including continuous infusions and PCA were converted to equivalent dosage. The total daily opioid dose for delirious patients was calculated for the 24 hours preceding the delirious episode and the highest 24h cumulative opioid dose for the first 3 postoperative days for non-delirious patients. The total number of patients who received opioid at the following dose ranges were as follows: below 10 mg: 38% (204/541); 10 to 30 mg: 36% (192/541); above 30 mg 23% (145/541). The study reported the pattern of opioid use in cognitively intact patients (44%: 242/541).

There was a significant effect of parenteral morphine sulphate equivalents on the incidence of delirium observed in patients receiving low doses (below 10 mg compared with the reference above 30mg): RR 5.4 (95% CI 2.4 to 12.3). There was no significant effect of the medium dose (10 to 30 mg) parenteral morphine sulphate equivalents on the incidence of delirium: RR 1.40 (95% CI 0.6 to 3.3); (figure 8.8, Appendix K).

The authors suggested that it is the untreated pain, as opposed to a low dose of opioid, that is the risk factor for developing delirium; the GDG concurred.

Preoperative morphine in addition to postoperative patient controlled analgesia

One RCT (Beaussier 2006) compared the additional effect of preoperative intrathecal morphine on the incidence of delirium in 52 older people recovering from major colorectal surgery. The study compared intrathecal (IT) morphine 0.3 mg (preoperatively) followed by patient controlled analgesia (PCA) morphine (postoperatively), versus preoperative subcutaneous saline plus PCA morphine postoperatively in the control group. The incidence of delirium was 35% (9/26) and 38% (10/36) in the IT morphine plus PCA morphine group and the placebo plus PCA morphine group, respectively. The result is imprecise [OR 0.85 (95% CI 0.27 to 2.62)] (figure 8.9, Appendix K).

Comparison of different routes of administration of opioids postoperatively

One low quality prospective cohort study (Leung 2006) compared the effects of different routes of delivery of postoperative opioids (PCA opioids versus oral opioids) on the incidence of delirium during recovery.

The multivariate analysis (adjusted for age, anaesthesia type, dependence on performing at least one ADL, postoperative analgesia, use of benzodiazepines) showed a higher risk of delirium in patients who received PCA, compared with oral opioids (figure8.10). PCA administration of opioids was a significant risk factor for delirium compared with oral opioids; OR 3.75 (95% CI 1.27, 11.01); the CI is wide, indicating some uncertainty in the magnitude of the effect (figure 8.10, Appendix K). No details were given regarding the oral opioids, and the doses were not reported for either route.

8.5.6 Anaesthesia

Three studies (Leung 2006; Papaioannou 2005; Pandharipande 2008) investigated the effects of anaesthesia on delirium: one RCT at higher risk of bias (Papaioannou 2005) compared general with regional anaesthesia (epidural or spinal), one RCT (Leung 2006) compared nitrous oxide and oxygen versus oxygen alone and one cohort study (Pandharipande 2008) evaluated the effect of anaesthetics on the incidence of delirium.

General anaesthesia versus regional anaesthesia

One RCT (Papaioannou 2005) compared the incidence of delirium in patients receiving general anaesthesia (n=25) versus those receiving regional anaesthesia (epidural or spinal) (n=25) for orthopaedic, urological, vascular or gynaecological surgery. Details on type of anaesthetic agents and dose were not stated. Duration of anaesthesia was over 120 min in over half the cases. Benzodiazepines were not administered for premedication or intraoperative sedation.

The incidence of delirium was 21% (6/28) and 16% (3/19) in the general and regional groups, respectively in the Papaioannou (2005) study. There was no significant effect of type of anaesthesia on delirium, although the results are very imprecise [OR 1.45 (95% CI 0.32 to 6.71)] (figure 8.11, Appendix K).

N_2O plus oxygen versus oxygen

In one RCT (Leung 2006) 228 patients were randomised to receive nitrous oxide plus oxygen or oxygen alone (as part of their intraoperative anaesthetic management) to evaluate if there was a difference in the incidence of delirium during recovery from general anaesthesia. There was no significant difference (figure 8.12, Appendix K), although the results are imprecise [OR 1.07 (95% CI 0.57 to 2.07)].

Anaesthesia

One study (Pandharipande 2008) reporting the effect of exposure to anaesthetics (type not reported) on the incidence of delirium showed no significant effect; OR 0.52 (95% CI 0.23 to 1.16); (figure 8.13, Appendix K).

8.5.7 Effect of benzodiazepines or opioids on the duration of delirium

One study (Pisani 2009) evaluated the use of benzodiazepines or opioids as a risk factor for the duration of delirium; 81% (247/304) of the patients were administered benzodiazepines or opioids. There was a significant effect of use of these drugs on the duration of delirium in ICU, but results were not reported separately for the two classes of drugs; RR 1.64 (95% CI 1.27 to 2.10)(figure 8.14, Appendix K). The GDG considered the results from this study set in the ICU had limited applicability when compared to other hospital populations.

8.6 Overall summary of results

Results for the pharmacological risk factors for delirium incidence and duration are summarised in tables 8.6 and 8.7.

Table 8.6: summary of the results: pharmacological risk factors on the incidence and duration of delirium

Incidence of delirium	
Risk factor	**Odds ratio (95% CI) Unless otherwise stated**
Benzodiazepines (as a continuous variable)	
Midazolam	1.7 (0.90 to 3.2)
Lorazepam	1.2 (1.1to 1.4)

Benzodiazepines (as a dichotomous variable)	
Benzodiazepines	1.0 (0.3 to 3.0)
Antihistamines with an anticholinergic effects	
Diphenhydramine (prospective cohort)	2.3 (1.43 to 3.69)
Diphenhydramine (case control)	1.8 (0.71 to 4.5)
H2 receptor antagonists	
H2 blockers	1.45 (0.80 to 2.62)
Individual opiates (as a continuous variable)	
Fentanyl	1.2 (1.0 to 1.5)
Morphine	1.1 (0.9 to 1.2)
Individual opiates (as a dichotomous variable)	
Meperidine	2.4 (1.3 to 4.5)
Oxycodone	RR 0.7 (0.3 to 1.6)
All opioids	
Opioids (reported at parenteral morphine sulphate equivalents)	<10mg vs reference 30mg:RR 5.4 (2.4 to 12.3)
	10 to 30mg vs reference 30mg: RR 1.4 (0.6 to 3.3)
Morphine	1.10 (0.95 to 1.27)
Oxycodone	RR 0.70 (0.30 to 1.62)
Preoperative morphine + postoperative PCA	0.85 (0.27 to 2.62)
Routes of administration	3.75 (1.27 to 11.01)
Anaesthesia	
General vs regional anaesthesia	1.45 (0.34 to 7.79)
N20 plus O2 vs O2	1.07 (0.57 to 2.07)
Anaesthesia (type not reported)	0.52 (0.23 to 1.16)

Table 8.7: summary of results for risk factors on the duration of delirium

Duration of delirium	
Risk factor	**Odds ratio (95% CI)**
Antipsychotics	

Haloperidol	1.35 (1.21 to 1.50)
Benzodiazepines/Opioids	
Benzodiazepines/Opioids	1.64 (1.27 to 2.10)

8.7 Health economic evidence

No relevant health economic papers were identified.

8.8 Clinical evidence statements

- There is moderate quality evidence to show no significant effect of midazolam on the incidence of delirium.
- There is moderate quality evidence to show there is a significant effect of lorazepam as a risk factor for the incidence of delirium (lorazepam was given to sedate in ITU).
- There is low quality evidence indicating that the use of benzodiazepines 5 days before admission was not a significant risk factor for the incidence of delirium.
- There is low quality evidence from one prospective cohort study to show that diphenhydramine (an antihistamine with anticholinergic activity) is a significant risk factor for the incidence delirium; there is some uncertainty with this result.
- There is very low quality evidence from one case control study to show diphenhydramine (an antihistamine with anticholinergic activity) is not a significant risk factor for the incidence delirium.
- There is moderate quality evidence to show no significant effect of H2 blockers on the incidence of delirium.
- There is inconsistent evidence for the effect of individual opioids on delirium.
    - There is moderate quality evidence to show no significant effect of fentanyl on the incidence of delirium.
    - There is moderate quality evidence to show meperidine is an important risk factor for the incidence of delirium.
    - There is moderate quality evidence to show no significant effect of morphine on the incidence of delirium.
    - There is very low quality evidence to show no significant effect of oxycodone on the incidence of delirium.

- There is moderate quality evidence to show a significant effect of parenteral morphine sulphate equivalents on the incidence of delirium observed in patients receiving low doses below 10 mg (compared with the reference dose above 30mg). There was no significant effect of the medium dose (10 to 30 mg) on the incidence of delirium.

- There is moderate quality evidence from one RCT to show preoperative morphine in addition to patient controlled analgesia in the postoperative period is not a significant risk factor for delirium. There is some uncertainty with this result.

- There is low quality evidence showing patient controlled administration of opioids was a significant risk factor for delirium compared with oral opioids. There is some uncertainty with this result.

- There is moderate quality evidence from one RCT to show there was no significant effect of type of anaesthesia (general compared with regional anaesthesia) on delirium. There is much uncertainty with this result.

- There is low quality evidence from one cohort study to show anaesthesia (type not reported) is not an important risk factor for the incidence of delirium.

- There is moderate quality evidence from one RCT to show no significant difference in the incidence of delirium in patients receiving nitrous oxide plus oxygen or oxygen alone.

- There is low quality evidence to show use of benzodiazepines or opioids is a significant risk factor for the duration of delirium in ICU.

8.9 From evidence to recommendations

The GDG noted all of the evidence reviewed was low to moderate quality. The evidence proved difficult to interpret because the studies often did not report what the drugs were being used for. Some of the drugs are normally used to treat delirium and could have been given for this purpose in the studies, so it was difficult to assess their contribution as a risk factor. In addition, for the ICU patient group, the methods of administration, dose and indication of drug use is often very different to other hospital populations.

The evidence for lorazepam was not used as the basis of a recommendation because the lorazepam study pertained to a specific population. The drug was used to sedate people who were receiving ventilation in ICU. However the GDG took into consideration the evidence for lorazepam when discussing the use of sleep enhancers within the multicomponent intervention section (recommendations 1.3.2 and 1.3.3) for sleep disturbance sub-recommendation 1.3.3.10.

The GDG noted the inconsistency in the evidence pertaining to opioids as a risk factor for delirium and deliberated whether untreated pain in itself was a contributing factor to delirious episodes. The GDG considered this as indirect evidence when developing the subrecommendation on pain 1.3.3.6 within the tailored multicomponent intervention package (section 10.25).

Although there was moderate quality evidence indicating that meperidine (also referred to as pethidine in the UK) is a potential risk factor for delirium the GDG did not wish to make a recommendation. Meperidine is not widely used in the UK; the evidence for meperidine came from one study, and overall the evidence for opioids was inconsistent.

The GDG did not wish to make a recommendation pertaining to patient controlled analgesia or benzodiazepine/opioids because these studies were low quality.

8.10 Recommendations

There are no recommendations for this section. In light of the evidence the GDG did not wish to make recommendations.

9 Consequences of delirium

CLINICAL QUESTION: What are the consequences of delirium in terms of morbidity and mortality in a person in hospital or long-term care?

9.1 Clinical introduction

Delirium has the potential to have an effect on a wide range of outcomes for the delirious person themselves, their family or carers, and health and social care organisations. Some of these may be a direct result of damage caused by the inflammatory response to delirium, whereas others may be a consequence of delirium affecting motor control and behaviour. In addition, many outcomes may also be affected by the index condition that is causing the delirium. Establishing the effect delirium has on outcomes can be challenging, with many potential confounding variables to be considered. This review examines the evidence for an independent effect of delirium on outcomes (i.e. delirium itself is a risk factor) affecting individuals (such as mortality, the development of dementia, falls) and organisations (length of hospital stay, institutionalisation) which will help to demonstrate the impact of delirium and identify areas for improvement.

9.2 Description of studies

Details of included and excluded papers together with study design are reported in table 9.1

Table 9.1: study inclusion, exclusion and design

Papers	Comments	Study
N= 36 evaluated for inclusion		
N= 12 excluded	Reasons for exclusion are reported in Appendix G.	
N= 1 identified in update searches	The study has not been reported in results as it was of low quality and would have been excluded in the sensitivity analysis	Bickel 2008
N= 24 reports of 19 studies were included	Study designs 24 reports of 19 studies were prospective cohort studies	Andrew 2005; Balas 2009; Bourdel-Marchasson 2004; Dolan 2000; Drame 2008; Ely 2004; Francis 1990; Francis 1992; Holmes 2000; Nightingale 2001; Inouye 1998; Leslie 2005; Levkoff 1992; Lin 2004; Lin 2008; McAvay 2006; Marcantonio 2000; Givens 2008;

		Marcantonio 2002; O'Keeffe 1997; Pitkala 2005; Rockwood 1999; Rudolph 2008; Thomason 2005
Three studies reported patients were part of either the intervention and/or control group in a trial: Leslie 2005: intervention and control groups enrolled in a delirium prevention intervention (Inouye 1999); McAvay 2006: control group of Delirium Prevention Trial (Inouye 1999); Marcantonio 2000: intervention and control arms of a trial described as a randomised trial on prevention of delirium [proactive geriatric consultation].		

Three studies had more than one report, which differed in the outcomes reported (Francis 1990 and Francis 1992; Holmes 2000 and Nightingale 2001; Marcantonio 2000, Marcantonio 2002 and Givens 2008). Hereafter, these studies are referred to by the first named reports, but are reported separately where appropriate and reported separately in the results section. One report (Lin 2008) included some of the same patients included in the Lin (2004) study but reported different outcomes and are reported separately. Two studies (Leslie 2005; McAvay 2006) included some of the same patients but reported different outcomes and are reported individually.

This review examines the evidence for the consequences associated with presence of prevalent or incidence delirium, increased delirium duration and increased delirium severity. Details of outcomes identified a-priori are reported in table 9.2a

Table 9.2a: outcomes of interest

Outcomes	Details	Study
Dementia/cognitive impairment/cognitive dysfunction	Cognitive impairment at discharge	Ely 2004
	Cognitive dysfunction at 7 days	Rudolph 2008
	Cognitive dysfunction at 3 months	Rudolph 2008
	Dementia at 3 years	Rockwood 1999
New admission to institution	At discharge	Balas 2009; Bourdel-Marchasson 2004; Inouye 1998; Levkoff 1992
	3 months	Inouye 1998
	6 months	O'Keeffe 1997
	2 years	Pitkala 2005
Mortality	In hospital	Inouye 1998; O'Keeffe 1997
	In ICU	Lin 2004
	In ICU and hospital	Lin 2008; Thomason 2005
	1 month	Marcantonio 2000
	6 weeks	Drame 2008
	3 months	Inouye 1998
	6 months	Ely 2004 [incidence and duration of delirium]; Francis 1990; Holmes 2000; Levkoff 1992; Marcantonio 2000; O'Keeffe 1997
	1 year	Leslie 2005 [incidence and severity of delirium]; Pitkala 2005
	2 years	Dolan 2000; Francis 1992; Nightingale 2001; Pitkala 2005
	3 years	Rockwood 1999

Outcomes	Details	Study
Length of stay	Hospital	Ely 2004 [incidence and duration of delirium]); Francis 1990; Holmes 2000*; Levkoff 1992; Thomason 2005; O'Keeffe 1997
	ICU	Thomason 2005
	Post ICU [defined as length of stay after first ICU discharge]	Ely 2004 [incidence and duration of delirium].

*The Holmes (2000) study reported the risk of being discharged sooner, which corresponds to decreased risk of remaining in hospital. This outcome will be grouped with studies reporting length of stay and the key confounding factors identified for length of stay would be applicable for this outcome.

The GDG agreed, *post-hoc*, outcomes identified during the course of the review, should also be included (table 9.2b).

Table 9.2b: Outcomes identified

Outcomes	Details	Study
Hospital acquired complications		O'Keeffe 1997
Mortality or new admission to institution	At discharge	Inouye 1998
	At 1 month	Givens 2008; Marcantonio 2000; Marcantonio 2002 [severity of delirium]
	At 3 months	Inouye 1998
	At 6 months	Givens 2008; Marcantonio 2000; Marcantonio 2002 [severity of delirium]
	At 1 year	McAvay 2006; Pitkala 2005
Mortality or functional decline	at discharge and at 6 months	Andrew 2005 [duration of delirium]

The GDG agreed that for incidence of delirium, cognitive dysfunction can be grouped with studies reporting dementia and cognitive impairment and that the key confounding factors identified for dementia would be applicable for this outcome.

One additional study (Francis 1992) reported the outcome 'loss of independent living' defined as 'patients institutionalised or needing assistance on 1 of 4 ADL'. The GDG thought that for this outcome, patients needing assistance on 1 of 4

ADL may be confounded by stroke (10% of patients with cerebrovascular diseases) and advised that this outcome should not be included in the review.

The Rudolph (2008) study also reported a subgroup analysis for two different durations of delirium, not allowing for duration of delirium in the multivariate analysis. This outcome will not be considered in this review.

The general characteristics of the studies including methodological quality are discussed for all studies first. These are reported separately for each outcome, where appropriate, and the results are reported separately for each consequence.

9.3 Characteristics of included studies

9.3.1 Study Design

Information on study sizes, geographical location and funding are described in table 9.3. For details of study quality, see appendix E.

Table 9.3: study characteristics

Study name	Size (N)	Geographical location	Funding
Andrew 2005	77	Canada	No funding
Balas 2009	117	USA	Not stated
Bourdel-Marchasson 2004	427	France	Non industry
Dolan 2000	682	USA	Non industry
Drame 2008	1036	France	Non industry
Ely 2004	275	USA	Non industry
Francis 1990	229	USA	Non industry
Holmes 2000	731	UK	Non industry
Inouye 1998	727	USA	Non industry
Leslie 2005	919	USA	Non industry
Levkoff 1992	325	USA	Non industry
Lin 2004	131	Taiwan	Not stated
Lin 2008	143	Taiwan	Not stated
McAvay 2006	433	USA	Non industry
Marcantonio 2000	126	USA	Non industry
Marcantonio 2002	122	USA	Not stated
O'Keeffe 1997	225	UK	Not stated
Pitkala 2005	425	Finland	Non industry
Rockwood 1999	203	Canada	Non industry
Rudolph 2008	1218	UK, Denmark, France, Germany, Greece, the Netherlands, Spain and USA.	Non industry
Thomason 2005	261	USA	Non industry

One study was conducted in both hospital and long-term care; the latter was the setting for 53% of the patients (Pitkala 2005). All the remaining studies were conducted in hospitals. Patients were in different types of wards:

- medical (Bourdel-Marchasson 2004; Dolan 2000; Drame 2008; Francis 1992; Leslie 2005; McAvay 2006; O'Keeffe 1997; Rockwood 1999). Where reported, the principal diagnoses of patients admitted to medical wards were:

- hip fracture (Dolan 2000);
- cancer, coronary artery disease, congestive heart failure, chronic lung disease, cerebrovascular disease, diabetes, hypertension (Francis 1992);
- pneumonia, chronic lung disease, congestive heart failure, ischemic heart disease, gastrointestinal disease, diabetes mellitus or metabolic disorder, cancer, cerebrovascular disease, renal failure, anaemia, and other conditions (Leslie 2005).

- surgical (Marcantonio 2000; Rudolph 2008). For these patients, the surgery was:
  - hip fracture repair (Marcantonio 2000) ;
  - non-cardiac surgery (Rudolph 2008).
- ICU (Balas 2009; Ely 2004; Lin 2004; Thomason 2005). Patients were in ICU for the following reasons:
  - mechanically ventilated patients (Ely 2004; Lin 2004);
    - Principal admission diagnoses of sepsis and/or acute respiratory distress syndrome (46%), pneumonia, myocardial infarction/congestive heart failure, hepatic or renal failure, chronic obstructive pulmonary disease, gastrointestinal bleeding, malignancy, drug overdose, and other diagnoses not stated (Ely 2004);
    - Principal admission diagnoses of pneumonia (34%), chronic lung disease, cerebrovascular disease, cancer, congestive heart failure, ischemic heart disease, gastrointestinal disease, diabetes mellitus or metabolic disorder, drug intoxication and other diagnoses not stated (Lin 2004);
  - non-ventilated [non invasive] patients. (Thomason 2005);
    - Diagnostic admission for pulmonary (27%), gastrointestinal, metabolic, cardiac, haematology/oncology, neurological, renaland other reasons not stated.
  - surgical ICU (Balas 2009)
    - 42.1% received mechanical ventilation at sometime during Surgical Intensive Care Unit (SICU) admission

        - Type of surgery included general (colorectal, surgical oncology and gastrointestinal surgery), vascular and trauma/emergency surgery.

- mixture of medical and surgical wards (Inouye 1998; Levkoff 1992).
    - reasons for admission included:
        - cancer, coronary artery disease, cardiac arrhythmias, congestive heart failure, chronic lung disease, pneumonia, gastrointestinal, cerebrovascular disease diabetes, renal disease and other conditions not reported (40%); number of surgical patients and type of surgery was not reported (Inouye 1998);
        - circulatory system disease (29.2%), digestive system disease, respiratory system disease, fracture, cancer, genitourinary system disease, endocrine, nutritional and metabolic diseases, diseases of skin or other reasons not stated. Type of surgery was not reported (Levkoff 1992).
- mixture of medical (32%), surgical (19%) and geriatric wards (48%) (Andrew 2005).

Eight studies reported the settings from which patients were admitted:

- community (Dolan 2000; Francis 1990);
- emergency units (Drame 2008);
- community (65%) and the remaining patients from long-term care (Levkoff 1992);
- community (41%), nursing homes(4%) and the remaining admissions were unclear (Inouye 1998);
- 6.1% from nursing home (Leslie 2005);
- community (93%) and the remainder from nursing homes (Marcantonio 2000);
- community (81%) and remaining patients from long-term care or residential home care (O'Keeffe 1997).

9.3.2 Population

The mean age, where reported, ranged from 55 years (Ely 2004) to 82.1 years (Holmes 2000). The age range was reported in four studies (Andrew 2005; Drame 2008; Holmes 2000; McAvay 2006) and the range was estimated from the mean ± 1 standard deviation in the remaining studies (table 9.4).

Table 9.4: patient ages

Study	Mean age and range (years)	Study	Mean age and range (years)
Andrew 2005	78.5 (64 to 93)	Leslie 2005	80 (73.5 to 86.5)$^{\pm}$
Balas 2009	75.4 (69.1 81.7) $^{\pm}$	Levkoff 1992	81.4 (73.7 to 89.1)$^{\pm}$
Bourdel-Marchasson 2004	85 (78.4 to 92.4)$^{\pm}$	Lin 2004	73.6 (70.5 to 77.4)$^{\pm}$
Dolan 2000	82 (72.6 to 90.1)$^{\pm}$	Lin 2008	76 (64 to 85.5)
Drame 2008	85 (75 to 103)	McAvay 2006	80 (70 to 99)
Ely 2004	55 (37 to 73)$^{\pm}$	Marcantonio 2000	79 (71 to 87)$^{\pm}$
Francis 1992	78 (72.1 to 85.0)$^{\pm}$	O'Keeffe 1997	82 (76 to 88)$^{\pm}$
Holmes 2000	82.1 (65 to 99)	Rudolph 2008	69 (62.9 to 76.3)$^{\pm}$
Inouye 1998	78.9 (72 to 85.8)$^{\pm}$	Thomason 2005	52.5 (32 to 74)$^{\pm}$

($\pm$) indicates that range was calculated from the mean $\pm$ 1 standard deviation

The age range was not stated and could not be calculated in two studies (Pitkala 2005; Rockwood 1999). The Pitkala (2005) study, however, reported that patients younger than 70 years were excluded and that 59% were over 85 years. In the Rockwood (1999) study patients over 65 years were enrolled and the mean age of 79 years was reported. In the Francis (1990) study patients over 70 years were enrolled and had a mean age of 78 years.

Where reported, all studies included both males and females. Two studies (Holmes 2000; Pitkala 2005) had less than 20% male patients, twelve studies had less than 50% (Andrew 2005; Dolan 1997; Drame 2008; Francis 1990; Inouye 1998; Leslie 2005; Levkoff 1992; McAvay 2006; Marcantonio 2000; O'Keeffe 1997; Rockwood 1999; Thomason 2005) and five studies had 50% or more male patients (Balas 2009; Ely 2004; Lin 2004; Lin 2008; Rudolph 2008). The Bourdel-Marchasson (2004) study did not report the number of male and female patients enrolled.

Fifteen studies reported including patients with cognitive impairment (Andrew 2005; Balas 2009; Bourdel-Marchasson 2004; Drame 2008; Francis 1990; Holmes 2000; Inouye 1998; Leslie 2005; Levkoff 1992; Lin 2008; McAvay 2006; Marcantonio 2000; O'Keeffe 1997; Pitkala 2005; Rockwood 1999), one study (Dolan 2000) reported patients with cognitive impairment were excluded, three studies (Lin 2004; Lin 2008; Rudolph 2008) reported that patients with dementia were excluded, and cognitive impairment was not reported in one study (Thomason 2005). Cognitive impairment ranged from 24% (Levkoff 1992) to 75% (Bourdel-Marchasson 2004). Assessment of cognitive impairment was based on the following scales:

- MMSE (range 0 to 30) (Holmes 2000; Inouye 1998; McAvay 2006; Pitkala 2005; Rudolph 2008);
    - one study (Inouye 1998) used a cut off score of 20 or below to define dementia; a cut off score of below 24 were used in two studies (Ely 2004; McAvay 2006); patients with score of 24 or

below were excluded in one study (Rudolph 2008) and the cut-off point was not reported in one study (Holmes 2000);

- The Inouye (1998) multicentre study used a 21 point scale MMSE at one of the three sites, and scores on the 21 point scale were adjusted to a denominator of 30 points;

- the Pitkala (2005) study used a score below 20 to define moderate cognitive impairment;

- Blessed's Dementia Rating Scale (Francis 1990; Leslie 2005; Lin 2008; Marcantonio 2000; O'Keeffe 1997);
  - The cut-off point was 4 or more in three studies (Francis 1990; Marcantonio 2000; O'Keeffe 1997); 2 or more in one study (Leslie 2005; modified version of Blessed scale); 3 or higher (Lin 2008)
- DSM- III-R criteria (Andrew 2005);
- cognitive status (MMSE, Blessed dementia rating scale) and functional assessment (Barthel Index, Physical Self-Maintenance Scale) to screen for cognitive impairment and assessment of dementia by geriatrician (Rockwood 1999);
- based on family interviews and physicians and checked if existed with respect to DSM-IV criteria (Bourdel-Marchasson 2004);
- IQCODE (Balas 2009);
- medical chart review or assessment of a senior practitioner (Drame 2008);
- medical chart review (Levkoff 1992).

Further details are reported in Appendix D.

Ten studies reported comorbidity scores, using the Charlson Comorbidity Index: (Bourdel-Marchasson 2004; Dolan 2008; Drame 2008; Ely 2004; Leslie 2005; McAvay 2006; Marcantonio 2000; O'Keeffe 1997; Pitkala 2005; Thomason 2005). Further details are reported in Appendix D.

Eight studies reported severity of illness assessed with an established scale (APACHE II: Balas 2009; Ely 2004; McAvay 2006; Leslie 2005; Inouye 1998; Thomason 2005; APACHE III: Lin 2004; Lin 2008). Two studies used a clinician based rating (Francis 1992; Levkoff 1992), severity of illness based on a rating scale (range 1 to 9, with 1= not ill and 9=moribund) (Francis 1992) and a sum of severity scores, calculated based on severity scores assigned to 15 medical conditions: one study (Levkoff 1992).

One study (Holmes 2000) reported using a researcher-rated scale, the modified Burvill scale to record concurrent physical illness (range:0 to 6, with 0 representing no physical illness and 6 representing severe chronic physical illness).

Further details are reported in Appendix D.

9.3.3 Incidence of delirium and its method of assessment

Overall rates of delirium ranged from 8% (Bourdel-Marchasson 2004; Rudolph 2008) to 48% (Thomason 2005).

All of the patients in one study (Andrew 2005: n=77) had delirium; this study was looking at the effects of increased duration of delirium.

The studies varied in whether they investigated the effects of prevalent delirium (occurring on admission to hospital) or incident delirium (appearing during the course of the hospital stay) or both.

- Nine studies included only prevalent delirium as a risk factor (Andrew 2005; Dolan 2005; Drame 2008; Holmes 2000; Inouye 1998; Lin 2004 (ICU study using delirium developed in first 5 days); Lin 2008 (ICU study using delirium developed in first 5 days); Pitkala 2005 (only recorded prevalent delirium); Rockwood 1999 (only recorded prevalent delirium))

- Four studies (Balas 2009; Leslie 2005 (patients with prevalent delirium were excluded); McAvay 2006 (patients with prevalent delirium were excluded); Marcantonio 2000 (reported to be incident delirium)) included only incident delirium rates

- One study (Bourdel-Marchasson 2004) included both prevalent and incident delirium and analysed them separately

- Four studies (Ely 2004; Francis 1990; Rudolph 2008; Thomason 2008) reported both incident and prevalent delirium, but combined them as ‘delirium’ in the analysis

- Two studies (Levkoff 1992; O’Keeffe 1997) reported both prevalent and incident delirium and combined these in some analyses (Levkoff 1992: mortality, length of stay; O’Keeffe 1997: mortality; length of stay; hospital acquired complications) but both reported only incident delirium for discharge to an institution.

Rates of delirium ranged from 8% (Rockwood 1999:16/203) to 82% (Ely 2004: 183/224).

The Bourdel-Marchasson (2004) study reported four categories of delirium: for patients classified as having prevalent delirium [8%:34/427] if the diagnosis of delirium was within the first 4 days of stay, subsequent delirium was classified as incident [3.5%:15/427], prevalent subsyndromal delirium [20.6%:88/427] and incident subsyndromal delirium [14%:60/427]. Patients having one or more CAM symptoms but not fulfilling the CAM algorithm were termed 'subsyndromal delirium'. Results for patients with only prevalent and incident delirium will be reported in this review.

In addition to examining the consequences of either prevalent and/or incident delirium, the GDG wanted to investigate the effect of persistent delirium on adverse outcomes. Persistent delirium was classified in accordance with the definition provided in the McAvay (2006) study. These authors defined persistent delirium as 'patients who met full criteria for delirium at the discharge interview, or had full delirium during the hospitalisation and partial symptoms at discharge'.

Four studies reported information on persistent delirium (Levkoff 1992; McAvay 2006; Marcantonio 2000; O'Keeffe 1997).

Persistent delirium rates were reported for the following time periods:

- discharge: ranged from 17% (Levkoff 1992: 54/325) to 32% (O'Keeffe 1997 [24%: 8/33 of those with prevalent delirium; 37%: 17/46 of those with incident delirium]);
- 1 month: 29% (Marcantonio 2000: 15/52);
- 3 months: 16.2% (Levkoff 1992);
- 6 months: ranged from 6% (Marcantonio 2000: 3/52) to 13.3% (Levkoff 1992);
- 1 year: 43% (McAvay 2006: 24/55).

In the Levkoff (1992) study only the percentages of patients with resolved delirium were reported from which the persistent delirium rates were calculated.

The method of assessment of persistent delirium differed from baseline assessment in one study (Levkoff 1992). At 3 and 6 months follow-up, relatives or carers were interviewed to determine if symptoms persisted. This was deemed an inadequate method of assessment.

In one study (Rockwood 1999), the study population was also separated into patients with delirium and dementia at baseline (11%: 22/203), prevalent dementia only (8%:17/203) and patients with neither delirium nor dementia (73%:148/203). For the outcome, dementia as a consequence of delirium, results were only presented for the combined groups, patients with delirium and patients with neither delirium nor dementia.

In one study (Ely 2004), 67% (123/183) of patients who had delirium for a median of 2 days (IQR 1 to 3) were in a coma for a median of 2 days (IQR 1 to 4).

The method of assessment of delirium varied between the studies. The GDG considered that 18 studies had an adequate method of assessment; one had a partially inadequate method and one was inadequate:

Adequate

- Ten studies used either the Confusion Assessment Method (CAM) (Bourdel-Marchasson 2004; Inouye 1998; Leslie 2005; McAvay 2006; Marcantonio 2000) or a variation (CAM-ICU: Balas 2009 ; Ely 2004; Thomason 2005; Chinese version of CAM ICU: Lin 2004; Lin 2008).
    - One study (Balas 2009) reported a patient was considered delirious if they scored positive on the CAM-ICU and the RASS (score ≥ -3)
- Three studies (Drame 2008; Pitkala 2005; Rockwood 1999) reported that delirium was classified based on DSM-IV criteria
    - One study (Rockwood 1999) used the Delirium Rating Scale
- Two studies (Andrew 2005; Francis 1990) reported that delirium was classified based on DSM III-R.
- One study (Holmes 2000) used the MMSE to identify patients with cognitive impairment and the Delirium Rating Scale was used to differentiate between delirium and dementia
- One study (Levkoff 1992) used the Delirium Symptom Interview (DSI) which assesses the domains of delirium specified in DSM III
- One study (O'Keeffe 1997) used the Delirium Assessment Scale (DAS), based on the DSM-III criteria for delirium

Partially inadequate

- One study (Rudolph 2008) reported that delirium was classified based on DSM III.
    - The method of delirium assessment was not consistent: patients were assessed with MMSE and medical records until postoperative day 3 and from day 4 until discharge, evaluation was based on the medical and nurse chart
    - Criterion 5 of the DSM-III was not a requirement ['evidence, from the history, physical examination, or laboratory tests of a specific organic factor judged to be etiologically related to the disturbance']. Primary caregiver or other informant was interviewed to identify symptoms that were new or had worsened within the week before hospital admission.

Inadequate

- One study (Dolan 2000) had a review of medical notes and/or proxy interview using CAM [proxies were family members or friends who could report on the patient's health]

The GDG considered the Dolan (2000) study to be biased because the method of assessment was based on review of medical notes and/or interview with proxy. The GDG agreed that the three studies (Levkoff 1992; O'Keefe 1997; Rudolph 2008) which used the DSM III (or methods based on DSM III) for assessment were acceptable if the method of assessment remained consistent throughout the duration of the study. However, in comparing with other studies, these studies should be treated with caution.

Assessment of severity

One study (Marcantonio 2002) used the Memorial Delirium Assessment Scale (MDAS) (range 0 to 30, with 30 indicating high severity) to assess severity of delirium and used 12.44 [the median of the average MDAS score for all patients with delirium] as the cut-off point between mild and severe delirium. Results were presented by severity of delirium.

9.3.4 Methodological quality of included studies

One study (Pitkala 2005) was considered to be truly representative of the population (i.e. adults in long-term and hospital settings) and the remaining studies were considered to be somewhat representative of the population.

The non-exposed cohort was drawn from the same community as the exposed cohort.

9.3.4.1 Missing data by outcome

- Dementia
  - One study (Rockwood 1999) reported less than 20% missing data (i.e. acceptable levels) for the outcome dementia;
  - One study (Rudolph 2008) reported less than 20% missing data (i.e. acceptable levels) for the outcome postoperative cognitive dysfunction at 7 days;
  - One study (Rudolph 2008) reported less than 20% missing data for the outcome postoperative cognitive dysfunction at 3 months, and here the authors showed that the 19% of missing data was not missing at random because those with delirium were twice as likely not to complete the testing, which indicates potential for bias;
  - One study (Ely 2004) was considered to have too high levels of missing data for the outcome cognitive impairment (28%) – these

patients were not tested because of their inability to complete testing or because of rapid discharge. This also indicates potential for bias.

- New admission to institution
    - Five studies (Balas 2009; Bourdel-Marchasson 2004; Inouye 1998: at discharge; O'Keeffe 1997; Pitkala 2005) reported less than 20% missing data (i.e. acceptable levels). In one study (Balas 2009) the missing data were due to patients remaining in hospital at the time of study closure and voluntary withdrawal from the study. In the remaining studies, the missing data were due to deaths;
    - One study (Inouye 1998) had about 20% missing data at 3 months follow up, but most of these were due to death or being lost to follow up: the missing group reportedly did not differ significantly from the completing group;
    - The level of missing data was not reported in one study (Levkoff 1992).

- Mortality
    - Seven reports of 6 studies had no missing data (Holmes 2000 [Nightingale 2001]; Inouye 1998- discharge; Levkoff 1992; Marcantonio 2000: 1 month; O'Keeffe 1997; Rockwood 1999);
    - Eleven studies stated there was less than 20% missing data (i.e. acceptable levels) (Dolan 2000; Ely 2004; Drame 2008; Francis 1990; Inouye 1998: 3 months; Leslie 2005; Lin 2004; Lin 2008; Marcantonio 2000: 6 months; Pitkala 2005; Thomason 2005).
- Length of stay
    - Three studies (Ely 2004: hospital; O'Keeffe 1997: hospital; Thomason 2008: hospital and ICU) reported less than 20% missing data (i.e. acceptable levels);
    - One study (Ely 2004: post ICU) had 29% missing data because of deaths in ICU and patients in a persistent coma. The former (10%) may have biased the outcome, but was at a low level;
    - Holmes (2001) reported no missing data.
- Hospital acquired complications

  - One study (O'Keeffe 1997) had no missing data.

- Mortality or new admission to institution

  - Three reports of two studies (Givens 2008 at 1 month and 6 months; Marcantonio 2000: 1 month; McAvay 2006: 1 year) had no missing data;

  - Two studies (Inouye 1998- discharge; 3 months; Marcantonio 2000: 6 months) reported less than 20% missing data.

- Mortality or functional decline

  - One study (Andrew 2005) reported no loss to follow up for the outcome at discharge and less than 20% loss to follow up at 6 months.

Assessment of delirium

As discussed above, the GDG considered that 18 studies had an adequate method of assessment; one had a partially inadequate method (Rudolph 2008) and one was inadequate (Dolan 2000).

Outcome of interest at baseline

- Dementia

  - One study (Rockwood 1999) excluded patients with dementia from the analysis.

  - One study (Ely 2004) assessing cognitive impairment reported the baseline modified Blessed Dementia rating score [range: 0 to 17] (mean (SD): 0.23(SD0.8) versus 0.14 (SD 0.6) for the delirious and non-delirious groups, respectively, indicating none of the patients were likely to have dementia.

  - One study (Rudolph 2008) assessing postoperative cognitive dysfunction reported that patients with a score of less than 23 on the MMSE were excluded but did not provide baseline scores for the neuropsychological tests used to assess postoperative cognitive dysfunction.

- New admission to institution

  - Five studies (Bourdel-Marchasson 2004; Inouye 1998; Levkoff 1992; O'Keeffe 1997; Pitkala 2005) reported patients in long-term care settings at admission were excluded from the analysis for this outcome.

    - In one study (Balas 2009) patients in long-term care setting at admission [3.5%: 4/114] were included in the analysis

- Hospital acquired complications (falls, pressure sores, urinary incontinence and any other complication)

    - One study(O'Keeffe 1997) reported patients with a pressure sore corresponding to Grade 2 of Shea's classification (Shea 1975) on admission were excluded; patients with frequent incontinence or with a catheter on admission were excluded from the analysis; history of falls was not reported;

- Mortality or new admission to institution

    - Mortality: not applicable;

    - New admission to institution:

        - One study (McAvay 2006) excluded patients admitted to hospital from a nursing home

        - Three reports of two studies (Inouye 1998; Marcantonio 2000; Marcantonio 2002) reported new admission to institutions for patients who had not been previously institutionalised at time of admission

- Mortality or functional decline

    - mortality: not applicable;

    - functional decline: the mean baseline Barthel index score was 86.6 (range 42 to 100) indicating some patients had less likelihood of living independently prior to hospitalisation (Andrew 2005)

Confounders taken into account:

The overall quality rating of the study was made taking into account the number of key risk factors, the method of delirium assessment, missing data in addition to the ratio of events to covariates.

All the included studies conducted multivariate analyses. The Marcantonio (2000) and Givens (2008) studies reported the same outcomes but adjusted for different variables in the multivariate analysis.

In relation to the events to covariate ratio, the GDG provided the following guidance:

- ratio of 1 or less: biased;

- ratio of 2 or 3: possibly confounded and rated as low quality;
- ratio of 4 to 7: moderate quality feature;
- ratio of 8 to 10: high quality feature.

The rest of this section examines the ratio of events to covariates and the number of key risk factors for each outcome.

A. Risk factor: presence of prevalent or incident delirium

1. Dementia/cognitive impairment/cognitive dysfunction

The GDG identified age, depression, and cognitive impairment as the key confounding factors. None of the studies included depression in the analyses, and studies were not downgraded if this risk factor was missing.

- One study had 2/3 of the important risk factors taken into account in the multivariate analysis, or held constant and the ratio of events to variables was at least 10. Patients with an MMSE score of 23 or less were excluded from the study.
    - Rudolph (2008) ratio: 66 [265/4]; [7 days postoperative dysfunction]; 24 [94/4]; [3 months postoperative dysfunction]; key factor was age, and cognitive impairment was constant
- Two studies had 2/3 of the important risk factors taken into account in the multivariate analysis but had an insufficient ratio of events to variables.
    - Ely (2004) ratio: 5 [63/12]; key risk factors were: age, cognitive impairment (dementia);
    - Rockwood (1999) ratio: 8 [32/4]; key factor was: age ; patients with dementia excluded from analysis

2. Progression of dementia

The GDG identified age and gender as the key confounding factors. There were no studies identified reporting this outcome.

3. New admission to an institution

The GDG identified ADL, cognitive impairment, and depression as the key confounding factors. None of the studies included depression in the analyses.

- Three studies had 2/3 of the important risk factors taken into account in the multivariate analysis and had a ratio of number of events to variables of at least 10.

  - Bourdel-Marchasson (2004) ratio: 10 [117/12]; key factors were: ADL, cognitive impairment [prevalent and incident delirium]
  - Inouye (1998) ratio:11 [77/7]; [3 month follow up]; key factors were: ADL, cognitive impairment
  - Pitkala (2005) ratio: 10 [72/7]; key factors were: ADL, cognitive impairment [dementia]

- Three studies had 2/3 of the important risk factors taken into account in the multivariate analysis but had insufficient ratio of events to variables.
  - Inouye (1998) ratio: 9 [60/7]; [at discharge]; key factors were: ADL, cognitive impairment
  - O'Keeffe (1997) ratio:5 [35/7]; key factors were: ADL, cognitive impairment
  - Balas (2009) ratio: 3 [35/13] ; key factors were: ADL, dementia

- One study had only one of the important risk factors taken into account in the multivariate analysis and had an insufficient ratio of events to variables.
  - Levkoff (1992) ratio: 6 [30/5]; key factor was: cognitive impairment

4. Falls

The GDG identified age, gender, polypharmacy and cognitive impairment as the key confounding factors. There were no studies identified reporting this outcome. Falls are, however, included as part of the hospital acquired complications outcome.

5. Hospital admission (for those who were in long-term care)

The GDG identified age, gender, cognitive impairment, severity of illness and/or comorbidity as the key confounding factors. There were no studies identified reporting this outcome.

6. Post discharge care

The GDG identified ADL, living alone and cognitive impairment as the key confounding factors. There were no studies identified reporting this outcome.

7. Post traumatic stress disorder

There were no studies identified reporting this outcome.

8. Pressure Ulcers

The GDG identified age, gender, and immobility as the key confounding factors. There were no studies identified reporting this outcome. Pressure ulcers are part of the hospital acquired complications outcome.

9. Mortality

The GDG identified age, gender, cognitive impairment, and severity of illness as the most important confounding factors. The GDG made a post-hoc decision to exclude gender as a confounding factor. Their decision was based on the findings from the non pharmacological risk factors, which showed gender was not a significant risk factor for the incidence of delirium.

- Three studies had all 3 important risk factors taken into account in the multivariate analysis and had a ratio of events to variables of at least 10
    - Inouye (1998): ratio: 14 [98/7] [3 months]; key risk factors were: age, severity of illness, cognitive impairment [dementia]
    - Levkoff (1992): ratio:12 [59/5]; key factors were: age, cognitive impairment, severity of illness
    - Nightingale (2001): ratio: 38 [347/10] [2 years]; key risk factors: age, dementia, physical illness [report of Holmes 2000]

- Four studies had 2/3 of the important risk factors taken into account in the multivariate analysis and had a ratio of events to variables of at least 10
    - Dolan (2000): ratio: 62 [369/6]; key factors were: age, cognitive impairment [cognitive impairment held constant as patients with cognitive impairment excluded]
    - Drame (2008): ratio: 11 [135/12]; key factors were: age, cognitive impairment [dementia]
    - Pitkala (2005): ratio: 15 [106/7][1 year]; ratio:28 [198/7] [2 years]; key factors were: age, cognitive impairment [dementia]
    - Rockwood (1999): ratio: 11 [101/9]; key factors were: age, cognitive impairment

- Four studies had all of the important risk factors taken into account in the multivariate analysis but had an insufficient ratio of events to variables.
    - Holmes (2000): ratio: 9 [195/ 22] [6 months]; key factors were: age, dementia, physical illness
    - Ely (2004): ratio:6 [69/12]; key factors were: age, severity of illness, dementia
    - Inouye (1998): ratio 5 [35/7][discharge]; key risk factors were: age, severity of illness, cognitive impairment [dementia]
    - O'Keeffe (1997): ratio: 3 [22/7] [in hospital]; 7 [49/7] [for 6 months]; key factors were: age, severity of illness, cognitive impairment [dementia]

- Three studies had 2/3 of the important risk factors taken into account in the multivariate analysis but had an insufficient ratio of events to variables.
    - Thomason (2005): ratio: 5 [32/7]; key factors were: age, severity of illness
    - Francis (1990): ratio: 4 [24/6]; key factors were: cognitive impairment, severity of illness [Unclear which factors were adjusted for in the multivariate analysis therefore factors reported for length of stay analysis were used]
    - Marcantonio (2000): ratio:1 [3/5] [1 month]; ratio: 3 [15/5] [6 months]; key factors were: age, cognitive impairment

- Two studies had only one of the important risk factors taken into account in the analysis and had a ratio of events to variables of at least 10
    - Francis (1992): ratio: 14 [55/4]; key factor was: cognitive impairment
    - Leslie (2005): ratio: 35 [208/6]; key factor was: age

- Two studies had only one of the important risk factors taken into account in the analysis and had an insufficient ratio of events to variables
    - Lin (2004): ratio: 6 [40/7]; key factor was: severity of illness, although patients with a history of chronic dementia were excluded from the study

- Lin (2008): ratio: 6 [59/10]; key factor was: age

10. Impact on carers

The GDG identified cognitive impairment and disability as the important confounding factors.
There were no studies identified reporting this outcome.

11. Length of stay

The GDG identified age, comorbidity and/or severity of illness as the important confounding factors:

- Five studies had all of the important risk factors taken into account in the multivariate analysis and had ratio of at least 10
    - Ely (2004): ratio: 19 [224/12] [length of stay-hospital]; key factors were: age, comorbidity and severity of illness
    - Ely (2004): ratio: 16 [196/12] [Post-ICU stay]; key factors were: age, comorbidity and severity of illness
    - Levkoff (1992): ratio: 42 [211/5] [community]; 23 [114/5] [institution]; key factors were: age, severity of illness
    - Holmes (2000): ratio: 33 [731/22] [risk of discharge sooner, i.e. decreased risk of remaining in hospital]; key factors were: age, physical illness
    - O'Keeffe (1997) ratio: 32 [225/7]; key factors were: age, severity of illness, comorbidity
    - Thomason (2005): ratio: 37 [260/7]; [length of stay-hospital and length of stay-ICU]; key factors were: age, comorbidity and severity of illness

- One study had one of the important risk factors taken into account in the multivariate analysis and had ratio of at least 10
    - Francis (1990): ratio: 38 [229/6]; key factor was: severity of illness

12. Quality of life

The GDG identified cognitive impairment and disability as the important confounding factors. There were no studies identified reporting this outcome.

13. Hospital acquired complication [urinary incontinence, falls, pressure sores or any other complications)

The GDG identified age, gender, polypharmacy, cognitive impairment [factors previously identified for falls] and/or age, gender, immobility [factors previously identified for pressure sores] as the important confounding factors:

- One study had 2/5 of the confounding factors taken into account in the multivariate analysis but had a ratio of at least 10
    - O'Keeffe (1997): ratio: 32 [225/7]; key factors were: age, cognitive impairment

14. Mortality or new admission to institution

The GDG identified ADL, age, cognitive impairment, comorbidity, severity of illness as the important confounding factors:

- Three studies had all/most (4 or 5) of the important risk factors taken into account in the multivariate analysis and had ratio of at least 10
    - Inouye (1998): ratio: 14 [95/7] at discharge; ratio: 24 [165/7] at 3 months; key factors were: ADL, age, cognitive impairment [dementia], severity of illness
    - McAvay (2006) ratio: 22 [198/9] key factors were: ADL, age, comorbidity, dementia, severity of illness
    - Pitkala (2005): ratio: 48 [336/7]; key factors were: age, ADL, dementia, comorbidity [outcome: mortality or *residing in institution at 2 years*]

- One study had all/most (4 or 5) of the important risk factors taken into account in the multivariate analysis but had insufficient ratio of events to variables.
    - Marcantonio (2000): ratio: 7 [33/5] [mortality or admission to nursing home at 1 month]; ratio: 6 [28/5] [mortality or admission to nursing home at 6 months]; key factors were: age, cognitive impairment, ADL, comorbidity

- One report of the Marcantonio (2000) study had 3/5 of the important risk factors taken into account in the multivariate analysis but had insufficient ratio of events to variables.
    - Givens (2008): ratio: 5 [33/7] [mortality or admission to nursing home at 1 month]; key factors were: age, ADL, comorbidity

- Givens (2008): ratio: 4 [28/7] [mortality or admission to nursing home at 6 months]; key factors were: age, ADL, comorbidity

B. Risk Factor: Increased duration of delirium

For this risk factor it was assumed that the other key risk factors for the various outcomes were the same as for the incidence of delirium

1. Mortality

- One study had all of the important risk factors taken into account in the multivariate analysis but had insufficient ratio of events to variables.
    - Ely (2004) ratio:6 [69/12]; key factors were: age, severity of illness, dementia

2. Length of stay

- One study had all of the important risk factors taken into account in the multivariate analysis and had ratio of at least 10
    - Ely (2004): ratio: 19 [224/12] [Length of stay: hospital]; key factors were: age, comorbidity and severity of illness
    - Ely (2004): ratio: 16 [196/12] [Length of stay: Post-ICU stay]; key factors were: age, comorbidity and severity of illness

3. Mortality or Functional decline

The GDG identified age, cognitive impairment and severity of illness as the key confounding factors for the composite outcome mortality or functional decline.

- One study had not enough risk factors (1/3) taken into account in the multivariate analysis but the ratio of events to covariate was at least 10
    - Andrew (2005): ratio: 12 [48/4] [6 months]; key factor was: age
- One study had not enough risk factors (1/3) taken into account in the multivariate analysis and the ratio of events to covariate was insufficient
    - Andrew (2005): ratio: 8 [32/4] [discharge]; key factor was: age

C. Risk Factor: Severity of delirium

For this risk factor it was assumed that the same key risk factors applied as for the incidence of delirium

1. Mortality

- One study had 1/3 confounding factors for mortality but the ratio of events to covariates was at least 10
    - Leslie 2005 ratio: 30 [208/7]; key factor was: age

2. Mortality or new admission to institution (for people who were in hospital)

- One report of the Marcantonio (2000) study had 2 of the 5 confounding factors for mortality or nursing home placement but had an insufficient ratio of events to variables.
    - Marcantonio (2002): ratio: 7 [22/3] [1 month]; ratio: 6 [17/3] [6 months]; key factors were: ADL and cognitive impairment

Overall quality assessment

Overall, the risk of bias was considered for each cohort study for each outcome, and a rating was given of high, moderate, low quality, and biased/confounded.

Four studies were judged to be biased for the following outcomes and therefore not considered further:

- Mortality (Dolan 2000: 2 years; Marcantonio 2000: 1 month)
- Dementia (Cognitive impairment: Ely 2004 at discharge; Cognitive dysfunction: Rudolph 2008)

The Marcantonio (2000) study was considered biased because there were more variables than events for the mortality outcome (at 1 month); the Dolan (2000) study was considered biased for the outcome mortality (at 2 years) because the method of delirium assessment was judged to be inadequate; the Rudolph (2008) study for the outcome cognitive dysfunction because of partially inadequate method of assessment of delirium and for the outcome cognitive dysfunction at 3 months, the study had missing data that was influenced by the presence of the prognostic factor; the Ely (2004) study had 29% missing data, which was attributed to an unexpected discharge or an inability to complete testing; inability to complete testing may have been related to the presence of delirium.

Thirteen reports of ten studies were given a low overall rating for the following outcomes and were treated with caution:

- Hospital acquired complications (O'Keeffe 1997)
- New admission to institution (Balas 2009; Levkoff 1992)
- Mortality (Francis 1990 - 6 months [Francis 1992- 2 years]; Leslie 2005 [incidence and severity of delirium]; Lin 2004; Lin 2008; Marcantonio 2000: 6 months; O'Keeffe 1997: in hospital; Thomason 2005)
- Mortality or new admission to institution (Givens 2008: 1 month and 6 months)
- Mortality or new admission to institution (Marcantonio 2002; severity of delirium)
- Mortality or functional decline (Andrew 2005; duration of delirium)
- Length of stay (Francis 1990)

Ten studies were given a moderate rating for the following outcomes:

- Dementia (Rockwood 1999)
- New admission to institution (Bourdel-Marchasson 2004; Inouye 1998: discharge and 3 months; O'Keeffe 1997)
- Mortality (Drame 2008: 6 week; Ely 2004 [incidence and duration of delirium]; Holmes 2000 - 6 months; Inouye 1998: discharge; 3 months; Levkoff 1992; O'Keeffe 1997: 6 months; Pitkala 2005: 1 year and 2 years; Rockwood 1998)
- Length of stay (Ely 2004:post ICU [incidence and duration of delirium]; Levkoff 1992)
- Mortality or new admission to institution (Inouye 1998: 3 months; Marcantonio 2000- 1 month and 6 months; Pitkala 2005- 2 years)

Eight reports of 7 studies were given a high rating for the following outcomes:

- New admission to institution (Pitkala 2005)
- Mortality (Nightingale 2001- 2 years)
- Length of stay (Ely 2004: hospital [incidence and duration]; Holmes 2000 [discharged from hospital earlier]; O'Keeffe 1997; Thomason 2005: hospital and ICU)

- Mortality or new admission to institution (Inouye 1998: discharge; McAvay 2006 - 1 year; Pitkala 2005: mortality or residing in long-term care at 2 years)

9.4 Results

Two studies (Andrew 2005; Ely 2004) reported the dependence of adverse consequences on the duration of delirium; two studies (Leslie 2005; Marcantonio 2002) reported the effects of increased severity of delirium and the remaining studies examined incidence of delirium as a prognostic factor.

Factors included in the multivariate analyses are given in Appendix F.

The following outcomes have been investigated:

- Risk Factor: Presence of prevalent and incident delirium
    - Dementia (1 study)
    - Progression of dementia (no studies)
    - New admission to an institution (6 studies)
    - Hospital admission (for those who were in long-term care) (no studies)
    - Post discharge care (no studies)
    - Pressure Ulcers (no studies) but see hospital acquired complications
    - Falls (no studies) but see hospital acquired complications
    - Mortality (16 reports of 14 studies)
    - Impact on carers (no studies)
    - Length of stay (6 studies)
    - Quality of life (no studies)
    - Hospital acquired complications (1 study)
    - Mortality or new admission to an institution (5 reports of 4 studies)

- Risk factor: Increased duration of delirium
    - Mortality (1 study)

  - Length of stay (1 study)
  - Mortality or functional decline (1 study)

- Risk factor: Severity of delirium
  - Mortality (1 study)
  - Mortality or new admission to an institution (1 study)

9.4.1 Risk factor: presence of prevalent or incident of delirium

Dementia

One moderate quality study (Rockwood 1999) reported dementia as a consequence of delirium at 3 year follow-up.

The Rockwood (1999) study reported 21% (32/154) of the patients developed dementia; the median follow-up period in the was 32.5 months.

Cognitive impairment was evaluated with MMSE (range 0 to 30), the Blessed dementia rating scale (range 0 to 17; higher score indicative of greater degree of dementia) and dementia was determined by a geriatrician. Information on patients who had died by follow-up was obtained through the IQCODE interviews from proxy informants. The study did not clarify who the proxies were.

This study in 203 patients showed that dementia was a significant consequence of delirium at 3 years follow up [OR 5.97 (95% CI 1.83 to 19.54)]; the confidence interval is wide (figure 9.1, Appendix K)

New admission to institution

Six studies (Balas 2009; Bourdel-Marchasson 2004; Inouye 1998; Levkoff 1992; O'Keeffe 1997[incident delirium only]; Pitkala 2005) reported new admissions to an institution following discharge. Two studies (Balas 2009; Levkoff 1992) were low quality, three were moderate quality (Bourdel-Marchasson 2004; Inouye 1998; O'Keeffe 1997 [incident and prevalent delirium]) and one study was high quality (Pitkala 2005).

The studies reported new admission to an institution following discharge from hospital (Inouye 1998; Levkoff 1992), at 3 months (Inouye 1998), 6 months (O'Keeffe 1997) and during 2 years (Pitkala 2005).

The number of patients admitted to an institution ranged from 3% (20/692) at discharge (Inouye 1998) to 36% (Pitkala 2005: 72/200) at 2 years.

The studies varied in their consideration of the key risk factors (ADL, cognitive impairment). Further information on these factors is reported in Appendix F. None of the studies reported including depression as a factor in the multivariate analysis.

Two studies (Inouye 19998; O'Keeffe 1997) reported excluding deaths for this outcome; one study (Balas 2009) reported patients who died within 24 hours of SICU admission were not considered for enrolment and one study (Bourdel-Marchasson 2004) reported the number of patients discharged either back to community or institution taking into account the number of deaths.

The odds ratio was generally around 2.8 and appeared to be fairly independent of when this was measured. The results suggest that new admission to an institution is a significant consequence of delirium (figure 9.2a, Appendix K).

A sensitivity analysis was undertaken (figure 9.2b, Appendix K) excluding the low quality studies. Three moderate quality study studies (Bourdel-Marchasson 2004 (n=427); Inouye 1998 (n=727); O'Keeffe 1997 (n=225)) and one high quality study (Pitkala 2005 (n=425)) were included. At discharge, the odds ratio ranged from 2.64 (95% 0.83 to 8.45) (Bourdel-Marchasson 2004: incident delirium) to 3.19 (95% CI 1.33 to 7.64) (Bourdel-Marchasson 2004: prevalent delirium). One study (Pitkala 2005) showed a significant effect of delirium on new institutionalisation at 2 years following discharge [adjusted OR 2.45 (95% CI 1.2 to 4.9)].

Mortality

Sixteen reports of 14 studies (Drame 2008; Ely 2004; Francis 1990 [Francis 1992:2 years]; Holmes 2000 [Nightingale 2001: 2 years]; Inouye 1998; Leslie 2005; Levkoff 1992; Lin 2004; Lin 2008; Marcantonio 2000; O'Keeffe 1997; Pitkala 2005; Rockwood 1999; Thomason 2005) reported mortality following delirium. Most studies did not state the cause of death, with the exception of two studies (Lin 2004; Drame 2008) which reported death from all causes.

Eight reports of seven studies were of low quality (Francis 1990: 6 months [Francis 1992: 2 years]; Leslie 2005; Lin 2004; Lin 2008; Marcantonio 2000: 6 months; O'Keeffe 1997: in hospital; Thomason 2005) and treated with caution; there were 8 studies of moderate quality (Drame 2008; Ely 2004; Holmes 2000: 6 months; Inouye 1998: hospital and 3 months; Levkoff 1992; O'Keeffe 1997: 6 months; Pitkala 2005; Rockwood 1998) and one report of the Holmes (2000) study was rated as high quality (Nightingale 2001: 2 years).

Information on the key factors (age, cognitive impairment, severity of illness) adjusted for in the multivariate analysis are presented in Appendix F.

Three studies reported death in hospital (O'Keeffe 1997; Inouye 1998; Thomason 2005). Of these, only the results from the O'Keeffe (1997) study will be considered as the GDG stated that only UK results are applicable for this outcome at discharge, however, the other studies are also shown on the forest plot for information (figure 9.3a, Appendix K).

Of the studies reporting mortality following discharge from hospital or ICU, eight reports of seven studies included hospital deaths (Drame 2008; Ely 2001; Francis 1990; Inouye 1998; Marcantonio 2000; Holmes 2000; Nightingale 2001; O'Keeffe 1997), three studies excluded death in hospital (Francis 1992 2.6% [6/229]; Leslie 2005: 1.5% [14/919]; Rockwood 1999 12.9% [32/247 enrolled]) and was unclear in two studies (Levkoff 1992; Pitkala 2005)

The number of patients who were in long-term care when they died was considered for the following time points:

- 6 weeks
    - In one study (Drame 2008), 17% of the patients [218/1306] were admitted from long-term care. It is unclear how many patients were discharged back into long-term care or if there were any new admissions and how many people died in long-term care.
- 3 months
    - In one study (Inouye 1998), of the 4% [29/77] patients admitted from long-term care it was unclear how many patients were discharged back into long-term care. Of those newly admitted to long-term care at discharge 8.7% [60/692], it is unclear how many people died there in the follow up period of 3 months. At 3 month follow-up, all deaths in hospital and at 3 months were excluded for the outcome new admission to long-term care.
- 6 months
    - In one study (Ely 2004) it was unclear if any patients were admitted to long-term care following discharge from ICU.
    - One study (Francis 1990) reported 7% (16/226: 16% vs 3.4% for the delirious and non delirious groups, respectively) of the patients were discharged to nursing homes, personal-care homes and rehabilitation facilities. The study also reported the percentages at 6 month follow-up [12% and 5% for the delirious and non delirious groups, respectively]. It is unclear how many patients in long-term care died.
    - In one study (Holmes 2000), of the patients who were diagnosed with delirium and living in non-residential setting at admission [76%: 82/108], 23% [19/63] were discharged to a residential or nursing home. It is unclear how many of these patients in long-term care died during the 6 month follow up period.
    - The Levkoff (1992) study reported 15% [30/203] of the community-dwelling patients with incident delirium were discharged to institution. It is unclear how many patients died in long-term care.

  - The Marcantonio (2000) study reported the proportion of patients who died was 12% [15/126] at 6 months.

- 1 year

  - In the Leslie (2005) study, of the 222 patients who died during the study period, 9.5% (21/222) were nursing home residents at admission. It is unclear whether all patients were discharged back into long-term care and subsequently how many died there.

  - In the Pitkala (2005) study, of the 53% [224/425] patients assessed in long-term care, it is unclear how many of these patients died in the first year during the course of the study.

- 2 years

  - In Francis (1992) it is unclear how many of the patients discharged to long-term care (as reported in Francis 1990) were followed up or how many died in the long-term care setting.

  - Pitkala 2005- Of the 53% [224/425] patients assessed in long-term care or the 36% of the patients [72/200] newly admitted to long-term care during the course of the 2 years, it is unclear how many of these patients died in long-term care. The study reported that 79% of the patients [336/425] were residing in institutional care or died during 2 years.

- 3 years

  - One study (Rockwood 1999) reported that, of the patients [101/203] who died during the 3 year follow-up, 79% (30/38) had delirium. Of the patients with delirium who died, the study reported 70% of the patients (21/30) were in institutional care.

The risk of mortality as a consequence of delirium varied with time as shown in the forest plot (figure 9.3a, Appendix K).

A sensitivity analysis was undertaken excluding the low quality studies. Results for one study (O'Keeffee 1997) set in the UK, irrespective of quality, has also been included for the outcome mortality in hospital (figure 9.3b, Appendix K). There is a significant effect of delirium incidence on mortality, which appears to be independent of time.

Length of stay

Two high quality studies (Holmes 2000; O'Keeffe 1997), one moderate quality study, (Levkoff 1992) and one low quality study (Francis 1990) reported length of stay in hospital. Two high quality studies (Ely 2004; Thomason 2005) reported length of stay in hospital (including the period in ICU), one high quality study

(Thomason 2005) reported length of stay in the ICU and one study (Ely 2004) reported length of stay post ICU (moderate quality for this outcome). The Ely (2004) study defined post ICU length of stay as the time after first ICU discharge.

The Holmes (2000) study, reported the relative risk of being discharged earlier, which corresponds to a decreased length of stay.

Three studies (Francis 1990; Levkoff 1992; O'Keeffe 1997) reported length of stay, adjusted for confounding factors in a multivariate analysis and gave p-values. The Levkoff (1992) study reported that delirium contributed to a longer length of stay both for patients admitted from the community ($t=4.03$; $p=0.0001$; 30.9 days and 7.4 days for the delirious and non delirious groups, respectively) and from long-term care ($t=4.48$; $p=0.0001$; 10.6 days and 6.9 days for the delirious and non delirious groups, respectively). The Francis (1990) study reported that delirious patients stayed in the hospital longer than the non delirious group (12.1 days versus 7.2 days, for the delirious and non delirious groups, respectively; $p< .001$). The O'Keeffe (1997) study reported that delirium was the only significant predictor of duration of hospital stay in a multivariate analysis (accounting for 6.7% of the variance; adjusted $t=3.8$, $p<.001$). The mean length of stay was 21 days and 11 days, for the delirious and non delirious groups, respectively ($p<.001$).

The median length of stay in hospital and interquartile range (IQR) were reported in the Ely (2004) study [21 days (IQR 19 to 25): 11 days (IQR 7 to 14) for the delirious and non delirious groups, respectively] and the Thomason (2005) study [median 5 days (IQR 2 to 8) and 3 days (IQR 2 to 6) for the delirious and non delirious groups, respectively]. In the Ely (2004) study, length of stay was measured from admission for prevalent delirium patients and from time of diagnosis for incident delirium patients.

The median length of stay in ICU and interquartile range (IQR) was reported in the Thomason (2005) study [median 4 days (IQR 3 to 5) and 3 days (IQR 2 to 4) for the delirious and non delirious groups, respectively].

The median length of post ICU stay and interquartile range (IQR) was reported in the Ely (2004) study [median 7 days (IQR 4 to 15.5) and 5 days (IQR 2 to 7) for the delirious and non delirious groups, respectively].

One study (Holmes 2000) reporting discharge from hospital, showed the likelihood of discharge was decreased in the presence of delirium, leading to an increased length of stay [RR 0.53 (95% CI 0.41 to 0.68); figure 9.4a, Appendix K].

The adjusted hazard ratio ranged from 1.41 (95% CI 1.05 to 1.89) to 2.0 (95% CI 1.4 to 3.0) showing increased length of stay in hospital to be a significant consequence of delirium for patients who had been in ICU (figure 9.4b, Appendix K).

There was no significant effect on length of stay in ICU [HR 1.29 (95% CI 0.98 to 1.69)] but there was an effect of delirium on post-ICU stay [HR 1.6 (95% CI 1.1 to 2.3); figure 9.4b, Appendix K].

Hospital acquired complication *[urinary incontinence, falls, pressure sores or any other complication]*

One low quality study (O'Keeffe 1997) reported results for hospital acquired complications. The percentages of patients with complications were as follows: urinary incontinence: 46% (86/206); falls: 12.4% (28/225); pressure sores: 4% (8/202) or any other complications: 44% (100/225). The multivariate analysis adjusted for age, chronic cognitive impairment, severity of illness, comorbidity, disability score and length of stay.

The study reported that falls, pressure sores (corresponding to grade 2 Shea classification) and urinary incontinence (new onset or worsening after admission to hospital) were identified based on interviews with nursing staff. The authors defined a fall as 'unintentionally coming to rest on ground … not as a result of an obvious major intrinsic event (such as stroke or syncope) or overwhelming hazard.'

The result showed that hospital acquired complications is a significant consequence of delirium [OR 2.3 (95% CI 1.7 to 5.0); figure 9.5, Appendix K].

Mortality or new admission to institutions

Five reports of four studies (Inouye 1998; McAvay 2006; Marcantonio 2000 [Givens 2008]; Pitkala 2005) reported a composite outcome of mortality or new admission to institution. The Givens (2008) report of the Marcantonio (2000) study and the Marcantonio (2000) study reported results for the same cohort but the multivariate analyses were adjusted for different factors. The Givens (2008) report only gave the adjusted odds ratio and p values. The standard error was calculated, on a trial and error basis, based on the reported p values.

Three studies were high quality (Inouye 1998 at hospital discharge; McAvay 2006; Pitkala 2005), two were of moderate quality (Inouye 1998 at 3 months; Marcantonio 2000), and the Givens (2008) report of the Marcantonio (2000) study was low quality. The Pitkala (2005) study reported mortality or residing in institution at 2 years.

Rates of the composite outcomes (mortality and new admission to institution) and the rates for each outcome, where reported, were as follows:

In hospital: 13% (Inouye 1998:95/727; mortality: 5% [35/727]; new admission: 9% [60/692])

- 1 month: 26% (Marcantonio 2000: 33/126; mortality: 2% [3/126])
- 3 months: 25% (Inouye 1998: 165/663; mortality: 14% [98/680]; new admission: 13% [77/600])
- 6 months: 23% (Marcantonio 2000: 28/123; mortality: 12% [15/123]);

- 1 year: (McAvay 2006)
    - o delirium at discharge: 83% [20/24]; (mortality: 38% [9/24]; new admission: 79% [19/24]);
    - o delirium resolved: 68% [21/31]; (mortality: 26% [8/31]; new admission: 45% [14/31]);
    - o never delirious: 42% [157/378]; (mortality: 20% [75/378]; new admission: 29% [111/378]).

At discharge from hospital, one multicentre study set in the US (Inouye 1998 - high quality) showed there was a significant effect of delirium on the composite outcome, mortality or new admission to institution [OR 2.1 (95% CI 1.1 to 4.0] however, the confidence interval is fairly wide.

At three months, one moderate quality study (Inouye 1998) showed a significant effect of delirium [OR 2.6 (95% CI 1.4 to 4. 5)]; however, the confidence interval is fairly wide.

One moderate quality study (Marcantonio 2000) and one low quality study (Givens 2008 showed a significant effect at one month with adjusted odds ratio ranging from 3.0 (95% CI 1.1 to 8.4)] to 4.26 (95% CI 1.49 to 12.16), however, the confidence interval was wide.

Two studies (Givens 2008; Marcantonio 2008) showed no significant effect at 6 months; adjusted odds ratio ranging from 1.80 (95% CI 0.62 to 5.25) to 2.17(95% CI 0.73 6.49)

The McAvay (2006) study reported the results at 1 year for those with delirium at discharge, resolved delirium and never delirious. There was a significant effect at 1 year [patients with delirium at discharge compared with those never delirious] [HR 2.64 (95% CI 1.60 to 4.35)] but the confidence interval is wide. In patients with delirium resolved compared with those never delirious and in patients with delirium at discharge compared with delirium resolved there was no significant effect at 1 year (figure 9.6, Appendix K).

9.4.2 Risk Factor: Increased duration of delirium as a continuous variable

Mortality

One moderate quality study (Ely 2004) reported mortality at 6 months as a consequence of duration of delirium. The study used duration of delirium as a continuous risk factor in the multivariate analysis. The results relate to each additional day of delirium for ICU patients.

There was a borderline significant effect of duration of delirium on mortality [HR 1.1 (95% CI 1.0 to 1.3); figure 9.7, Appendix K]. For each extra day with delirium, the hazard ratio increases by 1.10, so that if there were 3 extra days it would become $(1.10)^3$ (i.e. 1.33).

Length of stay

One study (Ely 2004) reported length of stay (hospital [high quality] and post-ICU stay[moderate quality]) as a consequence of increased duration of delirium.

The study used duration of delirium as a continuous risk factor in the multivariate analysis. The results relate to each additional day of delirium for ICU patients.

The length of ICU plus hospital stay was significantly greater for patients who had longer periods of delirium [HR 1.20 (95% CI 1.1 to 1.3)] and the post-ICU stay was of borderline significance [HR 1.10 (95% CI 1.0 to 1.2); figure 9.8, Appendix K].

Mortality or functional decline

One low quality study (Andrew 2005) reported a composite outcome of incomplete functional recovery or death following an episode of delirium. Functional decline was defined as a decrease by at least 10 points on the Barthel Index (BI) compared with the baseline BI score.

The results were presented for duration of delirium, adjusted for age, gender, and frailty. Frailty was assessed on the geriatric severity score (ranging from healthy and independent to terminally ill). Further information on these factors are presented in Appendix F. Mean duration of delirium was 6.3 days (range 1 to 35). The mean pre morbid (baseline) Barthel Index score was 86.6 (range 42 to 100), with an 8.9 point decrease at discharge and a 12.7 decline in score at 6 months.

The study reported that at discharge the mortality rate was 8% (6/77) and functional decline was reported in 37% (26/71) of the patients. At 6 months, 68% of the patients (48/71) had an outcome of death or functional decline.

Mortality or functional decline was a borderline significant consequence of increased duration of delirium at hospital discharge [OR 1.1 (95% CI 1.0 to 1.2)] and at 6 months [OR 1.2 (95% CI 1.0 to 1.4); figure 9.9, Appendix K].

9.4.3 Risk factor: severity of delirium as a categorical outcome

Mortality

One low quality study (Leslie 2005) reported the effect of severity of delirium, assessed during hospitalisation, on mortality at 1 year.

The mortality rate of patients with more severe delirium was 40% (16/40), 30.3% (80/264) for those with less severe delirium and 18.5% (110/596) for those who were never delirious.

At 1 year, increased severity (assessed during hospitalisation) had a significant effect on mortality compared with no delirium [HR 1.89 (95% CI 1.13 to 3.14)]. Less severe delirium (assessed during hospitalisation) also had a significant effect [HR 1.62 (95% CI 1.21 to 2.17); figure 9.10, Appendix K].

Mortality or New admission to institution

One low quality study (Marcantonio 2002) reported mortality or discharge to a care home at 1 month and 6 months. The study examined the effect of severity of delirium in patients with CAM defined delirium and those with non-delirious symptoms (some had subsyndromal delirium). The results for the former group (n= 49) are reported here.

Mortality or new admission to institution at 1 month was 33% (8/24) and 56% (14/25) for the mild and severe delirium groups, respectively. At 6 months mortality or new admission to institution was 17% (4/24) and 52% (13/25) for the mild and severe delirium groups, respectively

At 1 month, severe delirium compared with delirium had no significant effect on mortality or nursing home placement [OR 1.90 (95% CI 0.50 to 8.0)]. At 6 months, the confidence interval is very wide [OR 4.4 (95%CI 0.9 to 21.1); figure 9.11, Appendix K], and there is too much uncertainty to draw conclusions.

Overall summary

Table 9.5 shows the results for the key outcomes reported in the consequences of delirium review that were chosen as sources of data for the baseline risks considered in the economic model.

Table 9.5: summary of the results: consequences of delirium

Presence of incidence or prevalent delirium		
Consequences	**Details**	**Odds ratio (95% CI)**
Dementia	Assessed at 3 years	5.97 (1.83 to 19.54)
New admission to institution	Assessed at discharge	2.64 (0.83 to 8.45)
Mortality	Assessed mortality 'In hospital'	2.60 (0.7 to 6.2)
Hospital acquired complications	Including falls, urinary continence, pressure sores or any other complications	2.3 (1.7 to 5.0)
Mortality or new admission to institution	Assessed at 1 month	3.0 (1.1 to 8.4)

9.5 Health economic evidence

No relevant health economic papers were identified.

9.6 Clinical evidence statements

- Dementia
    - There is moderate quality evidence to show that dementia is a significant consequence of delirium at 3 year follow-up.

- New admission to institution
    - There is moderate quality evidence to show that new admission to institution is a significant consequence of delirium, which appears to be independent of time.

- Mortality
    - There is moderate quality evidence to show that mortality is a significant consequence of delirium, which appears to be independent of time.
    - There is low quality evidence to show that mortality was a significant consequence of increased severity of delirium (assessed during hospitalisation).
    - There is moderate quality evidence to show there is a borderline significant effect of duration of delirium on mortality.

- Length of stay
  - There is high quality evidence to show that:
    - the likelihood of discharge was decreased in the presence of delirium, leading to an increased length of stay in hospital.
    - an increased length of stay in hospital is a significant consequence of delirium for patients who had been in ICU.
    - post-ICU stay is a significant consequence of delirium.
  - There is moderate quality evidence to show that:
    - there is a significantly increased length of ICU plus hospital stay for patients who had longer periods of delirium.
    - there is a borderline significant effect on increased length of post-ICU for patients who had longer periods of delirium.
  - There is high quality evidence there is no significant effect of delirium on the length of stay in ICU.
- Hospital acquired complications
  - There is low quality evidence to show that hospital acquired complications [pressure sores, falls, urinary incontinence or any other complication] are a significant consequence of delirium.
- Composite outcome: Mortality or new admission to institution
  - There is high quality evidence to show that there is a significant effect of delirium on the composite outcome, mortality or new admission to institution at discharge from hospital and mortality or residing in institution at 2 years; there is some uncertainty around these results.
  - There is high quality evidence to show that there is a significant effect of persistent delirium on the composite outcome, mortality or new admission to institution, at 1 year; there is some uncertainty around this result.
  - There is moderate quality evidence to show that there is a significant effect of delirium on the composite outcome, mortality or new admission to institution at 3 months following discharge from hospital; there is some uncertainty around this result.
  - There is low to moderate quality evidence to show that there is a significant effect of delirium on the composite outcome, mortality or new admission to institution, at 1month and at 6 months

following discharge from hospital; there is some uncertainty around these results.

  - There is low quality evidence to show that an increased severity of delirium had no significant effect on the composite outcome, mortality or new admission to institution, at 1 month and at 6 months (there is too much uncertainty around this result) following discharge from hospital.

- Mortality or functional decline

  - There is low quality evidence to show that mortality or functional decline was a borderline significant consequence of increased duration of delirium at discharge from hospital and at 6 months following discharge.

9.7 From evidence to recommendations

There was low and moderate quality evidence from the consequences of delirium review for patients in hospital (this evidence informed the economic model), but no evidence for the consequences of delirium in long-term care.

The GDG considered the evidence noting that dementia, length of stay, death and new admission to long-term care were all significant consequences of delirium. The GDG felt that awareness of this information was very important, but acknowledged that a recommendation could not be made stating ‘be aware of the consequences of delirium’.

They recognised the difficulty of implementing and auditing a recommendation based on ‘awareness’. So as not to lose this important message, the GDG agreed that “Think delirium” should appear as a prominent statement at the start of the list of recommendations. The following paragraph was agreed by the GDG:

“THINK DELIRIUM”

Be aware that people in hospital or long-term care may be at risk of delirium. This can have serious consequences (such as increased risk of dementia and/or death) and, for people in hospital, may increase their length of stay in hospital and their risk of new admission to long-term care.

The GDG proposed a research recommendation (see below and Appendix H) to investigate the occurrence of delirium in the long-term care setting, and the consequences of delirium in that population.

Future research recommendation: How common is delirium and what are its adverse outcomes in people in

long-term care?

9.8 Recommendations

There are no recommendations for this section. In light of the evidence the GDG did not wish to make recommendations.

10 Prevention of delirium: non-pharmacological

Clinical introduction

Prevention of any harmful condition is clearly desirable, and delirium is no exception. Unfortunately, the introduction of delirium prevention protocols into routine care has been slow, partly because the existing research evidence base is fragmented and not well known to clinicians. Delirium prevention is similar in many respects to the issue of pressure sore prevention in the 1980s when the NHS was content to spend considerable amounts on the treatment of pressure sores and largely ignore prevention strategies. The prevention of pressure sores required specific and well-supported clinical policies to foster a new culture of prevention with the adoption of new procedures and skills in routine care.

A useful practical approach to the understanding of delirium has been to consider patient vulnerability (risk factors) in relation to stressor events (delirium precipitants). Thus, the precipitants do not alone cause an episode of delirium; they interact with the underlying risk factors. This clinical model suggests that interventions designed to reduce the impact of selected delirium risk factors might be associated with a reduction in delirium incidence. This section reviews the evidence for this approach – for single risk factors (single component interventions), and for multiple risk factors (multicomponent interventions).

10A) Single component prevention: hydration and music

CLINICAL QUESTIONS:

What are the most clinical and cost effective single-component, non-pharmacological interventions for the prevention of delirium in people in long-term care?

What are the most clinical and cost effective single-component, non-pharmacological interventions for the prevention of delirium in people in hospital?

10A. 1. HYDRATION FOR THE PREVENTION OF DELIRIUM (LONG-TERM CARE SETTING)

10.1 Description of studies

10.1.1 Study Design

Two papers were evaluated for inclusion and both were included: one (Mentes 2003) described a cluster randomised trial: four nursing homes were randomised to intervention or control groups; and the other (Robinson 2002) was a before-and-after study, in which the patients were monitored 2 weeks pre-intervention, then received 5 weeks of the intervention, followed by 2 weeks post-intervention study.

Both studies were conducted in the USA and both received funding from non-industry sources. There were 49 patients in the Mentes (2003) study and 51 in the Robinson (2002) study.

10.1.2 Population

Both studies took place in a long-term care setting. In the Mentes (2003) study, patients with acute confusion at baseline were excluded. Nine of 24 people in the intervention group and two of 25 in the control group had a diagnosis of cognitive impairment, although it was not specified how this was diagnosed or defined. In the Robinson (2002) study, it was unclear how many participants had cognitive impairment. Sensory impairment was not reported in either study.

In the Mentes (2003) study, the mean number of drugs daily was 6.4 in the intervention group compared with 7.1 among controls (not significantly different) and in the Robinson (2002) study 80% (41/51) had more than four drugs prescribed. It was not stated whether all eligible patients were included in either study.

The mean age in the Mentes (2003) study was around 82 years and it was 83.5 years in the Robinson (2002) study. The Mentes (2003) study included 22 men and 27 women, and the Robinson (2002) study had 8 men and 43 women. Ethnicity was reported in the Mentes (2003) study: all participants were Caucasian except for one who was African American. The Robinson (2002) study did not report ethnicity.

10.1.3 Interventions

In the Mentes (2003) study, the intervention was an 8-week hydration management intervention. This was based on calculating a daily individual fluid goal for each participant adjusted for his or her weight. For the intervention group, methods for ensuring that a participant met their goals included a standardised 180 ml fluid intake with each medication administration, fluid rounds morning and evening and ’happy hours‘ or ’tea time‘ twice a week in the late afternoon. The control group patients' fluid goals were also assessed and they received 'usual care', described as 'standard nursing care'.

The Robinson (2002) study gave the participants a hydration programme which consisted of the following components: a caregiver knowledgeable in techniques of fluid administration; an individualised plan of care incorporating the most effective techniques to administer fluids; a colourful beverage cart with colourful pitchers and glasses to enhance residents' interest in drinking; and a choice from 2 beverages at each encounter. Residents had a goal of 8 oz twice per day, but 47% did not achieve this goal every time.

10.1.4 Comparison

Hydration intervention versus usual care; outcomes recorded after 8 weeks (Mentes 2003). Concurrent medications were not reported in the Mentes (2003) study.

10.2 Methodological quality

10.2.1 RCTs

In the RCT (Mentes 2003), the method of randomisation to intervention or control was at the level of the nursing home and was by coin toss. Allocation concealment was unclear. No account was taken in the analysis of the fact that this was a cluster randomised trial, and there are likely to be unit of analysis errors.

It was assumed that patients were not blinded to treatment allocation. Blinding of outcome assessors was unclear. In the intervention group, the assessments appeared to be carried out by the research nurses involved in delivery of the intervention (i.e. not blinded), but in the control group, the assessment was carried out by the research nurses blinded to the patient's fluid goals; whether they were aware of the research question is not clear.

The study did not report an *a priori* sample size calculation and its small size and short duration suggest that it may have been underpowered.

The authors demonstrated baseline comparability of the groups on some measures (age, gender, number of diagnoses, mean number of daily medications, depression), but significant differences between the groups on several measures although there were confounders would be likely to negate differences between interventions. The intervention group scores on the NEECHAM Confusion Scale indicated that they were more at risk for delirium than the control group (mean 26.4 versus 28.4, p=0.005). This scale ranges from 0 to 30, where a score of less than 25 indicates confusion, and 26 to 27 indicates at risk of confusion. The treatment group had more patients with a diagnosis of cognitive impairment (9 versus 2, p=0.02) and the treatment group were more physically frail than the control group (mean scores 79.4 versus 112.2, p<0.001) on the Functional Independence Measure (FIM) instrument; (scale score ranges from 0 to 126; not specified for long-term care but higher values indicate better function). In addition, the mean length of stay for the intervention group in long-term care was 22.9 months compared with 94.9 months for control group patients.

It is noted that, cognitive impairment, a risk factor for delirium, was greater at baseline for the intervention group than the control group. The risk factors review had inconsistent evidence regarding whether long-term care was a risk factor for delirium, and functional status was not investigated as a risk factor for delirium.

All patients were followed up for the 8 weeks of the trial and all patients' data were analysed.

The primary outcome measure for the study was 'hydration-linked events', defined as acute confusion, urinary tract infection, upper respiratory infection, pneumonia or influenza, preceded by a urine specific gravity of 1.020 or above and decreased fluid intake as measured by intake records.

Delirium assessment was triggered if a participant exhibited a sudden change in mental status, or a cognitive or behavioural change. A participant was considered acutely confused if he or she scored lower than baseline on the MMSE and lower than 25 on the NEECHAM Confusion Scale. The GDG considered the MMSE to be an inadequate method of assessment of delirium.

The differences in baseline comparability between the groups, the randomisation by nursing home with only four nursing homes involved and the delirium assessment method mean that this study is at higher risk of bias.

10.2.2 Non-randomised study

The Robinson (2002) study was a before-and-after, prospective study. It was unclear if all eligible participants were included. In addition, the method of assessing delirium was not reported and, indeed, results for this outcome were not given. Overall, the nature of the design meant this was poor quality evidence.

10.3 Results

10.3.1 Hydration intervention versus usual care

Incidence of delirium

The Mentes (2003) study reported no delirium in the treatment group during the 8 weeks of treatment compared with 2 people in the control group (figure 10.1, Appendix K). The confidence interval is very wide and is consistent with both significant benefit and significant harm due to the small number of events and so there is uncertainty about the effect of the intervention on this outcome.

Other outcomes

The primary outcome measure of the Mentes (2003) study was 'all hydration-linked events', and these were urinary tract infections (1 in the control group), upper respiratory infections (2 in the control group), pneumonia (1 each in the intervention and control groups) and influenza (2 in the intervention group) (figure 10.2, Appendix K). The results are again very imprecise.

The non-randomised study (Robinson 2002) reported that the outcomes measured improved significantly with the hydration intervention: these were an increase in the number of bowel movements ($p = 0.04$); a reduction in laxative use ($p = 0.05$); and a decline in the number of falls ($p = 0.05$).

At 8 weeks, concordance was 95% in the intervention group for their fluid goals compared with 89% of controls ($p=0.08$), (Mentes 2008).

10.4 Health economic evidence

10.4.1 Single component non-pharmacological intervention for the prevention of delirium in a long-term care setting

One economic evaluation study was included as evidence (Robinson 2002). This was a before-and–after study of 51 older adults in the USA. The aim of the study was to determine the effect of a specific program on the level of hydration, and on the prevention of conditions associated with dehydration, namely, delirium, urinary tract infections, respiratory infections, falls, skin breakdown, and constipation. Patients in the intervention group were enrolled in a hydration programme to improve hydration. The programme included a hydration assistant to administer fluid, an individualised plan of care incorporating the most effective techniques to administer fluid, a colourful beverage cart with colourful pitchers and glasses to enhance residents' interest in drinking, and a choice from 4 beverages at each encounter. The goal was for each resident to consume an additional 8-ounce beverage mid-morning and mid-afternoon, which would increase fluid intake to 1.5L daily.

Patients in the control group received usual gray coloured institutional carts, white foam cups and limited variety of beverages. The cost of colourful cups and assorted beverages was \$154 per week, and \$3 per resident per week. The average cost of employee time per week per resident was \$8. The intervention resulted in a cost savings of \$103 over the 5 week period as a result of fewer negative outcomes for patients. There was no report on the delirium incidence or severity, mortality or HRQoL. This study did not adequately report the main outcomes of interest. The results of this study are not directly applicable.

10A. 2. HYDRATION FOR THE PREVENTION OF DELIRIUM (HOSPITAL SETTING)

10.5 Description of studies

One paper was included (O'Keeffe 1996).

10.5.1 Study Design

This study was an RCT conducted in the UK. The study did not report on funding, and 60 patients were included.

The study compared the effectiveness and tolerability of two methods of delivering fluids; it was not concerned with preventing delirium. The study is therefore included as indirect evidence, which may inform GDG discussion.

10.5.2 Population

The study took place in an acute geriatric unit. Patients suffering from mild dehydration or poor oral intake, requiring parenteral fluids for at least 48 hours and who had cognitive impairment were included. Cognitive impairment was defined as disorientation for time and place or an MMSE score of 20 or less. Patients were excluded if there was clinical evidence of poor tissue perfusion or if the amount of fluid administered would be critical (e.g. in those with renal or heart failure).

The mean age was 82.5 years and 38% were male. Ethnicity was not reported.

10.5.3 Interventions

In the O'Keeffe (1996) study the patients were randomised to receive either subcutaneous or intravenous fluids. Up to 2 litres of fluid were permitted in a 24 hour period.

10.5.4 Comparison

Subcutaneous fluids versus Intravenous fluids; outcomes recorded at 48 hours. Concurrent medications were not reported.

10.5.5 Outcome measures

The review's primary outcome measure was incidence of delirium. However, this included study did not give this outcome, but reported on agitation, serum urea and serum creatinine levels at 48 hours and the incidence of local oedema.

10.6 Methodological quality

The O'Keeffe (1996) study reported an adequate method of randomisation (table of random numbers) and a partially adequate method of allocation concealment (sealed envelope).

Blinding of patients would not have occurred due to the method of intervention. Blinding of outcome assessors was unclear.

The study reported an *a priori* sample size calculation. In order to detect a difference in serum urea of 1.5mmol/l between the two groups, at 80% power and 5% significance level, it was estimated that a sample size of 56 patients would be required; the study included 60 patients.

Baseline comparability was reported on age, gender, serum urea, serum creatinine levels, and baseline agitation levels. Agitation levels were assessed by a doctor using the modified Cohen-Mansfield Agitation Inventory based on personal observations and discussion with nurses or carers regarding the behaviour of the patient during the previous 48 hours.

There was less than 20% missing data, one patient in the subcutaneous group died and one patient in the intravenous group was switched to the subcutaneous route after 24 hours because of difficulties with venous access. These patients were not included in the analysis.

Overall, the study was considered not to be at higher risk of bias, although it only reported indirect outcomes.

10.7 Results

10.7.1 Subcutaneous versus intravenous hydration

Agitation

There was a large significant effect of the method of hydration in relation to agitated behaviour, with significantly fewer patients experiencing agitation related to the subcutaneous method of hydration; RR 0.46 (95% 0.28 to 0.76) (figure 10.3, Appendix K). There was some imprecision in the result.

Serum urea and creatinine levels

The study reported the serum urea and serum creatinine levels for both groups at 48 hours. For serum urea, there was no significant difference between interventions; mean difference (MD) -0.27 mmol/l (95% CI - 0.78 to 0.24)]. There was also no significant difference between the serum creatinine levels at 48 hours; MD 0.31 μmol/l (95% CI -0.82 to 0. 2); figure 10.4, Appendix K.

Local Oedema

The O'Keeffe (1996) study reported that local oedema was noted in two patients receiving fluids subcutaneously. The confidence interval is very wide due to the small number of events and there is insufficient evidence to draw conclusions about the effect of different hydration strategies on this outcome [OR 5.00 (95% CI 0.25 to 99.95)] (figure 10.5, Appendix K).

10.8 Health economic evidence

No relevant health economic papers were identified.

10.9 Clinical evidence statements

Long-term care setting

There is very low quality evidence showing that a hydration intervention had no significant effect on:

- the incidence of delirium.
- hydration linked events (urinary tract infection, upper respiratory, pneumonia, influenza).

However there is a lot of uncertainty around these results.

Hospital setting

There is moderate quality evidence comparing subcutaneous and intravenous methods of hydration to show:

- significantly lower levels of agitation in patients receiving fluids subcutaneously compared with intravenously.

- no significant difference was found in levels of serum urea or serum creatinine levels.

10.10 From evidence to recommendations

The GDG noted that the evidence on hydration in long-term care was limited. In addition to the evidence review, the GDG were aware that drinking-water regimens in long-term care settings indicated an improvement in the well-being of the residents. They considered writing a stand-alone hydration recommendation for all patients in long-term care but felt on balance the hydration recommendation would be better placed in the as part of the tailored multcomponent intervention package.

The GDG considered a single study (O'Keeffe 1996) of moderate quality evidence comparing hydration strategies in the hospital setting that reportedagitation as an outcome. Although the GDG agreed that the study was useful in the consideration of hydration strategies, the agitation outcome could not be extrapolated to delirium. It was agreed that strategies for hydration would be captured in the recommendation for tailored multicomponent prevention intervention package [1.3.3.2]

10.11 Recommendations

See recommendation 1.3.3.2.

10A. 3. MUSIC THERAPY FOR THE PREVENTION OF DELIRIUM (HOSPITAL SETTING)

10.12 Description of studies

Four papers were evaluated for inclusion. Two studies were excluded. Reasons for exclusions are reported in Appendix G. Two papers were included in this review (McCaffrey 2004; McCaffrey 2006).

10.12.1 Study Design

No studies were conducted in the UK; both were conducted in the USA. The study by McCaffrey (2004) used a non-probability convenience sample of 66 patients from a large tertiary care centre in south-east Florida. The McCaffrey (2006) study had a sample size of 124 patients from a hospital in Florida, but no further details were given. The McCaffrey (2004) study did not report the number of patients in the intervention or control groups.

10.12.2 Population

Both studies took place in a university hospital setting in the postoperative orthopaedic unit. Postoperative patients included were those undergoing elective hip or knee surgery, who were alert and oriented to provide consent, able to complete preoperative paperwork independently, and able to hear music.

Proportions of patients with low, intermediate and high risks of delirium at baseline were not reported in either of the studies. Neither delirium nor dementia status at baseline was reported.

The mean age of the patients was 75.7 years (SD 6; range 59 to 82 years) in the McCaffrey (2006) study and 73 years (SD 5) in the McCaffrey (2004) study.

In the McCaffrey (2006) study, there was a higher proportion of women (64.5%, 80/124) than men (35.5%, 44/124) and 67% of all patients had knee surgery (the rest had hip surgery). These details were not reported in the earlier study (McCaffrey 2004). Ethnicity was not reported in either of the studies.

10.12.3 Interventions

The interventions evaluated were:

- Music therapy: patients in individual rooms were given a bedside compact disc (CD) player that would automatically play music for a minimum of 1 hour, 3 times/day (McCaffrey 2004) or for a minimum of 1 hour, 4 times/day (McCaffrey 2006). The music started while the patient was awakening from anaesthesia and continued during the recovery period.
    - The McCaffrey (2004) study stated that the number of times that the CD could automatically be turned on was three times a day at the *most*, but that the *minimum* time was 1 hour, three times daily. In the study by McCaffrey (2006) the CD player would automatically play CDs for a minimum of 1 hour, 4 times daily.
    - In addition, nurses and family members were asked to turn on the music when they walked into the orthopaedic unit room.
    - Once awake and oriented, patients received the same instructions so they could play music when they desired.
    - The first CD placed in the player was chosen by the researcher. Other musical selections were available to the patients based on their musical preference.
    - Patients were visited by research assistants to ensure the CD players were working and that the times for automatic starting of the CD coincided with the patients' preference, and that the music playing was what the patient preferred.

Intervention and control groups in both studies had full access to in-room televisions, and both groups received standard postoperative care. Patients were not permitted to bring any electronic music devices into their hospital rooms.

10.12.4 Comparisons

The following comparison was carried out in both studies:

- Music therapy versus no treatment
  - Both groups received standard postoperative care
  - The total length of postoperative care was 3 days in both the intervention and control groups in one study (McCaffrey 2006), but was unclear in the other study (McCaffrey 2004).

10.13 Methodological quality

The method of sequence generation was not reported in either study; patients were randomly assigned to rooms that had been designated intervention or control; this was subject to room availability. Allocation concealment was considered to be adequate because the recovery room nurses who assigned patients to rooms were said to be unaware of the experimental and control group rooms' designation.

Blinding of the outcome assessor was unclear in both studies. It was not possible to blind the patients, but the GDG did not consider this to be important. *A priori* sample size and power calculations were not reported in either of the studies.

The McCaffrey (2006) study reported limited data on the demographic characteristics of the patients. Patients in each group were similar in age, proportion of men and women, and proportion of patients with hip and knee surgery. This was not reported in the McCaffrey (2004) study.

Only the McCaffrey (2006) study reported on withdrawals. 1.6% (2/126) patients were lost to follow-up due to cardiovascular complications during surgery, but missing data were not reported for individual groups. The McCaffrey (2004) study did not report whether an intention to treat (ITT) analysis was carried out, and the McCaffrey (2006) study used an available case analysis.

Both studies evaluated 'acute confusion' as a primary outcome, which was identified with delirium: nurses kept computerised notes, recording signs and symptoms of delirium. These nurse-identified signs and symptoms of delirium and confusion were reviewed retrospectively by researchers with the orthopaedic nursing staff to achieve consistency. In the McCaffrey (2004) study, the number of episodes of confusion and delirium were entered as a numerical score for that patient and the McCaffrey (2006) study recorded the number of patients with at least one episode of acute confusion. The GDG did not consider this to be a reliable measure of delirium assessment and so these studies were regarded with caution.

Overall, these studies were considered to have a higher risk of bias because neither had a validated method of assessing delirium incidence.

10.14 Results

10.14.1 Music therapy plus standard postoperative care versus standard postoperative care

Incidence of delirium

The McCaffrey (2004) study reported that patients receiving music therapy had significantly fewer periods of confusion or delirium during their hospitalisation than patients who received no additional therapy, and gave a p-value of 0.001 without detailing the results.

The McCaffrey (2006) study in 124 patients demonstrated that significantly fewer patients experienced acute confusion in the music therapy group. Although the CI was very wide, the results were not considered to be imprecise as far as decision making was concerned (figure 10.6, Appendix K); RR 0.06 (95% CI 0.01 to 0.22). This corresponds to an NNT of 2 (95% CI 2 to 3) for a control group rate of 58%.

Activities of daily life

Both studies assessed the patient's 'readiness to ambulate' during the postoperative period (McCaffrey 2004; McCaffrey 2006). An ambulation readiness profile was conducted by physiotherapists in both studies using postoperative scores ranging from 1 (indicating that patients were not ready to ambulate) to 10 (indicating that patients may be ready to ambulate that day or the next). The scores were based on: pain level; alertness; stable vital signs; ability to correctly identify person, place and time; ability to comprehend instructions; and willingness to participate in their own recovery.

The McCaffrey (2004) study found that patients receiving music therapy had significantly higher scores on the readiness to ambulate scale for the day of surgery than did patients who received no additional therapy, and reported a p-value of 0.001. No other details were given.

The McCaffrey (2006) study demonstrated that patients in the music therapy group had significantly higher scores for readiness to ambulate after undergoing surgery than patients in the control group (figure 10.7, Appendix K); MD 0.93 (95%CI 0.52 to 1.34). This is, however, a small effect even though significant.

Patient satisfaction

The McCaffrey (2006) study also measured patient satisfaction: the researcher phoned each patient 2 weeks after discharge from hospital to determine their satisfaction with their postoperative experience in the hospital. A scale of 1–10 was used (1 representing the worst experience and 10 the best experience they could imagine). Analysis showed a significantly higher score for the intervention group (figure 10.8, Appendix K); MD 2.77 (95%CI 2.38 to 3.16) for a control group score of 6.83.

10.15 Health economic evidence

No relevant health economic papers were identified.

10.16 Clinical evidence statements

- There is low quality evidence from one RCT comparing music therapy with usual care which showed:
    - A significantly lower incidence of delirium in the group receiving music therapy.
    - A higher score for readiness to ambulate after undergoing surgery in the music therapy group.
    - A higher score in patient satisfaction in the music therapy group.

10.17 From evidence to recommendations

The GDG considered the evidence which showed a significantly lower incidence of delirium in the group receiving music therapy compared with usual care. The GDG noted that the studies were at high risk of bias as an unvalidated method of assessing delirium incidence was used. The GDG did not want to make a recommendation based on this evidence and proposed music therapy should be considered as a future research recommendation (see below and Appendix H).

Future research recommendation:

Is music therapy that is tailored to the individual's preferences, more clinically and cost effective than non-tailored music or usual care in preventing the development of delirium in hospital patients at risk of delirium?

10.18 Recommendations

There are no recommendations for this section. In light of the evidence the GDG did not wish to make recommendations.

10B) Multicomponent prevention

CLINICAL QUESTIONS:

What are the most clinical and cost effective multicomponent interventions for the prevention of delirium in people in hospital?

What are the most clinical and cost effective multicomponent interventions for the prevention of delirium in people in long-term care?

10.19 Description of studies

Details of included and excluded papers together with study design are reported in table 10.1.

Table 10.1: study inclusion, exclusion and design

Papers	Comments	Study
N= 14 evaluated for inclusion		
N= 5 excluded	Appendix G.	
N= 0 identified in update searches	None Identified	
N= 9 reports of 8 studies included*	Study designs 3 RCTs	Landefeld 1995; Lundström 2005; Marcantonio 2001
	2 non-randomised designs	Inouye 1999; Wanich 1992
	3 historical controlled trials	Gustafson 1991; Harari 2007a; Wong 2005

*The Bogardus (2003) study was a six month follow up, post hospital discharge, of a sample of patients (705/852 (83%)) from the Inouye (1999) study. It appears that these patients were representative of the original sample; 133/852 (16%) had died.

10.19.1 Study Design

Information on study sizes and geographical location are described in table 10.2.

Table 10.2: study characteristics

Study	Size (N)	Geographical location
Gustafson 1991	214	Sweden
Harari 2007a	108	UK
Inouye 1999	852	USA
Landefeld 1995	651	USA
Lundström 2005	400	Sweden
Marcantonio 2001	126	USA
Wanich 1992	235	USA
Wong 2005	99	Australia

With the exception of the Wong (2005) study, all of the studies were supported by research grants not associated with industry. The Wong (2005) study did not state a funding source.

The unit of randomisation in all the RCTs was the patient. One of the non-randomised controlled studies (Wanich 1992) allocated patients to different wards (but did not say how this was done), and the Inouye (1999) study allocated patients by forming matched pairs, matched on age within 5 years, sex, and base-line risk of delirium (intermediate or high).

In the historical controlled trials (Gustafson 1991; Harari 2007a; Wong 2005), all eligible patients were enrolled at two different time periods.

All the studies compared a group of participants in the period before the intervention was given with a group who were given the intervention.

10.19.2 Population

All of the studies took place in hospital settings. In four studies, patients were undergoing surgery, either for hip fracture (Gustafson 1991; Marcantonio 2001; Wong 2005), or for hip, knee, or other replacements (Harari 2007a). The Harari (2007a) study intervention was targeted at-risk patients at higher risk of adverse events/illness (e.g. those with poorly controlled diabetes) and included those who had been assessed as being too 'medically unfit' to go on the waiting list; the control group were not selected in this way. The other studies included older people with acute medical illness (Inouye 1999; Landefeld 1995; Lundström 2005; Wanich 1992).

Comorbidities in patients undergoing surgery were reported in three studies: the Gustafson (1991) study reported that some patients also had cerebrovascular diseases, cardiovascular diseases, hypertension, diabetes, Parkinson's disease, renal failure, lung disease, on-going infection, urinary incontinence, constipation, prostatism, depression, and psychosis. The Harari (2007a) study reported that some of the surgical patients had rheumatoid arthritis, heart disease, heart failure, atrial fibrillation, diabetes, renal impairment, hypertension, chronic lung disease, prostate or bladder problems and cerebrovascular disease. The Wong

(2005) study reported that some patients had vascular disease, diabetes, chronic lung disease and/or depression/anxiety at baseline. Comorbidities were not specifically stated in the Marcantonio (2001) study; 39% in the intervention group and 33% in the control group were reported to have high medical comorbidity at baseline (Charlson index of at least 4).

Of the studies that examined older people with acute medical illness, reasons for hospitalisation included cardiac, respiratory, infection, metabolic, neoplasm, cerebrovascular, or other diagnoses (Inouye 1999; Landefeld 1995; Lundström 2005; Wanich 1992).

Medications taken at baseline were reported by the Gustafson (1991) and the Lundström (2005) studies. In the Gustafson (1991) study, drugs or groups of drugs taken by patients included digitalis, diuretics, antihypertensives, nitroglycerin, analgesics, steroids, antiasthma, sulfonylurea, insulin, warfarin, laxatives, antidepressants, neuroleptics, benzodiazepines, other sedatives, antiparkinson drugs and other drugs; in this study, 16% of patients were not taking drugs. The Lundström (2005) study also reported the proportions of patients taking digitalis, diuretics, beta-blockers, calcium blockers, insulin, analgesics, benzodiazepines and neuroleptics. None of the other studies reported details on medicine use at baseline (Harari 1997a; Inouye 1999; Landefeld 1995; Marcantonio 2001; Wanich 1992; Wong 2005).

All of the studies evaluated older patients. The age range across studies was 50 to 102 years, with the mean age, where given, ranging from 75 to 84 years. In almost all studies the majority of patients were women (Gustafson 1991: 74%; Harari 2007a: 60%; Inouye 1999: 61%; Landefeld 1995: 67%; Lundström 2005: 56%; Marcantonio 2001: 79%; Wong 2005: 72%). The Wanich (1992) study reported that the sex distribution was approximately equal. Ethnicity was reported in three studies (Inouye 1999; Landefeld 1995; Marcantonio 2001), in which 59 to 90% of patients were white. The Wanich (1992) study only reported that ethnic distributions were approximately equal.

The majority of studies (Gustafson 1991; Harari 2007a; Landefeld 1995; Lundström 2005; Marcantonio 2001; Wanich 1992; Wong 2005) did not explicitly report the proportions of patients with low, intermediate and high risks of delirium at baseline, although it may be inferred that many were at high risk. For example, the Marcantonio (2001) study included hip fracture patients. The Inouye (1999) study reported that 72% patients had an intermediate risk of delirium and 28% had a high risk: patients were defined as having intermediate risk if they had 1 or 2 risk factors and high risk if they had 3 or 4 risk factors from the following list: visual impairment, severe illness (APACHE II score more than 16), cognitive impairment (MMSE score below 24), high blood urea nitrogen to creatinine ratio of at least 18.

In the majority of studies, at least some patients were reported to have dementia: two studies (Inouye 1999; Lundström 2005) reported on cognitive function using the MMSE instrument (scale 0-30): the Inouye (1999) study reported a mean MMSE score of 24 (SD 5) in the treatment group and 23 (SD 5) in the control group. In the Lundström (2005) study, patients in the treatment and control groups both had an average score of 25 (SD 6). It is noted that a score of 20-26 indicates mild dementia or cognitive impairment. The Landefeld (1995) study reported using the MMSE scale for the first 21 items (scale of 0-21); they reported scores of 17 in both groups, and also reported that 11% had dementia at baseline. The Inouye (1999) study reported that 11% of the patients had dementia using a modified Blessed Dementia Rating Scale (>2), and the Marcantonio (2001) study reported that 40% of patients had dementia at baseline using the Blessed score (>4). The Lundström (2005) study reported that 5% of patients had dementia using DSM-IV criteria, and the Gustafson (1991) study reported that 22% in intervention group and 15% in the control group had dementia using the DSM-III criteria. The Wanich (1992) and the Wong (2005) studies reported using the MMSE score, but did not present any data. The Harari (1997a) study did not report cognitive function scores.

Three studies reported sight and hearing impairment at baseline (Gustafson 1991; Inouye 1999; Lundström 2005):

- In the Gustafson (1991) study, visual and hearing impairment was reported in 23% and 25% of the patients respectively (methods of assessment not stated).

- The Inouye (1999) study reported that visual and hearing impairment occurred in 23% and 26% of the patients respectively (as evaluated using the standard Jaeger test, and the Whisper test)

- The Lundström (2005) study reported that 2% of the intervention group and 4% of the intervention group had impaired hearing, and 15% to 17% had impaired vision. In this study, hearing impairment was considered if a patient could not hear a normal speaking voice within one metre or without a hearing aid, and impaired vision was considered if a patient could not read a newspaper without glasses.

It is also noted that 59% of patients in the Inouye (1999) study were dehydrated on admission.

10.19.3 Interventions

The interventions were largely education and/or management changes with structured protocols for patient care. Each intervention is described below. Additionally, in order to understand and compare the interventions more effectively we have carried out a themed analysis, breaking down the interventions by risk factors addressed, and whether or not a multidisciplinary team and educational interventions are described (see table 10.3).

Table 10.3: multicomponent interventions for the prevention of delirium: overview of different factors form each study identified

Study	Multi-disciplinary team	Education intervention	Care methods	assessment of patients	orientation	Dehydration nutrition	Sleep	Sensory impairment improvement	Early mobilisation	Environmental modifications	Medication management	Pain management	Other
Lundström (2005)	No; mainly nursing care	staff education on Ass; PTD: NPI; Med. Monthly guidance for staff	Patient-allocation care , with individualised care	yes: via education	only via education	No	No	No	No	No	only via education	no	No
Inouye (1999): Elder Life Program	Yes: N, Physio, G, TRS, V	yes: trained team; performance evaluated quarterly	not changed	Yes in order to determine risk factors addressed	Yes: schedule / name board; reorienting communi cation	Yes for those with dehydration: early recognition of dehydration and volume repletion	Yes: non-pharmacological sleep protocol; sleep-enhancement protocol	Yes for visually impaired and hearing impaired people	yes	yes: unit-wide noise reduction strategies	No	no	cognitively stimulating activities (e.g. discussion of current events)
Gustafson (1991): Geriatric-anesthesiologic intervention programme	Yes; nursing, anaesthetist and geriatrician care	No	not changed; task oriented	pre- and postop by geriatrician	No	No	No	No	No	No	individualised thrombosis prophylaxis	no	O2 therapy from admission; phenylephrine for low systolic bp; surgical policy
Harari 2007a: Proactive care of older people undergoing surgery (POPS)	Yes: N, Physio, G, OT, SW	Yes: patients preop (N, Ex, RT, PM); staff postop (TMC, EM, PM, BBF, N, DP)	no change	preop planning and postop review by geriatrician and nurse; targetting issues identified	No	Yes: nutrition	No	No	Yes	No	early detection and treatment of medical complications	Yes	discharge planning

Study	Multi-disciplinary team	Education intervention	Care methods	assessment of patients	orientation	Dehydration nutrition	Sleep	Sensory impairment improvement	Early mobilisation	Environmental modifications	Medication management	Pain management	Other
Landefeld (1995); Acute Care for Elders programme	yes: daily visits (D, N, SW, Diet, Physio, VNL)	No	patient centred care	Yes: daily assessment by nurses of physical, cognitive and psychosocial function; daily review of medical care	Yes: large clock, calendar	yes nutrition (no details)	yes (no details)	No	yes (no details)	Yes: specially designed environmt (carpeting, handrails, uncluttered hallways, elevated toilet seats, door levers)	yes: minimise medications (e.g. sedative-hypnotic agents	no	minimise effects of procedures (e.g. catheterisation); discharge planning
Wannich (1992):	yes: for discharge planning (N, Physio, OT, SW, Diet)	Yes: staff (Ass, SI, Mob, En); families (RC, O, En)	Not stated	Yes: assessment and management plans recorded on charts and shared with staff and families	Yes (e.g. day of week, current events, updated calendars in every room)	No	No	Yes for visually impaired and hearing impaired people only (glasses and hearing aid + encouragement to use them)	yes	Yes: lighting to decrease sensory deprivation; night lights	assess medicns contributing to delerium, e.g. neuroleptics, antidepressants, narcotic analgesics, sedative-hypnotics, and use discouraged	no	discharge planning; Communication: clear and slow, with repetition
Wong (2005)	Yes: project team supervised programme (N, G, Ph, D, QI, A, Diet)	Yes: staff on PTD, POD, Ass, MMD	not changed	Yes for identification of needs	Yes: clock, calendar	Yes: nutrition (including properly fitting dentures); maintenance of fluid/electrolyte imbalance	No	Yes: sensory stimuli - glasses, hearing aid	yes	Yes: soft lighting, not putting delirious patients in same room	treatment of major complications; stop unnecess benzodiazepines, antihistamines, anticholinergics	Yes	regulation of bladder / bowel function; O2; tmt of agitated delirium
Marcantonio (2001): Proactive geriatrics consultation	No consultation with geriatrician	no	Not stated	Yes: consultation with geriatrician preop / within 24 h postop. Geriatrician daily visits during hospitalisation => target recs made	Yes: clock, calendar	Yes: nutrition (including properly fitting dentures); maintenance of fluid/electrolyte imbalance; treat dehydration/ overload	No	Yes: sensory stimuli - glasses, hearing aid	yes	Yes: soft lighting, use of radio/tape recorder - not rec for any patient though	treatment of major complications; stop benzodiazepines, antihistamines, anticholinergics; eliminate medicn redundancies; tmt to raise bp	Yes	regulation of bladder / bowel function; O2; tmt of agitated delirium

Key: N = nurses; Physio = physiotherapists; OT = occupational therapists; D = doctor (generally); G = geriatrician; SW = social worker; TRS = therapeutic recreation specialist; V = volunteer; A = anaesthetist; QI = member of the quality improvement unit; Ph = pharmacist, Diet = dietitian / nutritionalist; VNL = visiting nurse liaison; Ass = assessment; PTD = prevention and treatment of delirium; CD = training on cognitive impairment; POD = prevalence and outcome of delirium; NPI = nurse patient interaction; N = nutrition; MMD = medication management of delirium; Ex = exercise; RT = relaxation therapy; PM = pain management; TMC = treatment of medical complications; EM = early mobilisation; PM = pain management; BBF = bowel bladder function; DP = discharge planning).

Education programme and reorganisation of nursing and medical care consisting of four parts (Lundström 2005):

- Two-day course for staff on geriatric medicine which focused on assessment, prevention and treatment of delirium and underlying causes (e.g. urinary tract infection); lectures started before the intervention, with a follow up during the first month of the study
    - training regarding medical interventions included focus on the prevention of hypoxaemia, hypercortisolism, and avoidance of drugs with anticholinergic properties
    - training regarding nursing interventions focused on interaction with patients with reduced attention and orientation in a stressful situation and optimisation of care for these patients
- Staff education on caregiver-patient interaction that focused on patients with dementia and delirium, particularly with respect to comprehension and orientation of the patients
- A patient-allocation nursing care system with individualised care (in which small teams of nurses had full responsibility for a small number of patients to promote continuity of care)
- Monthly guidance for nursing staff, focusing on caregiver-patient interaction
- The control ward received usual hospital care organised in a task allocated way.

'Elder Life Program' (Inouye 1999; Bogardus 2003)

This programme was implemented by a trained interdisciplinary team, consisting of a geriatric nurse-specialist, two specially trained Elder Life specialists, a therapeutic-recreation specialist, a physiotherapy consultant, a geriatrician and trained volunteers.

- The performance of each staff member was evaluated quarterly, with completion of checklists to ensure competency and consistent and complete adherence to protocols.
- This multidisciplinary team implemented the following interventions, which were targeted at particular risk factors:
    - Cognitive impairment; outcome: change in orientation score (first 10 items on MMSE)

    - an orientation protocol: schedule/name board; reorienting communication
    - therapeutic activities protocol: cognitively stimulating activities 3 times daily (e.g. discussion of current events, word games, structured reminiscence)

- Sleep deprivation: outcome: change in use of sedative drugs for sleep
    - non-pharmacological sleep protocol: at bedtime, warm drink, relaxation tapes/music, back massage
    - sleep-enhancement protocol: unit-wide noise-reduction strategies (e.g. vibrating beepers, quiet hallways) and schedule adjustments to allow sleep (e.g. medications)
- Immobility; outcome: change in Activities of Daily Living score
    - Early-mobilisation protocol: ambulation or active range-of-motion exercises 3 times daily; minimising use of immobilising equipment (e.g. bladder catheters; physical restraints)
- Visual impairment; outcome: early correction of vision up to 48 h after admission
    - vision protocol (for visually impaired people only): visual aids (e.g. glasses and magnifying lenses) and adaptive equipment (e.g. large illuminated telephone key pads. large print books, fluorescent tape on call bell), with daily reinforcement of their use
- Hearing impairment; outcome: change in Whisper Test score
    - hearing protocol (for hearing impaired people only): portable amplifying devices, earwax disimpaction, special communication techniques, with daily reinforcement of their use
- Dehydration; outcome: change in ratio of blood urea nitrogen to creatinine
    - dehydration protocol (for those with evidence of dehydration, i.e. ratio of blood urea nitrogen to creatinine of at least 18): early recognition of dehydration and volume repletion (e.g. encouragement of oral fluid intake)

Usual care was standard hospital services provided by a multidisciplinary team.

Education and multicomponent intervention (Wanich 1992), which consisted of:

- Nursing staff education in the month before the start of the study and repeated once during the study on mental and functional status assessments, nursing management of deficits in sensory-perceptual function, mobility and environmental modifications

- Patient assessment and management plans recorded on charts and shared with staff and families to assist in nursing care and discharge planning

- Families education and consultation including reassurance and coping skills; orientation and personalising the environment

- 2 geriatricians assigned to intervention group

- Orientation: provision of orientation cues to patients (e.g. day of week, current events, a discussion of their condition, information about upcoming diagnostic or therapeutic measures); updated calendars in every room; favourite TV programmes determined)

- Communication (families and nurses taught to communicate clearly and slowly, and to use repetition and orientation clues)

- Mobilisation (e.g. getting patients out of bed each day, ambulation daily, physical and occupational therapy as needed)

- Sensory stimuli (glasses and hearing aids available and nurses encouraged patients to use them)

- Environmental modifications (lighting to decrease sensory deprivation; night lights used)

- Medical management (to assess medications suspected of contributing to delirium, e.g. neuroleptics, antidepressants, narcotic analgesics, sedative-hypnotics, and their unnecessary use discouraged)

- Discharge planning (with multidisciplinary team: primary nurse, social worker, discharge planning nurse. physiotherapist, occupational therapist and dietitian)

- The control group received usual care, but also received the physical and occupational therapy components in similar proportion to the intervention group.

'Acute Care for Elders' programme (Landefeld 1995)

This was carried out in a special unit and consisted of:

- Daily assessment by nurses of physical, cognitive and psychosocial function; daily review of medical care
- Daily rounds by multidisciplinary team: medical and nursing directors, a primary nurse, a social worker, a nutritionalist, a physical therapist and a visiting-nurse liaison
- Protocols to improve self-care, continence, nutrition, mobility, sleep, skin care, mood, cognition (implemented by the primary nurse based on the daily assessment)
- Specially designed environment (carpeting, handrails, uncluttered hallways, elevated toilet seats and door levers)
- Orientation (large clocks and calendars)
- Patient-centred care
- Planning for discharge including early involvement of a social worker and home healthcare nurse if indicated
- Protocols to minimise the adverse effects of selected procedures (eg. urinary catheterisation) and medications (e.g. sedative-hypnotic agents)

The comparator was usual care in another general medical unit.

'Proactive care of older people undergoing surgery (POPS)' (Harari 1997a)

This was a multidisciplinary, preoperative, comprehensive geriatric assessment service with postoperative follow-through:

- Multidisciplinary team consisting of a consultant geriatrician, a nurse specialist in older people, an occupational therapist, a physiotherapist and a social worker
- Preoperative assessment: Abreviated Mental Test Score, Geriatric Depression Scale, Barthel Index, Timed Up and Go, 180° turn, body mass index, continence screen, orthostatic blood pressure, pain score, and peak expiratory flow rates. Then investigation and treatment targeted the identified issues and medical comorbidities were optimised according to evidence based practice.
- Management plans and goals were agreed with the patient, and post-discharge plans made preoperatively

- Most patients had preoperative home visits from the occupational therapist and the physiotherapist, providing aid and equipment

- Preoperative education of patients in optimising postoperative recovery including home exercises, good nutrition, relaxation techniques and pain management; mean number of preoperative clinic visits was 1.79 (range 1-4)

- Postoperative staff education on early detection and treatment of medical complications, early mobilisation, pain management, bowel-bladder function, nutrition and discharge planning

- Postoperative early detection and treatment of medical complications, early mobilisation, pain management, bowel-bladder function, nutrition and discharge planning

- Follow-up therapy home visit in those with functional difficulties, and outpatient clinic review in those with ongoing medical problems

Quality improvement programme (Plan-do-study-act methodology with interventions introduced incrementally) (Wong 2005)

- Project team consisting of a consultant and registrar geriatricians, a consultant anaesthetist, two clinical nurse managers, a member of the quality improvement unit, and representatives of allied health staff (pharmacist, dietitian) met approximately fortnightly to supervise the programme

- Staff education on definition of delirium, predisposing and precipitating factors, investigations (including use of CAM) and management of delirium

- Geriatric team made recommendations for each person, based on the following:

    - Regulation of bladder and bowel function (remove indwelling catheters, screening for constipation, retention) [recommended in 24%]

    - Early detection/treatment of major complications (myocardial ischaemia, infection, pulmonary embolism, etc) [recommended in 22%]

    - Maintenance of fluid and electrolyte balance [recommended in 14%]

    - Discontinuation of unnecessary medications (especially benzodiazepines, antihistamines, drugs with anticholinergic effects) [recommended in 14%]

- Maintenance of adequate oxygen delivery (oxygen and blood transfusion)
- Pain management
- Treatment of agitated delirium (including low dose haloperidol or lorazepam)
- Use of appropriate environmental stimuli (soft lighting, avoid putting delirious patients in the same room)
- Sensory impairment improvement (glasses, hearing aids)
- Orientation (clock, calendar)
- Adequate nutritional intake (dentures used properly, adequate positioning, dietitian review and intervention
- Early mobilisation and rehabilitation

Proactive geriatrics consultation (Marcantonio 2001)

This consisted of:

- A consultation with a geriatrician that began preoperatively, or within 24 hours postoperatively. Geriatrician made daily visits during hospitalisation at which time target recommendations were made using the following (it is noted that the recommendations were only made if the consultants noticed something that was not already being done):
- Adequate CNS oxygen delivery
    - oxygen therapy to keep saturation above 90%, treatment to raise systolic bp to above 2/3rds that at baseline or above 90 mm Hg; blood transfusion to keep haematocrit above 30% [applied to 73%]
- Fluid/electrolyte balance
    - Treatment to restore serum sodium, potassium, glucose to normal limits
- Treatment of dehydration or fluid overload
    - Detected by examination or blood tests [applied to 43%]
- Treatment of severe pain (regular paracetamol) and treatment of break through pain
- Elimination of unnecessary medication

- Discontinuation of benzodiazepines, anticholinergics, histamines [applied to 56%]
- Elimination of medication redundancies

- Regulation of bowel/bladder function
  - Removal of urinary catheter by postoperative day 2, with screening for retention or incontinence [applied to 63%]
- Nutritional intake
  - Dentures used properly [applied to 37%]
  - Nutritional supplements
  - Temporary nasogastric tube
- Early mobilisation [applied to 47%] and rehabilitation
- Prevention, detection and treatment of major postoperative complications
  - Including myocardial infarction/ischaemia, pneumonia/COPD, pulmonary embolism [applied to 50%], urinary tract infection
- Environmental stimuli
  - soft lighting and use of radio/tape recorder
  - but wasn't implemented for any patient in practice
- Sensory stimuli (glasses and hearing aid)
- Orientation (clock and calendar)
- Treatment of agitated delirium (including haloperidol or lorazepam)

The usual care group received management by the orthopaedics team, including internal medicine or geriatrics consultations, but on a reactive rather than proactive basis.

Geriatric-anaesthesiologic intervention programme (Gustafson 1991)

This involved the following:

- Surgical policy (patients were operated on as soon as possible after admission)

- Preoperative assessment: for all patients, mostly by a specialist in geriatric and internal medicine

- Individualised thrombosis prophylaxis: heart failure patients given Heparin, rest Dextran (c.f. control group all given Dextran)

- Diuretics: patients with clinical signs of heart failure were treated with extra doses of diuretics

- Oxygen therapy: nasal oxygen given soon after admission (1 l/min). Oxygen enriched air was given throughout the operation and the first postoperative day, and then continued or not depending on oxygenation levels

- Anaesthetic technique: all patients had sc morphine premedication and spinal anaesthesia; patients who had systolic blood pressure below 90 mm Hg were aggressively treated with phenylephrine

- Postoperative assessment: all patients were assessed several times by a geriatrician

- Treatment of patients developing delirium for complications associated with acute coronary syndrome (e.g. anaemia, heart failure, urinary retention) – this is expected to confound measurements on the duration of delirium and incidence of delirium at 7 days

- Wards: all patients admitted to the same ward (but not part of the study protocol)

- Nursing care in both groups treated according to task allocation system

10.19.4 Comparisons

The following comparison was carried out in all studies:

Multicomponent intervention versus usual hospital care

In Lundström (2005), 'usual hospital care' was task-oriented care (i.e. the same nurse handling particular tasks for all patients; meaning that several nurses could care for each patient each day) – for this study, the intervention was patient oriented care

10.20 Methodological quality

10.20.1 Randomised trials

The method of sequence generation was adequate in two RCTs: the Landefeld (1995) study employed a computer-generated sequence and the Marcantonio

(2001) study used a random numbers table. The Lundström (2005) study did not describe sequence generation.

Allocation concealment was partially adequate in the Marcantonio (2001) study, in which sealed envelopes were used. The method of allocation concealment was not stated in Landefeld (1995). The study by Lundström (2005) was an RCT in which patients were randomly allocated to any ward with an accessible bed (i.e. this may constitute some selection bias), so that intervention patients and controls were on different wards. The study stated that the staff and assessors knew to which wards the patients were allocated, i.e. there was inadequate allocation concealment.

Due to the nature of the interventions, none of the RCTs were patient blinded. The Marcantonio (2001) study reported that the outcome assessor was blinded to the intervention status of the patients, and the Landefeld (1995) study stated that data were obtained by means of interviews and the interviewers were not blinded to the patients' group assignments. The Lundström (2005) study stated that the outcome assessors were blinded for delirium diagnosis, but were not blinded otherwise.

The Marcantonio (2001) study reported an *a priori* sample size calculation to detect the incidence of delirium; they required a sample size of 125 to detect a 33% decrease in risk with 80% power (they had sample size of 126). The Landefeld (1995) and the Lundström (2005) studies did not report *a priori* sample size calculations.

In the Landefeld (1995) study, 36% (651/1974) of eligible patients were randomised; 1143 eligible patients were not enrolled because beds were not available in the intervention or control wards at the time of their admission. In the Marcantonio (2001) study, 85% of eligible patients were included; of 149 eligible patients, 23 refused to participate. In the Lundström (2005) study, all eligible patients were randomised.

The Marcantonio (2001) and the Landefeld (1995) studies demonstrated baseline comparability of the groups. In the Lundström (2005) study, there were more females in the intervention ward ($p = 0.04$), a higher mean age in the control ward ($p = 0.02$), a greater proportion of patients previously diagnosed with diabetes mellitus on the intervention ward ($p<0.001$), and a greater proportion of patients diagnosed with myocardial infarction on the intervention ward ($p = 0.03$). The GDG did not consider these to be important differences.

In the Landefeld (1995) study, 7% of patients in both the intervention and control groups were lost to follow-up. In both these studies, the authors only analysed data from available patients. The Lundström (2005) and the Marcantonio (2001) studies reported no missing data, and all patients were included in their analyses.

Two studies evaluated delirium as a primary outcome (Marcantonio 2001; Lundström 2005). The primary outcome in the Landefeld (1995) study was the change from admission to discharge in the number of activities of daily living (ADL) that patients could perform independently.

The Marcantonio (2001) study evaluated delirium using the CAM diagnostic algorithm. The Marcantonio (2001) study also assessed individual symptoms of delirium using the DSI and severity of delirium was evaluated using the MDAS (scored 0-30, 30 best). In the Lundström (2005) study, delirium was diagnosed using the DSM-IV criteria. Delirium was also measured using a modified version of the Organic Brain Syndrome (OBS) scale, which incorporated the MMSE to assess disorientation, and the Katz ADL index to assess ADL. The Landefeld (1995) study only reported a mental status score based on the Mini-Mental State scale (using a score from 0-21, with higher scores indicating better cognitive function). This was considered to be a partially adequate method of measuring delirium.

Overall, the Lundström (2005) study was considered to be at higher risk of bias due to inadequate allocation concealment, and non-blinding of outcome assessors. The Landefeld (1995) study was at higher risk of bias because of non-blinding of outcome assessors, incomplete recruitment and the use of the MMSE for diagnosis of delirium. With the exception of the Landefeld (1995) study, the RCTs were relatively small and not highly powered.

10.20.2 Non-randomised studies

Five non-randomised studies were included in the review (Gustafson 1991; Harari 2007a; Inouye 1999; Wanich 1992; Wong 2005).

Three studies reported that all eligible patients were recruited consecutively to the study (Gustafson 1991; Harari 2007a; Wong 2005). The Inouye (1999) study stated that, of the 2434 patients meeting the inclusion criteria, 1265 (52%) were excluded because of inability to participate in interviews: because of a hospital stay of less than 48 hours (219); prior enrolment in their study (324), dementia (154), patient not available, etc. The 1265 excluded patients did not differ significantly from those included in terms of age, sex, risk of delirium, but a larger proportion were excluded from the control group than the intervention. The remaining patients had 250/1169 (21%) patients/family/physician who refused consent and an additional 67 who could not be matched. These unmatched patients were significantly older, had a higher risk of delirium at baseline, and were more likely to be admitted to a usual-care unit.

In the Wanich (1992) study, 117/354 (33%) patients/physicians refused consent.

The Inouye (1999) study was a non-randomised controlled study, and patients were allocated to groups by matching on age, sex, and baseline risk of delirium. The Wanich (1992) study was also a non-randomised controlled study in which patients from different wards were compared; it was not stated if the patients were matched.

The Gustafson (1991) study was a historical controlled trial in which a group of patients given the intervention in December 1986 to January 1988 were compared with a group of patients in the same hospital from March 1983 to June 1984.

The Harari (2007a) study was a historical controlled trial in which a group of patients given the intervention in August 2003 to February 2004 were compared with a group of patients in the same hospital from May to July 2003.

The Wong (2005) study was a historical controlled trial where baseline data were collected for 28 days on one group of patients, and further data were collected on another group of patients during the subsequent three months.

The Inouye (1999) study took account of possible confounders, by matching patients on the basis of age, sex and baseline risk of delirium; patients were included only if their risk of delirium was intermediate or high, as defined in the Inouye (1993) study. This Inouye (1993) study used a predictive model to define intermediate and high risk, based on risk factors of visual impairment, severe illness, cognitive impairment and a high ratio of blood urea nitrogen to creatinine. In order to appraise the accuracy of the matching on the basis of delirium risk, we need to assess the quality of the predictive model. We note that the prognostic factor review classified the Inouye (1993) study as low quality and that the predictive model did not include the full set of risk factors for delirium as identified in the risk factors review (section 7.2.1). Therefore, we can conclude that the possible confounders have not been completely accounted for in the matching process, although this may not be an important difference.

The method involved prospective individual matching of patients that had already been assigned to treatment groups; patients were admitted to one of three units (two control and one intervention) and matching was carried out using a computerised algorithm, based on logistic regression methods. The authors stated that randomisation of patients to intervention or usual care units was not feasible because of the large number of patients in all medical units at the time of the study; a pilot study found that beds in the intervention group were often unavailable. This pilot study does not appear to have been reported. The authors contend that their method of prospective matched pairing was chosen as an alternative to randomisation, but we note that the matching is only on the basis of known confounders whereas randomisation theoretically matches on known and unknown. There were no significant differences at baseline for age, sex, race, married, residence in a nursing home, education, APACHE II score, impairment in activities of daily living, MMSE score, patients with dementia, immobility, visual impairment, dehydration, comorbidities. However, the authors stated that contamination between groups was evident, because of the low rates of delirium in the control group, and because it was stated that intervention protocols were carried across to the usual care wards. This contamination would have underestimated the effect.

In the Harari (2007a) study, the patients in the intervention group were selected to be at-risk: those on the waiting list, aged 65 years and older, were sent a preoperative questionnaire and those with any risk factor (e.g. significant medical problems) were invited to the ‘proactive care of older people undergoing surgery (POPS)’ clinic. The control group was not selected in this way and patients were included regardless of case-mix. At baseline, there was a significant difference in renal impairment and hypertension), but the study used linear multiple regression to adjust for any baseline differences. We note that the percentages of people with hypertension were 80% and 52% in the intervention and control groups respectively (p=0.01); there were 22% and 4%.respectively with renal impairment (p=0.007). These are highly significant differences.

In the Wanich (1992) study, the intervention group had significantly more people with cardiac disease and cerebrovascular accidents and the control group had significantly more with neoplasm as the primary diagnosis. Adjustments were not made for the delirium outcome. The study also reported some contamination because some intervention techniques (e.g. medication management and physiotherapy) were also given to control patients.

The Wong (2005) study reported no significant differences in the age, sex, mental scores, Barthel indices, types of surgery or comorbidities between the baseline group and the post intervention group.

The Gustafson (1991) study reported no significant differences between groups in impaired vision, impaired hearing, dementia, depression, psychosis, many comorbidities, but significantly more people in the intervention group had cerebrovascular diseases and significantly more had urinary incontinence; the intervention group also received significantly fewer antiparkinsonian drugs, but significantly more of other drugs (e.g. penicillin); the control group also had more patients walking without walking aids before the fracture. The Gustafson (1991) study did not consider potential confounders in their analyses. Although these are important differences, it is not clear what would be their effects on delirium risk.

The historical comparison studies did not have blinded outcome assessors, nor did the Wanich (1992) study. However, the Inouye (1999) study reported that outcome assessors were blinded.

All the non-randomised studies, with the exception of the Harari (2007a) study, evaluated delirium as a primary outcome. The primary outcome in the Harari (2007a) study was hospital length of stay.

Two studies (Inouye 1999; Wong 2005) reported that delirium had been assessed using the CAM, and two studies (Gustafson 1991; Wanich 1992) diagnosed delirium using the DSM-III criteria. One study (Harari 1997a) assessed delirium as 'acute change in mental status postoperatively with improvements pre-discharge', but did not say how this was done. Therefore, the GDG down graded this study.

Five non-randomised studies reported no missing data and all patients were included in their analyses. In the Inouye (1999) study, 6 (1%) patients in the intervention group and 7 (2%) patients in the control group died during hospitalisation, but information on delirium was available for all patients. In the 6 month follow up study (Bogardus 2003), baseline data were available for 705/852 (83%) patients, 133 (16%) of whom had died. This study reported some additional missing data for some outcomes (for example, only 580 (68% of original sample) reported cognitive impairment).

Overall, none of the non-randomised studies were of high quality: the study by Inouye (1999) had the best study design, but large numbers of patients were not recruited and the matching of patients had limitations. The Bogardus (2003) study was considered at higher risk of bias for some outcomes because of missing data.

All of the other studies were considered to have a higher risk of bias:

- The Harari (2007a) study appeared to compare different types of patient, as well as not using a recognised method of assessing delirium and being a historical comparison.
- Two other studies had baseline differences (Gustafson 1991; Wanich 1992), but all the confounders in these studies appeared to disfavour the intervention group.
- The Wong (2005) study was considered at risk of bias because of its study design
- The Wanich (1992) study also reported some contamination
- In all studies, except the Inouye (1999), none of the outcome assessors were blinded.

10.21 Results

10.21.1 Multicomponent hospital care versus usual treatment

In summarising the results we have decided to indicate with one, two or three asterisks, studies which are considered to be at some, higher or much higher risk of bias respectively (i.e. moderate, low and very low quality studies, respectively). High quality studies have no asterisks. Where possible, we have separated the high quality (zero asterisk) and moderate quality studies (one asterisk) in the forest plots, or have outlined the forest plots in black.

Incidence of delirium

With the exception of the RCT by Landefeld** (1995), all studies evaluated the incidence of delirium. This outcome was evaluated differently between studies (e.g. cumulative incidence versus incidence at defined time point):

- the Gustafson** (1991) study reported acute confusional state in the postoperative period from 8 hours to 7 days and at 7 or more days
- the Harari*** (2007a) study reported outcomes measured during the hospitalisation period (mean 11.5 to 15.8 days)
- the Inouye* (1999) study appeared to report the rate of incidence of delirium up to 7 days and the number of patients were calculated from percentages

- the Lundström** (2005) study reported the incidence of delirium at 24 hours, 3 days and 7 days after admission. For the latter two days, the authors reported the data as the number of delirious patients on day 3 or 7 divided by the number with delirium on day 1. In our analyses, we have used the total number of patients in each group as the denominator

- the Marcantonio (2001) study reported the cumulative incidence during hospitalisation (mean about 3 days)

- the Wanich**(1992) study recorded the incidence of delirium at some time during their hospital stay (about 9 days), 38/48 within 24 h of admission

- the Wong** (2005) study recorded delirium in hospital (median stay 8-10 days)

Figure 10.9 (Appendix K) shows all studies separately for outcomes up to 7 days. Considering all the studies, we note that, generally, there was a significant effect of multicomponent interventions on the incidence of delirium. Considering only the reasonably reliable studies, Marcantonio (2001) and Inouye* (1999), each had a relative risk of about 0.66. In general these results were lacking in precision: the confidence interval was consistent with both a clinically important difference and no clinically important difference [see Grading evidence section 2.4.7 in the methodology chapter (chapter 2) for further information on imprecision].

Follow up

The six month follow-up study by Bogardus* (2003) (following the Inouye* 1999 study) found no significant difference between the groups (figure 10.10, Appendix K).

The confidence limits were consistent with significant harm and significant benefit, so the evidence quality was considered to be very low, on the grounds of being imprecise. Please refer to the Grading evidence section 2.4.7 in the methodology chapter (chapter 2) for further information on when evidence is considered imprecise.

Duration of delirium

One RCT reported on the mean number of days with delirium per episode of delirium (Marcantonio 2001). The results demonstrate that there was no difference in the mean duration of delirium per episode (not per person) between the treatment and control group; MD –0.20 days (95%CI –0.95, 0.55); figure 10.11, Appendix K. The results were considered to be precise for this outcome, although the study was small.

One non-randomised study reported on the number of patients with delirium for 7 days or more (Gustafson** 1991). There was no significant difference between groups (figure 10.12, Appendix K).

The non-randomised study by Inouye* (1999) reported that the total number of days of delirium amongst all patients in each group was significantly lower in the intervention group than in the usual-care group (105 versus 161 days, p=0.02).

Severity of delirium

One non-randomised study evaluated severity of delirium (Inouye* 1999), using an additive score for four symptoms (symptom fluctuation, inattention, disorganised thinking and an altered level of consciousness), ranging from 0 to 7 with higher scores indicating increased severity; the GDG were uncertain whether this was a validated scale, although it uses individual CAM items.

There was no difference in severity of delirium between the intervention and control groups (figure 10.13, Appendix K); MD 0.33 (95%CI 0.15 to 0.51); this is a precise result.

Length of hospital stay

Length of hospital stay was reported in three RCTs (Landefeld** 1995; Lundström** 2005; Marcantonio 2001), and five non-randomised studies (Gustafson** 1991; Harari*** 2007a; Inouye* 1999; Wanich**1992; Wong** 2005).

Three non-randomised trials reported the mean number of hospital days (Gustafson** 1991; Harari*** 2007a; Wanich**1992).

Five studies reported the mean length of stay (Gustafson** 1991; Harari*** 2007a; Lundström** 2005; Wanich**1992), but in each case, at least one of the groups had a skewed distribution.

The RCT by Landefeld** (1995) reported mean lengths of hospital stay of 7.3 and 8.3 days respectively for the intervention and control groups respectively, but standard deviations were not reported; the authors also reported that the median length of stay (6 days) was the same for both groups. We note that the Landefeld** (1995) study did not report the incidence of delirium.

The Lundström** (2005) study reported that patients in the treatment ward stayed in hospital for significantly fewer days than those in the control group; MD –4.05 (95% CI, -6.05, -1.95) (figure 10.14, Appendix K) . Due to a higher risk of bias, however, this result should be interpreted with caution.

With the exception of the Wanich** (1992) study, patients in the intervention group stayed in hospital for significantly fewer days than patients in the control group. In the Wanich** (1992) study there was no significant difference in hospital stay.

Four studies reported median length of stay:

- The Marcantonio (2001) RCT found no significant difference in length of hospital stay; both groups had a median stay of 5 days (with an interquartile range of 2); p = 0.95.

- The Inouye* (1999) study reported that the median length of stay was 7 days in the intervention group and 6.5 days in the control group; this was not a significant difference (p = 0.95).

- The Wong** (2005) study reported that the median length of stay was 10 days (2-44) in the intervention group and 8 days (3-41) in the control group; this was not a significant difference.

- The Harari*** (2007a) study reported a median length of stay of 10.0 days (range 4-26) and 14.5 (2-80) days for the intervention and control groups respectively (this was not a significant difference; p=0.058).

Cognitive impairment

The Inouye* (1999) study reported an adjusted orientation score (10 items on the MMSE) at reassessment (day 5 or at discharge if earlier); adjustment was for the patients' baseline score. We note that all patients received the cognitive impairment protocol once daily and those with an MMSE score below 20 or an orientation score below 8 received the protocol 3 times daily (advanced protocol); results were only reported for 253 of the original 852 patients (as two groups) – we assume this included the patients receiving the advanced protocol and their matched pairs in the control group. There were significantly more patients who had improved by 2 points on the MMSE at 5 days or at discharge : RR 1.51 (95% CI 1.05 to 2.17)(figure 10.15, Appendix K).

There was no significant difference in MMSE score in 580 patients (i.e. more than 20% missing data) at 6 months follow up in the Bogardus* (2003) study: adjusted mean difference –0.3 (95%CI -0.7 to 0.1) on a scale of 0–23. This study reported the MMSE score for all patients available, regardless of whether they had the advanced protocol.

One low quality RCT (Landefeld** 1995) reported no significant difference (p = 0.3) in MMSE scores (0 to 21) between the intervention (17.3) and control (17.7) groups for patients surviving to hospital discharge.

Number of patients discharged to new long-term care placement

One low quality RCT (Landefeld** 1995) reported that, of the patients admitted from private homes who survived to discharge, significantly fewer patients in the intervention group were discharged to new long-term care (figure 10.16, Appendix K); RR 0.64 (95% CI 0.46 to 0.90) which corresponds to a number needed to treat of 13 (95% CI 8 to 50), for a control group rate of 22%.

In addition, two studies (Marcantonio 2001; Wanich**1992) presented percentages of patients discharged to institutional settings (e.g. nursing home, rehab hospital); however, it was not clear how many of the patients were in long-term care settings at baseline.

In a non-randomised study (Wong** 2005), no significant difference in the number of patients discharged to higher level care was found between the intervention and control groups RR 0.96 (95% CI 0.45, 2.06); figure 10.16, Appendix K.

The Bogardus* (2003) study reported the number of patients with a new long-term placement at 6 months follow up of the Inouye* (1999) study. The denominators used were the number of patients in the original study. There was no significant difference between interventions RR 0.98 (95% CI 0.75 to 1.28); figure 10.16, Appendix K.

Mortality

Two low quality RCTs (Landefeld** 1995; Lundström** 2005) and four non-randomised studies reported on mortality (Harari*** 2007a; Inouye* 1999/Bogardus* 2003; Wanich**1992; Wong** 2005).

The Inouye* (1999) non-randomised study reported mortality during the hospitalisation period and the Bogardus* (2003) study reported mortality between hospital admission and 6 months follow up. In the latter case, the denominators used were the number of patients in the original study. There was no significant difference between interventions, but the confidence interval was consistent with significant benefit and significant harm.

The Lundström** (2005) study reported on mortality but only in patients with delirium. They found that mortality was less *in delirious patients* who received the intervention, than in delirious patients who received usual care (2/63 (3.2%) compared to 9/62 (14.5%), p=0.03).

In the Harari*** (2007a) study, the figures reflect the number of patients who died within 30 days of surgery. The Landefeld** (1995) also reported the number of deaths post discharge and up to 3 months and we used these data to calculate the number of deaths between admission and 3 months.

Overall none of the studies showed an effect on mortality, but often the CIs were wide and the results imprecise (figures 10.17 and 10.18, Appendix K).

Activities of daily living

Three non-randomised studies evaluated ADL (Inouye* 1999/Bogardus* 2003; Landefeld** 1995; Wanich**1992); figure 10.19. The Lundström** (2005) study also examined the patients using the Katz ADL scale but no results were reported.

The Inouye* (1999) study reported an adjusted Katz ADL score, on a scale of 0–14 (low scores indicate functional impairment), at reassessment (day 5 or at discharge if earlier); adjustment was for their baseline score. Although the study reported that standard deviations were given, this did not agree with the p value reported and it was assumed that the SDs were standard errors. Accordingly we calculated standard deviations. There was no significant difference between interventions (figure 10.19, Appendix K); MD 0.40 (95%CI -0.43, 1.23) on a scale of 0 to 14. There was no significant difference in the number whose immobility improved by 2 points but this result was imprecise (figure 10.19). We note that all patients had ambulation where possible and additional measures were provided when patients were non-ambulatory, Results were only reported for 194/852 patients.

In the Wanich** (1992) study a change in functional status was determined as an increase or decrease in two or more levels of function (e.g. Katz level C to E or C to A). By comparing the proportion of patients who were 'better', 'same' and 'worse', more patients in the intervention group had improved functional status and fewer had deteriorated in function compared to patients in the control group (p=0.02). The Wanich (1992) study also carried out a multiple logistic regression analysis to take into account baseline differences; the adjusted odds ratio was still significant; OR 3.29 (95%CI 1.26 to 8.17) (figure 10.20, Appendix K).

The Landefeld** (1995) study also reported on the change from admission to discharge in the number of basic activities performed independently (using the Katz index); the authors reported the number of patients with improved or much improved levels of function (figure 10.20, Appendix K) and the mean number of basic activities that could be performed at discharge (up to 5); this was 3.6 and 3.3 for the intervention and control groups respectively, which was of borderline significance (p = 0.05).

Post-discharge follow up

There was no significant difference in ADL score in 704 patients at 6 months follow up in the Bogardus* (2003) study: adjusted mean difference 0.1(95%CI –0.2 to 0.4) on a scale of 0–14. There was also no significant difference in the mean number of basic activities that could be performed in the 3 months after discharge in the Landefeld** (1995) study; this was 4.0 and 3.8 for the intervention and control groups respectively, (p = 0.3).

Severe falls

One study (Gustafson** 1991) reported the number of people with severe falls. The confidence interval was too wide [RR 0.80 (0.00 to 1.45) to determine if there was a difference between interventions (figure 10.21, Appendix K).

Infections

Urinary tract infections

Two studies (Gustafson** 1991; Harari*** 2007a) reported the number of patients with urinary infections). There was no significant difference between the intervention and control studies in the number of patients with urinary tract infections, although the results were imprecise in the Gustafson** (1991) study (RR 1.37 (95% CI 0.88 to 2.12)) and very imprecise in the Harari*** (2007a) study (RR 0.44 (95% CI 0.15 to 1.36)) (figure 10.22, Appendix K).

Wound infection

One study (Harari*** 2007a) reported the number of patients with wound infections. There was a clinically significant difference (RR 0.17 (95% CI 0.04 to 0.71) but there was imprecision in this small study (figure 10.23, Appendix K).

Pressure ulcers

Two non-randomised studies (Gustafson** 1991; Harari*** 2007a) reported the number of people with pressure ulcers. There was a significant difference between interventions in both studies (Gustafson** 1991: RR 0.31 (95% CI 0.10 to 0.91); Harari*** 2007a: RR 0.20 (0.05 to 0.87), but the results are imprecise (figure 10.24, Appendix K)

Sensory impairment

Visual impairment

The Inouye* (1999) study reported the number of patients with early vision correction at reassessment (day 5 or at discharge if earlier). There was no significant difference between interventions (figure 10.25, Appendix K); RR 1.34 (95% CI 0.79 to 2.28), but the results are imprecise. We note that only patients who had a visual acuity of less than 20/70 on binocular near vision testing received the vision protocol; results were only reported for 119/852 patients.

Hearing impairment

The Inouye* (1999) study reported an adjusted Whisper test score at reassessment (day 5 or at discharge if earlier); adjustment was for the patients' baseline score. Although the study reported that standard deviations were given, this did not agree with the p value reported and it was assumed that the SDs were standard errors. Accordingly we recalculated standard deviations. There was no significant difference between interventions MD 0.80 (95%CI -0.19, 1.79) on a scale of 0 to 12 (good hearing) (figure 10.26, Appendix K). There was no significant difference in the number (RR 1.28 (95% CI 0.95 to 1.72) whose score improved by 1 point (figure 10.27, Appendix K). We note that only patients who had a Whisper test score below 7 received the protocol once daily; results were only reported for 218/852 patients.

Dehydration

Two non-randomised studies reported on dehydration (Harari*** 2007; Inouye* 1999).

The Inouye* (1999) study reported the number of patients assessed to be improved by 5 points for the adjusted ratio of blood urea nitrogen to creatinine at reassessment; adjustment was for the patients' baseline score. There was no significant difference in the number who were assessed to be improved RR 1.16 (95% CI 0.94 to 1.43) although the results are imprecise (figure 10.28, Appendix K). We note that only patients who had a ratio of blood urea nitrogen to creatinine of at least 18 received the protocol; results were only reported for 494/852 patients.

The Harari*** (2007a) study reported the number of patients with dehydration. The CI was very wide and consistent with both important benefits and important harms RR 0.67 (95% CI 0.20 to 2.23) (figure 10.29, Appendix K).

Urinary incontinence

Two studies investigated urinary incontinence (Gustafson** 1991; Bogardus* 2003; Inouye* 1999). There was no significant difference between the intervention and control studies in the number of patients with urinary infections in the Gustafson** 1991 study (RR 0.62 (95% CI 0.35 to 1.11), but the 6 months follow up of the Inouye* (1999) study showed a significant difference in the number of people with incontinence compared with the usual care group (RR 0.80 (95% CI 0.65 to 0.99). Both studies showed imprecision (figure 10.30, Appendix K).

Adherence

One study (Inouye* 1999) reported the overall rate of adherence to all protocols (87%) and the rate of adherence to individual protocols: orientation 96%; vision 92%; hearing 92%; therapeutic activities 86%; early mobilisation 84%; volume repletion 81% and non-pharmacological sleep 71%. No adverse effects were associated with the intervention protocols. The Marcantonio (2001) study reported an overall adherence to recommendations of 77%, and the Wong** (2005) study reported 90%.

10.21.2 Overall summary

Summary of results for multicomponent prevention of delirium in hospital setting are reported in table 10.4.

Table 10.4: summary of results - non pharmacological multicomponent intervention for the prevention of delirium.

Outcome	**Education and reorganisation of nursing & medical care** (Lundstrom** 2005)	**Elder Life Program** (Inouye* 1999)	**Education and multicomponent** (Wanich** 1992)	**'Acute Care for Elders' programmes** (Landfeld** 1995)	**Proactive care of older people undergoing surgery** (Harari*** 2007)	**Quality improvement programme** (Wong** 2005)	**Proactive geriatrics consultation** (Marcantonio 2001)	**Geriatric-anaesthesiologic intervention** (Gustafson** 1991)
Incidence of delirium	RR 0.51 (95% CI 0.31 to 0.86)	RR 0.66 (95% CI 0.46 to 0.95); at 6 month follow-up:1.25 (0.55to 2.84)	RR 0.88 (95% CI 0.53 to 1.45)		RR 0.30 (95% CI 0.09 to 1.03)	RR 0.35 (95% CI 0.16 to 0.78)	RR 0.65 (95% CI 0.42 to 1.00)	RR 0.78 (95% CI 0.60 to 1.00)
Duration of delirium							MD -0.20 (-0.95 to 0.55)	RR 0.73 (95% CI 0.50 to 1.07)
Severity of delirium		MD 0.33 (95% CI 0.15 to 0.51)						
Outcome	**Education and reorganisation of nursing & medical care** (Lundstrom** 2005)	**Elder Life Program** (Inouye* 1999)	**Education and multicomponent** (Wanich** 1992)	**'Acute Care for Elders' programmes** (Landfeld** 1995)	**Proactive care of older people undergoing surgery** (Harari*** 2007)	**Quality improvement programme** (Wong** 2005)	**Proactive geriatrics consultation** (Marcantonio 2001)	**Geriatric-anaesthesiologic intervention** (Gustafson** 1991)
Length of hospital stay	MD-4.00 (95% CI-6.05 to -1.95)		MD -1.20 (95% CI -3.67 to 1.27)		MD-4.30 (95% CI -8.08 to -0.52)			MD -5.80 (95% CI -8.85 to -2.75)
Cognitive Impairment		RR 1.51(95% CI 1.05 to 2.17)						
Discharge to new LTC	RR 1.05(95% CI 0.93 to 1.18)	RR 0.98 (95% CI 0.75	2.04 (0.67 to 6.21)	0.64 (0.45 to 0.90)		0.96 (0.45 to 2.06)		

		to 1.28)						
Outcome	**Education and reorganisation of nursing & medical care** (Lundstrom** 2005)	**Elder Life Program** (Inouye* 1999)	**Education and multicomponent** (Wanich** 1992)	**'Acute Care for Elders' programmes** (Landfeld** 1995)	**Proactive care of older people undergoing surgery** (Harari*** 2007)	**Quality improvement programme** (Wong** 2005)	**Proactive geriatrics consultation** (Marcantonio 2001)	**Geriatric-anaesthesiologic intervention** (Gustafson** 1991)
Mortality		RR 0.86 (95% CI 0.29 to 2.53) At 6months: RR 1.22 (95% CI 0.89 to 1.67)	RR1.63 (95% CI 0.58 to 4.54)	RR0.99 (95% CI 0.57 to 1.71)	RR 0.33 (95% CI 0.01 to 8.01)	RR 0.59 (95% CI 0.10 to 3.35)		
ADL		0.47 (0.19 to 1.19); adjusted ADL: MD 0.40 (-0.43 to 1.23)	2.16 (1.23 to 3.80)	2.17 (1.07 to 4.42)				
Outcome	**Education and reorganisation of nursing & medical care** (Lundstrom** 2005)	**Elder Life Program** (Inouye* 1999)	**Education and multicomponent** (Wanich** 1992)	**'Acute Care for Elders' programmes** (Landfeld** 1995)	**Proactive care of older people undergoing surgery** (Harari*** 2007)	**Quality improvement programme** (Wong** 2005)	**Proactive geriatrics consultation** (Marcantonio 2001)	**Geriatric-anaesthesiologic intervention** (Gustafson** 1991)
Post-discharge follow up		Adjusted : MD 0.1 (0.2 to 0.4)		Mean number of basic activities 3				

				months after discharge: p=0.3				
Severe falls								0.08 (0.00 to 1.45)
UTI					0.44 (0.15 to 1.36)			1.37 (0.88 to 2.12)
Wound infection					0.17 (0.04 to 0.71)			
Pressure ulcers					0.20 (0.05 to 0.87)			0.31 (0.10 to 0.91)
Outcome	**Education and reorganisation of nursing & medical care** (Lundstrom** 2005)	**Elder Life Program** (Inouye* 1999)	**Education and multicomponent** (Wanich** 1992)	**'Acute Care for Elders' programmes** (Landfeld** 1995)	**Proactive care of older people undergoing surgery** (Harari*** 2007)	**Quality improvement programme** (Wong** 2005)	**Proactive geriatrics consultation** (Marcantonio 2001)	**Geriatric-anaesthesiologic intervention** (Gustafson** 1991)
Visual impairment		RR 1.34 (95% CI 0.79 to 2.28)						
Hearing impairment		MD 0.80 (-0.19 to 1.79) Improvement by one point: RR 1.28 (95% CI 0.95 to 1.72)						
Outcome	**Education and reorganisation**	**Elder Life Program**	**Education and multicomponent**	**'Acute Care for Elders'**	**Proactive care of older**	**Quality improvement programme**	**Proactive geriatrics consultation**	**Geriatric-anaesthesiologic**

	of nursing & medical care (Lundstrom** 2005)	(Inouye* 1999)	(Wanich** 1992)	**programmes** (Landfeld** 1995)	**people undergoing surgery** (Harari*** 2007)	(Wong** 2005)	(Marcantonio 2001)	**intervention** (Gustafson** 1991)
Dehydration		Improvement in dehydration: RR 1.16 (0.94 to 1.43)			Number of patients with dehydration: RR 0.67 (95% CI 0.20 to 2.23)			
Urinary incontinence		RR 0.62 (95% CI 0.35 to 1.11)						RR 0.80 (95% CI 0.65 to 0.99)
Adherence		Overall rate of adherence to all protocols: 87%				Overall adherence to recommendations: 90%	Overall adherence to recommendations: 77%	

10.22 Health economic evidence

10.22.1 Multicomponent interventions for the prevention of delirium in a hospital setting

One economic evaluation study was included as evidence (Rizzo 2001). This was a non-randomised study of 70 year old patients with no evidence of delirium but who had intermediate or high risk of delirium. It was conducted in the USA in 2001 with the following objectives to:

- determine the impact of the multicomponent intervention strategy on total hospital costs, average daily costs, and length of stay,
- estimate the impact of the multicomponent intervention on specific hospital cost components,
- describe the intervention costs associated with the intervention strategy, and
- combine the results of cost and effectiveness analyses to assess the cost-effectiveness of the intervention strategy.

Patients in the intervention group were those who met the inclusion criteria of being 70 years and older with no evidence of delirium but had intermediate or high risk of delirium. Control patients were prospectively selected and matched on age, gender, and baseline delirium risk. The intervention group received the multicomponent intervention (Hospital Elder Life Program) strategy which consisted of interventions targeted toward six delirium risk factors (cognitive impairment, sleep deprivation, immobility, visual impairment, hearing impairment and dehydration). The core interventions included orienting communication, therapeutic activities, sleep enhancement strategies, exercise and mobilisation, provision of vision and hearing aids, and oral volume repletion for dehydration. Others included geriatric nursing assessment and interdisciplinary rounds. The control arm did receive usual hospital care.

The cost of the intervention was based on personnel and equipment costs during the three year study period. The total personnel and equipment costs over this period were $252,885 and $257,385 respectively. The non-intervention costs in the intervention and usual care groups were reported as $6,484 and $7,300 respectively. The additional cost of the intervention was $592 per patient (standard error, se=21). Unit cost and resources use were reported and the perspective of the analysis was third party (hospital healthcare provider). The multicomponent intervention was estimated to result in cost savings (excluding intervention costs) of $831 for intermediate risk patients after multivariate adjustment for confounding variables but there was no significant difference for the high risk group. In the intermediate delirium risk patients the net cost saving attributable to the intervention was $99 if intervention costs were included. This was statistically insignificant after multivariate adjustment. The intervention had a statistically significant cost increase of $1,308 in high risk patients.

The overall incidence of delirium was 9.9% and 15.0% in the intervention and control groups respectively. The incidence of delirium in the intermediate risk group was 6.5% with intervention and 11.7% without intervention. In the high risk group, it was 18.5% and 23.5% respectively. Incidence of delirium was based on CAM, MMSE and digital span test. A mortality rate of 1% and 2% were reported in the respective groups. Costs were not assessed from a UK NHS and PSS perspective. The measure of health benefit from the intervention was not in QALY units. The results of this study were judged to be not applicable to the guideline population.

10.23 Clinical evidence statements

There is low quality evidence to show the following results for a multicomponent intervention based on targeting 6 modifiable risk factors (cognitive impairment, sleep deprivation, immobility, vision impairment, hearing impairment, dehydration), with training (Inouye 1999) in patients at high or intermediate risk of delirium a:

- significant reduction in the incidence of delirium; RR 0.66 (95% CI 0.46 to 0.95)
- significant reduction in the total number of days of delirium amongst all patients in the group (105 versus 161 days)
- significant difference in the number with urinary incontinence after 6 months follow up; RR 0.80 (95%CI 0.65 to 0.99).

No significant difference in the:

- incidence of delirium after 6 months follow up; the evidence was very low quality for this outcome
- MMSE score after 6 months follow up
- delirium severity
- median length of stay in hospital
- number of patients with a new long-term care placement
- number of patients who died, either during the hospitalisation period or in the time between hospital admission and 6 months follow up; the evidence for hospitalised patients was very low quality.

There is low quality evidence to show the following results for a multicomponent intervention based on targeting 6 modifiable risk factors with training (Inouye

1999) in subgroups of patients who were targeted to receive the part of the multicomponent intervention appropriate to that outcome (the proportion receiving the targeted component is given in brackets)

- A significant increase in the number of patients with an improvement of 2 points on their MMSE score after 5 days or at discharge if earlier (253/852).
    - No significant difference in the number of patients:
    - with an improvement in activities of daily living (194/852)
    - with early vision correction at reassessment (day 5 or at discharge if earlier) (119/852)
    - whose hearing improved at reassessment (day 5 or at discharge if earlier) (218/852)
    - whose dehydration improved at reassessment (day 5 or at discharge if earlier) (494/852).

There is moderate quality evidence to show the following results in patients undergoing surgery for hip fracture (i.e. higher risk), and receiving a multicomponent intervention based on targeting 7 modifiable risk factors (orientation, dehydration, sensory impairment, immobility, environmental modifications and medication management) following consultation with a geriatrician preoperatively (Marcantonio 2001) showed the following results:

- A borderline significant reduction in the incidence of delirium; RR 0.65 (95%CI 0.42 to 1.00)
- No significant difference in the:
    - mean duration of delirium per episode; this is an indirect outcome measure
    - median length of stay in hospital
    - number of patients discharged to long-term care (it was unclear if this was a new placement).

There was very low quality evidence for the effectiveness of an intervention consisting of an education programme for staff and reorganisation of nursing and medical care, such that the patients received patient centred care, rather than task allocated care. Results for this study (Lundström 2005) showed:

- A significant reduction in the:

- incidence of delirium; RR 0.51 (95%CI 0.31 to 0.86); this was considered low quality evidence rather than very low because the outcome assessors were blinded for the delirium of diagnosis
- mean length of stay in hospital, although the data were skewed.

The remaining evidence is from studies with a low quality study design (Gustafson 1991, Harari 2007a, Wanich 1992, Wong 2005) or from a low quality RCT that did not record the incidence of delirium as an outcome measure (Landefeld 1995).

For the outcome, incidence of delirium, there is very low quality evidence to suggest that the following interventions may have potential to reduce the incidence of delirium in hospital patients:

- Multidisciplinary team, pre- and post-operative assessment and targeting of identified issues including pain management, early mobilisation, nutrition, and early detection and treatment of medical complications (Harari 2007a). There is much uncertainty around this result
- Geriatric-anaesthesiologic intervention programme, including pre- and postoperative assessment by specialist in geriatric and internal medicine (Gustafson 1991)
- Plan-do-study-act programme, including staff education and geriatric team assessments to address 12 modifiable risk factors (Wong 2005).

There is very low quality evidence to suggest that the following intervention did not have a significant effect on the incidence of delirium: education of staff and assessment by geriatricians to address 6 modifiable risk factors (Wanich 1992).

10.24 Health economic evidence statements

The results of the economic model (chapter 16) showed the following:

- The use of two multicomponent targeted interventions was cost effective in:
    - elderly patients at intermediate or high risk of delirium and who were admitted to the general medicine service.
    - elderly patients who were admitted emergently for surgical repair of hip fracture.

These findings were robust as the interventions remained cost-effective after a series of sensitivity analyses were conducted.

10.25 From evidence to recommendations

10.25.1 Interventions to prevent delirium (recommendations 1.3.2, 1.3.3 and 1.3.3.1-1.3.3.10)

Recommendations 1.3.2 and1.3.3 derive from high (Marcantonio 2001), moderate (Inouye 1999), and low quality evidence in other studies from the multicomponent prevention review for patients in hospital (primary evidence source). This is supported by mixed quality evidence from the non-pharmacological risk factors review, low quality evidence from the hydration review, low to moderate quality evidence from the pharmacological risk factors review and GDG consensus. The latter was also informed by three other NICE guidelines ['Nutrition support in adults' (NICE clinical guideline 32), 'Infection control' (NICE clinical guideline 2; this guideline is currently being updated) and 'Parkinson's Disease' (NICE clinical guideline 35)]

Economic evidence for the hospital setting was obtained by modelling the preventative pathway and was informed by both the evidence from the multicomponent prevention and consequences of delirium reviews. It was also informed by evidence on cost, quality of life and baseline risks.

There was no clinical or cost-effectiveness evidence for the long-term care population. Recommendations for this setting were based on indirect evidence from the hospital population.

10.25.2 GDG considerations: multicomponent interventions in a hospital setting for the prevention of delirium

The evidence from two studies was of moderate and high quality (Inouye 1999 and Marcantonio 2001). Each of the multicomponent interventions (and not each study) were incorporated into the economic model (using the same risk profiles as those described in the studies) and was found to be cost effective. There was a degree of uncertainty around the cost-effectiveness estimates, but this uncertainty was not judged by the GDG to be sufficient to affect the general conclusion.

The GDG discussed whether the preventative intervention should be given to all patients, or only to those at risk of delirium, or whether to carry out sensitivity analyses to determine separately the cost effectiveness for intermediate and high risk groups. They concluded that the recommendation should be restricted to patients who are at-risk of delirium (corresponding to the intermediate and high risk groups in the Inouye (1993) study), but that healthcare professionals should be encouraged to give the intervention to all patients in that category. They defined the at-risk group according to the risk factors review (see section 7.8).

The GDG recognised that the initial stage of the multicomponent intervention was assessment of the person's needs, and a recommendation was made for multicomponent intervention that is tailored to the individual's needs. Both of the higher quality intervention studies (Inouye 1999 and Marcantonio 2001) included this initial assessment stage, and the GDG agreed this was very important. The GDG also concurred with the evidence from the Marcantonio (2001) study, that this assessment should be made within 24 hours of admission.

In line with evidence from the Inouye (1999) study, the GDG agreed that a multidisciplinary team should carry out the multicomponent intervention, and considered it important that the healthcare team should be trained and competent in carrying out these tasks.

The GDG discussed whether to recommend one or both of the multicomponent intervention 'packages' (described by the two reviewed studies) or whether to produce a more general recommendation that selected individual elements from each package, together with evidence from the other reviews.

The GDG concluded that the latter course of action should be taken and that the two packages could be used to make a broad recommendation since the studies showed that when risk factors were addressed by providing better quality care, outcomes were improved. Hence the studies were deemed by the GDG to be 'proof of concept' studies.

The GDG discussed which clinical factors should be addressed by the multicomponent interventions. The agreed list was closely based on the two multicomponent prevention packages that were modelled, supplemented by the additional clinical evidence and GDG clinical experience. Each factor that was included and the evidence for them is listed below:

- Cognitive impairment/disorientation – evidence from the Inouye (1999) study, the non-pharmacological risk factors review and from GDG expertise. The GDG recognised that the evidence from the Inouye (1999) study was for cognitive impairment which entailed both reorientation and therapeutic activities. In addition to cognitive impairment, the GDG felt it was important to address disorientation, because this is a specific manifestation of people who have underlying cognitive impairment.

- Dehydration / constipation – evidence from the Inouye (1999) and Marcantonio (2001) studies, from the hydration review and from GDG expertise. See also section 10.10 for the GDG rationale relating to hydration.

- Hypoxia- evidence from the Marcantonio (2001) study and GDG expertise.

- Infection – evidence from the Marcantonio (2001) study, the non-pharmacological risk factors review and GDG expertise; cross reference to the NICE Infection Control guideline. For catheterisation, the evidence came from the Marcantonio (2001) and Inouye (1999) studies, the non-pharmacological risk factors review, and GDG clinical expertise

- Limited mobility or immobility – evidence from the Inouye (1999) and Marcantonio (2001) studies and GDG expertise.

- Pain – evidence from the Marcantonio (2001) study, indirect evidence from the pharmacological risk factors review and GDG expertise. The GDG emphasised that both verbal and non verbal signs of pain should be assessed, particularly in patients with dementia or learning difficulties.

- Polypharmacy effects - evidence from the Marcantonio (2001) study, from both the pharmacological and non-pharmacological risk factors reviews and GDG expertise. The GDG advised recommending a drug review that addressed the type of drugs as well as the number; the GDG also supported the principle that if clinicians add a new long-term drug, another should be taken away.

- Poor nutrition - some evidence from the Marcantonio (2001) study and from lower quality multicomponent prevention studies, and GDG expertise; cross reference to the NICE nutrition guideline

- Sensory impairment – evidence from the Inouye (1999) and Marcantonio (1999) studies, from the non-pharmacological risk factors review for visual impairment and GDG expertise.

- Sleep disturbance – evidence from the Inouye (1999) study and GDG clinical expertise; evidence from the pharmacological risk factors review; cross reference to the NICE Parkinson's disease guideline. Although the GDG considered it important that patients slept well in hospital, they decided to exclude the use of sleep enhancers (which was part of the Inouye (1999) study intervention) because low quality evidence from the pharmacological risk factors review suggested that lorazepam may also cause delirium

10.25.3 GDG considerations: multicomponent interventions in the long-term care setting for the prevention of delirium

There was no evidence for multicomponent preventative interventions in a long-term care setting, and very limited evidence for the consequences of delirium. Clinical effectiveness was therefore extrapolated from the hospital setting and GDG experience. Health economic modelling was not carried out because there was a lack of data for this setting and a large number of assumptions would have had to be made by the GDG, leading to serious uncertainty in outcomes. GDG consensus was that a multicomponent intervention for long-term care could have large potential cost-savings, was unlikely to do any harm to patients, and could probably be fairly easily accommodated within current care without incurring high costs. The GDG decided to recommend that the tailored multicomponent intervention package should also be applied in the care setting, and that further research should be carried out. This led to writing a research recommendation (see below and Appendix H). The GDG considered it important that the care staff concerned should be trained and competent in carrying out the tasks in the multicomponent intervention. In the long-term care setting 'multidisciplinary team' should be interpreted as appropriate.

Future research recommendation:

For patients in long-term care, is a multicomponent non-pharmacological intervention more clinically and cost effective than usual care in preventing the development of delirium?

The GDG noted that some of the low quality multicomponent prevention studies examined the effectiveness of an educational intervention for staff. The GDG felt that this showed some potential, not least in the prevention of delirium resulting from increased staff awareness and this is reflected in a research recommendation (see below and Appendix H).

Future research recommendation:

Does an education programme for staff reduce the incidence of delirium and improve the recording of delirium for patients in hospital, compared with an education leaflet or usual care?

The GDG also wished to know what was the cost to the NHS of implementing a multicomponent prevention intervention, compared to the care that is currently given to people in hospital and long-term care. They therefore proposed a receommendation for future research:

Future research recommendation:

What is the resource use and cost of implementing a muticomponent prevention intervention in hospital or long term care settings as compared to usual care?

10.26 Recommendations

Give a tailored multicomponent intervention package:

- Within 24 hours of admission, assess people at risk for clinical factors contributing to delirium.

- Based on the results of this assessment, provide a multicomponent intervention tailored to the person's individual needs and care setting as described in recommendations 1.3.3.1-1.3.3.10. **[1.3.2]**

The tailored multicomponent intervention package should be delivered by a multidisciplinary team trained and competent in delirium prevention. **[1.3.3]**

[1.3.3.1] Address cognitive impairment and/or disorientation by:

- providing appropriate lighting and clear signage; a clock (consider providing a 24-hour clock in critical care) and a calendar should also be easily visible to the person at risk

- talking to the person to reorientate them by explaining where they are, who they are, and what your role is

- introducing cognitively stimulating activities (for example, reminiscence)

- facilitating regular visits from family and friends.

[1.3.3.2] Address dehydration and/or constipation by:

- ensuring adequate fluid intake to prevent dehydration by encouraging the person to drink – consider offering subcutaneous or intravenous fluids if necessary

- taking advice if necessary when managing fluid balance in people with comorbidities (for example, heart failure or chronic kidney disease).

[1.3.3.3] Assess for hypoxia and optimise oxygen saturation if necessary, as clinically appropriate.

[1.3.3.4] Address infection by:

- looking for and treating infection

- avoiding unnecessary catheterisation
- implementing infection control procedures in line with 'Infection control' (NICE clinical guideline 2).

[1.3.3.5] Address immobility or limited mobility through the following actions:

- Encourage people to:
    - mobilise soon after surgery
    - walk (provide appropriate walking aids if needed – these should be accessible at all times).
- Encourage all people, including those unable to walk, to carry out active range-of-motion exercises.

[1.3.3.6] Address pain by:

- assessing for pain
- looking for non-verbal signs of pain, particularly in those with communication difficulties (for example, people with learning difficulties or dementia, or people on a ventilator or who have a tracheostomy)
- starting and reviewing appropriate pain management in any person in whom pain is identified or suspected.

[1.3.3.7] Carry out a medication review for people taking multiple drugs, taking into account both the type and number of medications.

[1.3.3.8] Address poor nutrition by:

- following the advice given on nutrition in 'Nutrition support in adults' (NICE clinical guideline 32)
- if people have dentures, ensuring they fit properly.

[1.3.3.9] Address sensory impairment by:

- resolving any reversible cause of the impairment, such as impacted ear wax
- ensuring hearing and visual aids are available to and used by people who need them, and that they are in good working order.

[1.3.3.10] Promote good sleep patterns and sleep hygiene[13] by:

- avoiding nursing or medical procedures during sleeping hours, if possible
- scheduling medication rounds to avoid disturbing sleep
- reducing noise to a minimum during sleep periods.

[13] For more information on good sleep hygiene, see ‘Parkinson’s disease’ (NICE clinical guideline 35)

11 Prevention of delirium: pharmacological

11.1 Clinical introduction

The serious nature of delirium and its consequences makes all methods of prevention important to establish. Pharmacological agents are a recognised cause of delirium and so the use of these agents for prevention needs to be approached cautiously. Antipsychotic, benzodiazepines, acetylcholinesterase inhibitor classes of drugs in particular, and products that influence the immune system, may prove useful, based on early evidence from small studies, or from a theoretical perspective.

People at risk of delirium are already vulnerable to the adverse effects of pharmacological products. It will be essential to establish the efficacy and risks of preventative drug treatment from well conducted clinical trials before they might be considered for routine use in clinical practice.

11 A) Prevention in hospital

CLINICAL QUESTION: What are the most clinical and cost effective and safe pharmacological interventions for the prevention of delirium in people in hospital?

11.2 Description of studies

Details of included and excluded papers together with study design are reported in table 11.1.

Table 11.1: study inclusion, exclusion and design

Papers	Comments	Study
N= 10 evaluated for inclusion	2 Cochrane reviews were identified and are updated within this review	Lomergan 2007; Siddiqi 2007
N= 2 excluded	Reasons for exclusion are reported in Appendix G.	
N= 1 identified in update searches	1 RCT included	Gamberini 2009
N= 6 included	Study designs 5 RCTs	Aizawa 2002; Kalilsvaart 2005; Kaneko 1999; Liptzin 2005; Prakanrattana 2007

Two Cochrane Reviews were identified (Lonergan 2007; Siddiqi 2007) and updated. The Lonergan (2007) review examined the effectiveness of cholinesterase inhibitors in one study (Liptzin 2005) and the Siddiqi (2007) review examined both pharmacological (Aizawa 2002; Berggren 1987; Diaz 2001; Kalisvaart 2005; Liptzin 2005) and non-pharmacological (Marcantonio 2001) interventions for the prevention of delirium. Studies which did not meet our search criteria (Berggren 1987) or examined interventions not licensed for use in the UK (Diaz 2001) were not inlcuded. One study reporting non pharmacological intervention (Marcantonio 2001) has been reported in Chapter 10B (multicomponent prevention). This evidence review also includes outcomes not reported within the Cochrane reviews and has been updated to include papers published up to 2009.

11.2.1 Study Design

None of the studies were conducted in the UK. Information on study size, geographical location and funding are described in table 11.2.

Table 11.2: study characteristics

Study	Size (N)	Geographical location	Funding
Aizawa 2002	42	Japan	Not Stated
Gamberini 2009	120	Switzerland	Pharmaceutical industry
Kalisvaart 2005	430	The Netherlands	No Funding
Kaneko 1999	80	Japan	Not Stated
Liptzin 2005	90	USA	Pharmaceutical industry
Prakanrattana 2007	129	Thailand	Hospital research grant

Study duration was reported in four studies (Aizawa 2002: 7 days; Gamberini 2009: 6 days postoperatively; Kalisvaart 2005: varied to a maximum of six days depending on the onset of delirium; Liptzin 2005: 28 days).

11.2.2 Population

The age range across the studies was 51 years to 90 years. All studies included men and women. The patients' ethnicity was described as being 95% white and 5% other in one study (Liptzin 2005) and was not reported in the remaining studies.

All of the studies were conducted in hospital settings in patients undergoing surgery. The type of surgery included resection for gastric or colorectal cancer (Aizawa 2002); hip surgery for acute fractures or hip replacements (Kalisvaart 2005); gastrointestinal surgery (Kaneko 1999); total joint replacement surgery of the knee or hip (Liptizin 2005); cardiac surgery with cardiopulmonary bypass (Prakanrattana 2007), cardiac surgery (Gamberini 2009). The Kaneko (1999) study reported that all patients were admitted to an ICU before the scheduled surgery.

Cognitive status was not reported in two studies (Aizawa 2002; Prakanrattana 2007), one study (Liptzin 2005) reported that at baseline patients did not have dementia, and one study (Gamberini 2009) reported that patients with an MMSE score of less than 15 were excluded. Three studies reported that the method used to assess dementia was the Mini Mental State Examination (MMSE) (Gamberini 2009; Kalisvaart 2005; Liptzin 2005). The reported MMSE scores indicated that at least some of the patients had dementia. One study did not report the method used for the assessment of dementia (Kaneko (1999).

One study reported the risk of postoperative delirium (Kalisvaart 2005). In this study, 84% of the patients had an intermediate risk for postoperative delirium and 16% had a high risk for postoperative delirium (as based on four predictive risk factors not specifically described); low risk patients were excluded. Patients with delirium at hospital admission were excluded from the study.

The Kalisvaart (2005) study also described their patients as having light dehydration.

11.2.3 Interventions

Acetylcholinesterase

One study (Liptzin 2005) investigated the acetylcholinesterase inhibitor, donepezil.

- 5–10 mg donepezil per day.

One study (Gamberini 2009) investigated the acetylcholinesterase inhibitor, rivastigmine

- 1.5 mg oral rivastigmine three times per day every 8 hours, starting on the evening preceding surgery and continuing until the sixth postoperative day; each patient received 22 doses in total.

11.2.3.1 Atypical antipsychotics

One study (Prakanrattana 2007) investigated the atypical antipsychotic, risperidone.

- 1 mg (orally disintegrating tablet) sublingually as a one-off dose when patients started to wake up in the ICU.

11.2.3.2 Typical antipsychotics

Two studies (Kalisvaart 2005; Kaneko 1999) investigated the typical antipsychotic drug haloperidol. The interventions included:

- 0.5 mg haloperidol tablet three times per day, starting on hospital admission and continued until 3 days after surgery; a maximum delay from admission of 72 hours was permitted before surgery (Kalisvaart 2005)

- 5 mg intravenous haloperidol once per day, starting on the first postoperative day (Kaneko 1999)

Benzodiazepines

One study (Aizawa 2002) investigated the use of a 'Delirium Free Protocol (DFP)' which was designed to address the risk factor of insomnia. The DFP included:

- a combination of two benzodiazepines with pethidine: (diazepam 0.1 mg/kg per day intramuscularly given at 20.00h and a drip infusion of flunitrazepam 0.04 mg/kg) and pethidine 1 mg/kg (both given from 20.00 to 04.00h), for the first 3 days postoperatively, starting on the day of the operation.
- The GDG expressed concern that the method of delivery of the drug (IM diazepam), and the addition of pethidine made the effect of benzodiazepines unclear, the study was addressing symptoms of improving insomnia, which in turn is a risk factor for delirium; this study was therefore not considered further.

11.2.4 Comparisons

The following comparisons were carried out:

Acetylcholinesterase inhibitors

- Donepezil versus placebo (Liptzin 2005)
  - The intervention was given for 14 days preoperatively and a further 14 days postoperatively; patients were not admitted to hospital until the day before surgery.
  - The control group received placebo once a day at breakfast, and again, where symptoms of delirium were experienced, the placebo dose was doubled.
- Rivastigmine versus placebo (Gamberini 2009)

- The intervention was given the evening before surgery, three times per day every 8 hours thereafter until the evening of the sixth postoperative day.

- The control group was administered the placebo (liquid identical to rivastigmine solution) following the same dosing scheme.

- If postoperative delirium occurred, patients received haloperidol (starting with 0.5 mg every 6 to 8h) and lorazepam (1 mg per day)

11.2.4.1 Atypical antipsychotics

- Risperidone (orally disintegrating tablet) versus placebo (an antiseptic strip applied sublingually). The interventions were a one-off dose. (Prakanrattana 2007)

11.2.4.2 Typical antipsychotics

- Haloperidol versus placebo
    - 0.5 mg haloperidol tablet three times per day, up to 6 days pre and postoperatively (Kalisvaart 2005)
        - all patients received a proactive geriatric consultation (geriatric medical attention; enhancement of orientation and cognition; sensory and mobility improving advice; attention to pain and sleeping problems; extra attention to food and fluid intake; patient, family and nursing staff education). This study also gave the patients haloperidol and/or lorazepam 3 times a day if postoperative delirium occurred.
    - 5 mg intravenous haloperidol once per day, 5 day intervention period postoperatively (Kaneko 1999)

Concurrent medications were not reported in three studies (Liptzin 2005; Kalisvaart 2005; Kaneko 1999). Comorbidities were not reported in three studies (Kalisvaart 2005; Kaneko1999; Liptzin 2005). One study (Prakanrattana 2007) reported that 67% of the patients were suffering from coexisting diseases including hypertension, diabetes mellitus, cerebrovascular accident, renal failure, or atrial fibrillation and another study (Gamberini 2009) reported that patients had arterial hypertension (78%) and were being treated for diabetes mellitus (7%) and for chronic pulmonary obstructive disease (4%).

11.3 Methodological quality

The Liptzin (2005) study reported that initially 1038 patients were contacted and 732 were not followed up or refused to participate. The remaining 306 were contacted 2–3 weeks before surgery and underwent screening. From these, 90 patients were randomised, although 10 were not operated on and the results are based upon 80 patients. The study reported there were no significant differences between the randomized patients and the non participants, in relation to age, gender, ethnicity, and site of operation (knee or hip joint surgery).

The method of sequence generation was adequate in three studies (computer generated blocks of 20: Gamberini 2009; computer-generated sequence: Kalisvaart 2005; Prakanrattana 2007). Sequence generation was not reported in two studies (Kaneko 1999; Liptzin 2005).

Allocation concealment was partially met in all of the studies. Gamberini (2009) reported that optically identical solutions in identical bottles were delivered by the hospital pharmacy, labelled with a number. Kalisvaart (2005) used identical containers prepackaged by a hospital pharmacist, which were sequentially assigned; Kaneko (1999) used sealed envelopes. In the Liptizin (2005) study the patients were randomised by the research pharmacist, but no further details were given, and in the Prakanrattana (2007) study, a concealed envelope was used.

Four studies (Gamberini 2009; Kalisvaart 2005; Liptzin 2005; Prakanrattana 2007) were described as double-blind (Kalisvaart 2005: blinding was checked by interviewing the study assessors). Although in the Prakanrattana (2007) study the patients' placebo was an antiseptic strip rather than tablet, the authors stated that the assessors were blind to treatment. The Kaneko (1999) study did not report on blinding, although a placebo was used.

An *a priori* sample size calculation was reported in three studies (Kalisvaart 2005; Liptizin 2005; Prakanrattana 2007). The Gamberini (2009) study reported that a sample size of 120 was required to detect a relative risk reduction of 50%, with 80% power at a 5% significance level. One study (Kalisvaart 2005) reported a sample size of 206 patients per group was required to detect a 13% decrease in risk with 80% power at a 5% significance level. The sample sizes included in this study (n= 430), slightly exceeded this sample size estimate. The Liptzin (2005) study reported that a sample of 80 was required to have an 80% power to detect a difference of 22% in the study groups at a one-sided significance level of 5% assuming a delirium rate of 44% in the placebo group. Another study (Prakanrattana 2007) required a sample size of 63 per group to detect a 30% reduction in risk with 90% power at a 5% significance level; 63 patients per group were recruited and completed the study.

All studies demonstrated baseline comparability.

The Kalisvaart (2005) study reported no significant differences in mean age, proportion of males to females, Mini-mental state examination scores, visual acuity, health scores, geriatric depression scores, Barthel Index, or baseline risk of delirium between treatment and control groups. The Kaneko (1999) study reported no differences in the proportion of males to females by group, pre-existing diseases, preoperative medicines, duration of operation and anesthesia. They did observe that fewer patients in the haloperidol group had premorbid cognitive impairment (5% versus 10% in the placebo group), but the difference was not statistically significant. In the Liptzin (2005) study patients were comparable at baseline for age, gender, ethnicity, the surgeon who operated, the joint operated on and the MMSE questionnaire and clock-drawing test scores. The Prakanrattana (2007), study demonstrated baseline comparability between intervention groups for age, proportion of males to females, weight, New York Heart Association functional class, coexisting disease, type of operation (coronary artery bypass graft, valve or others), anaesthesia time, cardiopulmonary bypass time, and aortic cross-clamp time. In the Gamberini (2009) study patients were comparable for age, gender, baseline MMSE, baseline clock-drawing test scores, pre-existing diseases, type of operation (CABG, valve repair).

One study (Prakanrattana 2007) reported no missing participants; all patients were included in the analysis.

Three studies (Gamberini 2009; Kalisvaart 2005; Kaneko 1999) reported acceptable missing levels of data (that is less than 20%).

- The Gamberini (2009) study reported there was missing data for 25% (15/61) and 24% (14/59), in the intervention and control groups respectively. The study reported that only patients who were not assessed with CAM within 6 days after surgery (4/61: 3/59) were excluded from the analysis; however, the authors reported that an intention to treat analysis was carried out.
- In the Kaneko (1999) study 5% (2/40) in the intervention group and 0% in control group were missing, and the authors analysed all available participants in their analyses (n = 78).
- In the Kalisvaart (2005) study, 5% (11/212) were lost to follow-up in the treatment group and 11% (24/218) were lost to follow-up in the placebo group. However the authors analysed all patients who were randomised (ITT analysis).

One study (Liptzin 2005) had inadequate levels of missing data (more than 20% missing data in each group). Originally 90 patients were included in the study, but ten patients were not included in the final analyses because they were not operated on, or took no further part in the analysis; the groups to which they were assigned were not reported. Of the remaining 80 patients, a further 11/39 (28%) and 11/41 (27%) did not complete the study. A per protocol analysis was reported based on the 80 patients, although it was not clear what was assumed about the missing data.

Methods to assess concordance were partially reported in Kalisvaart (2005). They stated that clinical staff recorded the level of adherence to the intervention, but it was not stated how this was done. Concordance was determined by patients keeping records of their medication usage, and this was assessed by a research assistant (Liptzin 2005). Methods to assess concordance were not reported in the remaining studies.

The method of delirium assessment was:

- **adequate** in three studies (Kalisvaart 2005; Liptizin 2005; Prakanrattana 2007)
  - One study used the DSM-IV criteria (Liptzin 2005)
  - One study used the CAM and DSM-IV criteria (Kalisvaart 2005)
  - One study used the CAM-ICU instrument (Prakanrattana (2007)

- **partially adequate** in one study (Gamberini 2009). The Gamberini (2009) study used the CAM instrument in both the surgical and ICU setting.

Method of delirium assessment was unclear in one study (Kaneko 1999). The DSM-IV and DSM III-R criteria were used for 'psychotic diagnoses' and also stated that delirium was 'clinically diagnosed'. Data were collected from the patients and nursing charts on the fifth day after surgery; it was not clear if the charts were used to record delirium.

One study (Kalisvaart 2005) assessed severity using the DRS-R-98 [range 0 (no severity) to high 45 (high severity)], MMSE, and the Digit Span test [assessment of attention, range 0 (no attention) to 42 (good attention)].

All studies evaluated the incidence of delirium as a primary outcome. Secondary outcomes were: severity of delirium (Kalisvaart 2005), duration of delirium (Gamberini 2009; Kalisvaart 2005; Kaneko 1999; Liptzin 2005) and adverse events (Kalisvaart 2005; Kaneko 1999), length of hospital stay (Gamberini 2009; Kalisvaart 2005; Liptzin 2005; Prakanrattana 2007), length of ICU stay (Gamberini 2009; Prakanrattana 2007), and sleep-wakefulness rhythm (Kaneko 1999).

Overall two studies were considered to have a higher risk of bias for the following reasons:

- The method of measurement of delirium was unclear (Kaneko (1999).
- Inadequate levels of missing data [over 20%] (Liptzin 2005)

The use of rescue medication in the Kalisvaart (2005) study may have led to confounding for the following outcomes: duration of delirium, severity of delirium and length of stay.

11.4 Results

11.4.1 Acetylcholinesterase inhibitor versus placebo

1. Incidence of postoperative delirium (endpoint 28 days)

Meta-analysis of two studies (Gamberini 2009; Liptzin 2005) with 193 patients, comparing acetylcholinesterase (ACH) with placebo showed_no significant difference in the incidence of delirium between the groups (RR 1.11 (95% CI 0.69 to 1.79)); although the results are very imprecise (figure 11.1, Appendix K);

2. Duration of postoperative delirium

Two studies (Gamberini 2009; Liptzin 2005) reported the duration of postoperative delirium.

The Gamberini (2009) study compared rivastigmine versus placebo, in 113 patients and reported there was no difference in the duration of delirium. The results from this study are not shown on the forest plot because study reported values for the median and range. The reported median and range were as follows: 2.5 days (range 1 to 5) and 2 days (range 1 to 6) for the rivastigmine and placebo groups respectively (reported p value= 0.3).

The remaining study (Liptzin 2005) comparing donepezil with placebo in 80 patients found no significant difference in the duration of postoperative delirium (end point) (figure 11.2, Appendix K); mean difference (MD) –0.30 days (95%CI –0.67 to 0.07), for a placebo group duration of 1.3 days; the results are imprecise. The standard deviation in the donepezil group was stated to be zero, but for the purposes of analysis this was assumed to be 0.001.

3. Length of hospital stay

Two studies (Gamberini 2009; Liptizin 2009) reported the length of stay. The Gamberini (2009) study reported the median and range and the results for this study are not shown on the forest plot. The (Gamberini 2009) study comparing rivastigmine versus placebo in 113 patients reported there was no difference in the length of hospital stay; the median and range was 13 days (range 7 to 39) for both the rivastigmine and placebo groups respectively (reported p value = 0.3).

One study (Liptzin 2005) comparing donepezil with placebo in 80 patients found no significant difference in the length of hospital stay (endpoint 28 days) between the groups (figure 11.3, Appendix K); MD 0.20 days (95%CI –0.10 to 0.50). There was imprecision because of the small sample size.

4. Length of ICU stay

One study (Gamberini 2009) comparing rivastigmine versus placebo in 113 patients reported there was no difference in the length of ICU stay; the median and range were as follows: 2 days (range 2 to 7) and 2 days (range 2 to 6) for the rivastigmine and placebo groups respectively (reported p value: 0.9). This outcome is not included in the GRADE evidence summary.

5. Number of patients discharged to a rehabilitation facility (endpoint 28 days)

Analysis of one study comparing donepezil with placebo in 80 patients found no significant difference between the groups for the number of patients discharged to 'a rehabilitation facility', but it was not clear what this facility was (figure 11.4, Appendix K); RR 0.87 (95%CI 0.68 to 1.10). There was some imprecision in this outcome.

6. Use of rescue medications

The Gamberini (2009) study reported the use of haloperidol and lorazepam rescue medications. 32%: and 30% of the patients receiving rivastigmine and placebo respectively were given haloperidol (p=0.9). 61% and 68%, of the patients receiving rivastigmine and placebo, respectively were given lorazepam; p=0.3). There were no significant differences between the two groups in the number of patients who received the rescue medications.

Please refer to table 11.3 for the GRADE evidence summary for cholinesterase inhibitors.

Table 11.3: GRADE evidence summary - cholinesterase inhibitors vs placebo

Outcome/ Author (year)	*Meta-analysis details*	*Summary statistics*	*Comments*	*GRADE details*	*GRADE Comments*
Incidence of delirium (Liptzin 2005; Gamberini 2009)	2 trials; 193 patients; from RCT p=0.84; I^2 =0%	RR=1.11 (95%CI 0.69,1.79)	No significant difference in effect between the acetylcholinesterase and placebo groups.	• Study quality: serious limitation incomplete follow up • Directness: Direct • Imprecision: CI crosses appreciable harm/benefit • Inconsistency: consistent • Reporting bias: Adequate	
GRADE evidence rating:Very low					
Duration of delirium (Liptzin 2005)	1trial; 90 patients; from RCT	MD=-0.3 (95% CI -0.67,0.07)	No significant difference at endpoint 28 days	• Study quality: serious limitation - incomplete follow up • Directness: Direct • Imprecision: CI crosses MID • Inconsistency: consistent • Reporting bias: Adequate	MID=1 day; OIS=260;
GRADE evidence rating:Low					
Length of stay in hospital (Liptzin 2005)	1trial; 90 patients; from RCT	MD=0.2 (95% CI -0.1,0.5)	No significant difference in length of hospital stay (endpoint 28 days)	• Study quality: serious limitation - incomplete follow up • Directness: Direct • Imprecision: CI crosses MID • Inconsistency: consistent • Reporting bias: Adequate	MID: 1 day; More than 20% missing data;
GRADE evidence rating: Very low					
Number of patients discharged to rehab facility (Liptzin 2005)	1trial; 90 patients; from RCT	RR=0.87 (95% CI 0.68,1.1)	No significant difference between the donepezil and placebo groups at endpoint 28 days	• Study quality:serious limitation - incomplete follow up • Directness: Direct • Imprecision: CI crosses appreciable harm/benefit • Inconsistency: consistent • Reporting bias: Adequate	More than 20% missing data;
GRADE evidence rating: Very low					
Use of rescue medications-haloperidol (Gamberini 2009)	1trial; 113 patients; from RCT	RR=0.96 (95% CI 0.56,1.66)	No significant difference between the acetylcholinesterase and placebo groups.	• Study quality:no serious limitations • Directness: Direct • Imprecision: CI crosses appreciable harm/benefit • Inconsistency: consistent • Reporting bias: Adequate	
GRADE evidence rating: Low					
Use of rescue medications-lorazepam (Gamberini 2009)	1trial; 113 patients; from RCT	RR=0.89 (95% CI 0.67,1.17)	No significant difference between the acetylcholinesterase and placebo groups.	• Study quality:no serious limitation • Directness: Direct • Imprecision: CI crosses appreciable harm/benefit • Reporting bias: Adequate	
GRADE evidence rating: Low					

Typical antipsychotics

Typical antipsychotics versus placebo

1. Incidence of postoperative delirium

Two studies (Kaalisvaart 2005; Kaneko 1999) reported the use of haloperidol versus placebo on incidence of postoperative delirium. The Kaalisvart (2005) study reported that all patients received a proactive geriatric consultation, thus the study was investigating the adjunctive effect of haloperidol. Therefore, these two studies are reported separately on the forest plots (figure 11.5, Appendix K)

- One study (Kalisvaart 2005) with 440 patients showed no significant difference in the incidence of postoperative delirium; RR 0.91 (95% CI 0.59 to 1.42).
- The Kaneko (1999) study with 78 patients showed a small significant effect [0.32 (95% CI 0.12 to 0.91)]. We note this study was at higher risk of bias.

2. Severity of delirium

Two studies (Kalisvaart 2005; Kankeo 1999) evaluated the severity of delirium, and only Kalisvaart (2005) presented data for analysis. In 78 patients who had delirium, Kalisvaart (2005) used the highest value obtained during delirium, on the DRS-R-98 scale, (maximum value on this scale is 39) to assess the severity of delirium. The analysis demonstrates a significant effect in favour of haloperidol: MD –4.01 (95% CI –5.87 to -2.15; figure 11.6, Appendix K). It is noted that the severity of delirium may have been confounded by the use of rescue medication.

The Kaneko (1999) study reported that the postoperative delirium was more severe in the placebo group (no data or statistical analyses were presented).

3. Duration of delirium

Two studies (Kalisvaart 2005; Kaneko 1999) evaluated the duration of delirium, and only Kalisvaart (2005) presented data for analysis. The analysis demonstrates that patients who received haloperidol, had, on average, significantly fewer days of delirium (of those who had delirium): MD –6.40 (95% CI –9.38 to –3.42; figure 11.7, Appendix K). It is noted that the duration of delirium may have been confounded by the use of rescue medication and that results were reported only for those with delirium. We also note that the distribution for the duration of delirium is skewed for both the intervention and placebo groups (mean values less than twice the standard deviation). The Kaneko (1999) study reported that the duration of postoperative delirium was longer in the placebo group (no data or statistical analyses were presented).

4. Length of hospital stay

The Kalisvaart (2005) study demonstrated that the number of days spent in hospital was significantly shorter in patients who received haloperidol compared to patients who received placebo in addition to the proactive geriatric consultation; MD –5.50 (-8.17 to -2.83; figure 11.8, Appendix K). The study included the results for hospital length of stay in a table that was stated to apply to patients with delirium only. However, we have assumed this should refer to all patients; we also note that the summary statistics are incorrectly noted in the table in the report (the upper confidence limit is lower than the mean). Furthermore, the distribution for length of stay is skewed for both intervention and placebo groups.

5. Adverse events

Two studies (Kalisvaart 2005; Kaneko 1999) evaluated adverse events. Kalisvaart (2005) reported that there were no drug-related side effects and no sedation events were reported, other than those related to morphinomimetics. Only Kaneko (1999) presented data for analyses; they observed that one patient in the treatment group developed transient tachycardia. The results are very imprecise (figure 11.9, Appendix K).

Please refer to table 11.4 for the GRADE evidence summary for typical antipsychotics.

Table 11.4: GRADE evidence summary: Typical antipsychotics vs placebo

Outcome/ Author (year)	*Meta-analysis details*	*Summary statistics*	*Comments:*	*GRADE details:*	*GRADE Comments*
Incidence of delirium (Kalisvaart 2005)	1 trial; 430 patients from RCT p=0.07; I^2=70%	RR=0.91 (95%CI 0.59,1.42)	No significant difference between the haloperidol and placebo groups.	• Study quality:no serious limitation • Directness: Direct • Imprecision: CI crosses appreciable harm/benefit • Inconsistency: consistent • Reporting bias: Adequate	all pts received proactive geriatric consultation; downgraded by 2 for imprecision
GRADE evidence rating: Low					
Incidence of delirium (Kaneko 1999)	1trial; 78 patients; from RCT	RR=0.32 (95%CI 0.12,0.91)	No significant difference between the haloperidol and placebo groups	• Study quality:serious limitation method of assessment of delirium unclear • Directness: Direct • Imprecision: CI crosses appreciable harm/benefit • Inconsistency: consistent • Reporting bias: Adequate	
GRADE evidence rating:Low					
Duration of delirium (Kalisvaart 2005)	1trial; 430 patients from RCT	MD=-6.4 (95%CI -9.38,-3.42)	Statistically significant fewer days of delirium in the haloperidol group	• Study quality:serious limitation some confounding • Directness: Direct • Imprecision: Number of patients < 400 • Inconsistency: consistent • Reporting bias: Adequate	MID=1 day; Use of rescue meds may have confounded this outcome; Patients also received proactive geriatric consultation;
GRADE evidence rating: Low					
Severity of delirium (Kalisvaart 2005)	1trial; 430 patients from RCT	MD=-4.01 (95%CI -5.87,-2.15)	Statistically significant in favour of the haloperidol group on the DRS-R-98(0-39)	• Study quality:serious limitation some confounding • Directness: Direct • Imprecision:CI crosses MID • Inconsistency: consistent • Reporting bias: Adequate	MID=7.8;Severity of delirium (of those who had delirium) may have been confounded by the use of rescue medication. Pts also received proactive geriatric consultation; Results reported for those who had delirium (n=68); a-priori sample size calc=208
GRADE evidence rating: Low					
Length of stay in hospital (Kalisvaart 2005)	1trial; 430 patients from RCT	MD=-5.5 (95%CI -8.17, -2.83)	Statistically significantly shorter length of stay in patients who received haloperidol	• Study quality:serious limitation some confounding • Directness: Direct • Imprecision: Number of patients < 400 • Inconsistency: consistent • Reporting bias: Adequate	MID=1; Use of rescue meds may have confounded this outcome; Patients received proactive geriatric consultation
GRADE evidence rating: Low					

Table 11.4: GRADE evidence summary: Typical antipsychotics vs placebo (continued)

Outcome/ Author (year)	Meta-analysis details	Summary statistics	Comments	GRADE details	GRADE Comments
Adverse event (extrapyramidal) (Kalisvaart 2005)	1trial; 430 patients; from RCT	RR=1	Study reported no extrapyramidal events	• Study quality:no serious limitation • Directness: Direct • Imprecision: Number of events < 300 • Inconsistency: consistent • Reporting bias: Adequate	
GRADE evidence rating: Low					
Adverse events (sedation) (Kalisvaart 2005)	1trial; 430 patients; from RCT	RR=1	Study reported no sedation event in either group	• Study quality:no serious limitation • Directness: Direct • Imprecision: Number of events < 300 • Inconsistency: consistent • Reporting bias: Adequate	
GRADE evidence rating: Low					
Adverse events (tachycardia) (Kaneko 1999)	1trial; 78 patients; from RCT	RR=3.15 (95%CI 0.13, 75.12)	No significant difference	• Study quality: serious limitation-method of assessment of delirium unclear • Directness: Direct • Imprecision: CI crosses appreciable harm/benefit • Inconsistency: consistent • Reporting bias: Adequate	
GRADE evidence rating: Low					

11.4.2 Atypical antipsychotics

Atypical antipsychotics versus placebo

1. Incidence of delirium

In the Prakanratta (2007) study, delirium was recorded twice daily in the ICU and once daily on discharge from the ICU. The study reported results as percentages, so we calculated the number of patients with delirium.

In one study (Prakanrattana 2007) comparing risperidone with placebo in 126 patients, there were significantly fewer patients with delirium in the risperidone group compared with placebo, although the result was imprecise (figure 11.10, Appendix K); RR 0.35 (95%CI 0.16 to 0.77) which corresponds to a number needed to treat of 5 (95%CI 3 to 14), for a control group rate of 32%. The authors reported that all episodes of delirium occurred within the first three postoperative days.

2. Length of ICU stay

There was no significant difference between the treatment groups for the number of days spent in ICU; MD 0.10 (95% CI –0.64 to 0.84; figure 11.11, Appendix K). The results are very imprecise (clinically important difference: 0.5 days).

3. Length of hospital stay

There was no significant difference between the treatment groups for the number of days spent in hospital; MD 0.20 (95% CI –1.74 to 2.14; figure 11.12, Appendix K). The results are very imprecise.

Please refer to table 11.5 for the GRADE evidence summary for atypical antipsychotics.

Table 11.5: GRADE evidence summary - A*typical antipsychotics vs placebo*

Outcome/ Author (year)	*analysis details*	*Summary statistics*	*Comments*	*GRADE details*	*GRADE Comments*
Incidence of delirium (Prakanrattana 2007)	1trial; 126 patients; from RCT	RR=0.35 (95%CI 0.16,0.77)	Significantly fewer patients with delirium in the risperidone group	• Study quality:no serious limitation • Directness: Direct • Imprecision: CI crosses appreciable harm/benefit • Inconsistency: consistent • Reporting bias: Adequate	
GRADE evidence rating: Moderate					
Length of stay in hospital (Prakanrattana 2007)	1trial; 126 patients; from RCT	MD=0.2 (95%CI -1.74, 2.14)	No significant difference in length of hospital stay	• Study quality: no serious limitation • Directness: Direct • Imprecision: CI crosses MID • Inconsistency: consistent • Reporting bias: Adequate	MID=1; CI crosses both threshold so downgraded by 2 for imprecision
GRADE evidence rating: Very low					
Length of stay in ICU (Prakanrattana 2007)	1trial; 126 patients; from RCT	MD=0.1 (95%CI -0.64, 0.84)	No significant difference in number of days spent in the ICU	• Study quality:no serious limitation • Directness: Direct • Imprecision: CI crosses MID • Inconsistency: consistent • Reporting bias: Adequate	MID=0.5 days
GRADE evidence rating: Low					
Adverse event (cardiovascular instability) (Prakanrattana 2007)	1trial; 126 patients; from RCT	RR=0.76 (95%CI 0.18, 3.27)	Not significant	• Study quality:no serious limitation • Directness: Direct • Imprecision: Wide CI • Inconsistency: consistent • Reporting bias: Adequate	
GRADE evidence rating: Low					

Overall summary of results

Overall summary of results for pharmacological prevention of delirium in hospital setting are reported in table 11.6.

Table 11.6: summary of results: pharmacological prevention of delirium in hospital setting

Outcomes [Summary statistic]	Aceytlcholinesterase vs placebo	Typical anitpsychotics vs placebo	Typical anitpsychotics vs placebo (proactive geriatric consultation for all)	Atypical antipscyhotics vs placebo
Incidence of delirium RR(95% CI)	1.11 (0.69 to1.79)	0.32 (0.12 to 1.91)	0.91 (0.59 to 1.42)	0.35 (0.16 to 0.77)
Severity of delirium MD (95% CI)			-4.01 (-5.87 to -2.15)	
Duration of postoperative delirium MD (95% CI)	-0.3 (-0.67 to 0.07)		-6.40 (-9.38 to -3.42)	
Length of hospital stay MD (95% CI)	0.20 (-0.10 to 0.50)		-5.50 (-8.17 to -2.83)	0.20 (-1.74 to 2.14)
Length of ICU stay	Median and range: 2 days (2 to 7) vs 2 days (2 to 6) for the rivastigmine and placebo groups, respectively			0.10 (-0.64 to 0.84)
Discharge to rehabilitation unit RR(95% CI)	0.87 (0.68 to 1.10)			
Use of rescue medicine (haloperidol) RR(95% CI)	0.96 (0.56 to 1.66)			
Use of rescue medicine (lorazepam) RR(95% CI)	0.89 (0.67 to 1.17)			
Adverse events RR(95% CI)		Tachycardia: 3.15 (0.13 to 75.12)	Reported no adverse events	

11.5 Health economic evidence

11.5.1 Pharmacological interventions for the prevention of delirium in a hospital setting

One economic evaluation study was included as evidence (Bracco 2007). This was a non-randomised clinical trial of 1293 patients who underwent cardiac surgery in Canada. The objective was to examine outcomes and use of intensive care resources for a cohort of consecutive patients who underwent cardiac surgery with or without thoracic epidural anaesthesia. The intervention group received thoracic epidural anaesthesia for cardiac surgery. The control group did not receive thoracic epidural anaesthesia. Detailed description of intervention and control strategies is given in Appendix J (table J1). The intervention shortened ventilation time and the length of stay in the ICU by 9.6 hours and 12.7 hours respectively after adjusting for type of surgery in a multivariate analysis. This reduction decreased the ICU and mechanical ventilation costs by US$2700 and US$700 respectively, per patient. The additional cost of thoracic epidural use was given as US$82. Post-operative delirium complication rate was reported as 24/506 in the intervention arm, and 20/787 in the control arm. This was measured using CAM-ICU scale. A relative risk of 0.3 was reported. Intensive care unit mortality rate of 2/506 was also reported in the intervention arm and 14/787, in the control arm. A multivariate analysis for mortality was not statistically significant. Cost data was taken from the literature and QALY estimates were not reported. The study sample was not randomised and there was no sensitivity analysis on variables whose values will probably be uncertain. The results are not directly applicable and should be cautiously interpreted.

11.6 Clinical evidence statements

11.6.1 Acetylcholinesterase inhibitor versus placebo

- Meta-analysis of 2 RCTs comparing acetylcholinesterase with placebo showed:
    - no significant effect on the incidence of delirium (very low quality).

11.6.1.1 Donepezil versus placebo

- 1 RCT comparing donepezil with placebo showed:
    - no significant effect on the length of hospital stay and the number of patients discharged to a rehabilitation facility (low quality).

11.6.2 Typical antipsychotics

Haloperidol versus placebo

- 1 RCT comparing haloperidol with placebo as an adjunct to a proactive geriatric consultation (non-pharmacological intervention) showed:
    - no significant effect on the incidence of postoperative delirium (low quality).
    - a significantly lower severity of delirium and fewer days of delirium in favour of the haloperidol group (low quality)
    - a significantly shorter length of hospital stay in patients who received haloperidol (low quality).
- 1 RCT comparing haloperidol with placebo showed:
    - no significant effect on the incidence of postoperative delirium (low quality)
    - no difference between the groups for the number of adverse events (transient tachycardia); (insufficient evidence).

11.6.2.1 Atypical antipsychotics versus placebo

- 1 RCT conducted in ICU, comparing risperidone with placebo showed:
    - a lower incidence of delirium in patients receiving risperidone (moderate quality).
- 1 RCT comparing risperidone with placebo showed:
    - no significant difference between the groups for length of stay in ICU and hospital (low quality).

11.7 From evidence to recommendations

The GDG discussed the evidence from the pharmacological prevention reviews and noted that it was limited and of low quality. The evidence was mainly from single studies and each of these had risk of bias issues; in addition, the evidence was often imprecise, sometimes indirect and showed inconsistency where there was more than one study.

- Donepezil: the study was unrepresentative of the population (patients were fit and healthy with no cognitive impairment)
- Risperidone: the study was unrepresentative of the intervention or the population [the dose used was very different from that used in clinical practice, and the study included a relatively young population (age range: 51 to 71 years) undergoing cardiac surgery].
- Haloperidol: two studies investigated haloperidol. One study had a high risk of bias and the other assessed haloperidol as an adjunct to a proactive geriatric consultation intervention. There was explained inconsistency between the studies.

The GDG was not confident in the evidence and did not make a recommendation but agreed a research recommendation for typical antipsychotics, atypical antipsychotics, acetylcholinesterase inhibitors and benzodiazepines (and Appendix H). For ethical reasons, research should only be carried out in a population at high risk of delirium.

Future research recommendation:

In hospital patients at high risk of delirium, which medication (atypical antipsychotics, typical antipsychotics, benzodiazepines, or acetycholinesterase) compared with placebo or each other is more clinically and cost effective, in preventing the development of delirium?

11.8 Recommendations

There are no recommendations for this section. In light of the evidence the GDG did not wish to make recommendations.

11 B) Prevention in long-term care

CLINICAL QUESTION: What are the most clinical and cost effective and safe pharmacological interventions for the prevention of delirium in people in long-term care?

11.9 Description of studies

One paper was evaluated for inclusion Moretti (2004). The study was an RCT.

11.9.1 Study Design

The RCT was conducted in Italy in a community based setting; this was treated as an indirect setting for long-term care. Patients without reliable carers were excluded from the trial. The funding source was not reported. Two hundred and forty six patients were randomised; the unit of randomisation was the patient.

11.9.2 Population

The patients all had an MMSE score of at least 14, indicating patients had mild to moderate dementia. All patients met the DSM-IV criteria for dementia. Patients also satisfied the criteria for probable vascular dementia, or multi-infarct dementia with the NINDS-AIREN criteria (National Institute of Neurological Disorders and Stroke and Association Internationale pour la Recherché et l'Enseignement en Neurosciences). Their ages ranged from 65–80 years with a mean age of 76 years. One hundred and sixteen men and 130 women were included in the study, although 12 patients died during the study and four refused to participate; all data were based on the remaining groups of 115 in the rivastigmine group and 115 in the aspirin group. All were ambulatory outpatients living in the community. Their delirium risk was not stated in the study. The comorbidity was vascular dementia, although other comorbidities were implied because of the drugs patients were taking; patients with previous psychiatric illness or central nervous system disorders or alcoholism were excluded.

11.9.3 Interventions

The included study investigated rivastigmine, a cholinesterase inhibitor, compared with cardio-aspirin (considered as usual care). Participants were ambulatory outpatients and were given the interventions for 2 years after randomisation. Rivastigmine was titrated to the higher dose after the first 16 weeks. The interventions included:

- 3–6 mg/day rivastigmine
- aspirin 100 mg/day

It was assumed that the cardio-aspirin was representing usual care and was not an active intervention.

11.9.4 Comparisons

The following comparison was carried out:

- rivastigmine versus cardio-aspirin for 2 years (Moretti 2004)

The patients were allowed to continue taking their existing drug therapies, anti-hypertensives, anti-dyslipidemic, anti-diabetic drugs, diuretics and bronchodilators.

Patients received benzodiazepines or neuroleptic drugs during delirium, which were significantly less in the intervention group. This may have led to confounding for some outcomes, but would serve to underestimate the size of the effect.

11.10 Methodological quality

The methods of sequence generation and allocation concealment were not described, although the patients were matched for age and education level. It was not reported if all eligible patients were recruited.

The study did not report whether patients and investigators were blinded to treatment allocation. An *a priori* sample size calculation was not reported.

Originally 246 patients were included in the study, but 16 were not included in the final analyses (7% missing data; 12 patients died during the follow up and four refused to participate in the follow up). The groups to which they were assigned were not reported. The remaining 230 patients completed the two year follow up. Patients were found to be comparable at baseline on the following scales: BEHAVE-AD (Behavioural Pathology in Alzheimer's Disease Rating); Clinical Dementia Rating; and the Cumulative Illness Rating Scale. Concordance was monitored by care givers, who controlled the intake of drugs.

Delirium was assessed using the Confusion Assessment Method (CAM).

Overall, the study may have been at a higher risk of bias because allocation concealment and blinding were unclear; appear to have a higher potential for bias, although the differential use of rescue medication may have led to confounding for some outcomes.

11.11 Results

11.11.1 Rivastigmine versus usual care (aspirin)

Incidence of delirium (endpoint 2 years)

Analysis of one study in 230 patients showed that the incidence of delirium was significantly lower in the rivastigmine group compared with usual care; RR 0.65 (95%CI 0.50 to 0.85), which corresponds to a number needed to treat of 5 (95%CI 4 to 12), for a control group rate of 62%. The result was imprecise (figure 11.13 , Appendix K).

Duration of delirium

Analysis of one study in 230 patients showed that the duration of delirium was significantly shorter in the rivastigmine group compared with usual care (figure 11.14a , Appendix K); MD –3.86 days (95%CI –4.44 to –3.28), for a control group duration of 7.86 days. It was unclear whether the duration of delirium was reported just for those who had delirium or was a mean across all patients: the paper describes 'the main duration of the delirium'. In addition, the different standard deviations across the groups, indicates the mean may just be for those with delirium. Figure 11.14b (Appendix K) shows the analysis with this assumption; the only difference is a slightly wider CI; MD –3.86 days (95%CI –4.66 to –3.06).

Cognitive impairment

The study assessed global performance using the Clinical Dementia Rating (scale 0–3), and reported the change from baseline at 12 months. Analysis of 230 patients showed there was no significant difference between the groups, although the table in Moretti (2004) stated the difference was significant (figure 11.15 , Appendix K);

Behavioural disturbance (change score at 1 year)

Analysis of one study in 230 patients showed that behavioural disturbance was significantly lower in the rivastigmine group compared with usual care (figure 11.16a , Appendix K). The study used the BEHAVE-AD to assess individual behavioural items on this scale (delusions, hallucinations, activity alterations, aggressiveness, anxiety/phobia, sleep disturbances, affective disturbancesand anxiety). All individual items were stated to be statistically significant, with the exception of delusions. The overall score showed a statistically significant mean difference, favouring the intervention; MD –39.66 (95%CI –40.06 to –39.26). This seems to be a very narrow CI, even for a change score from baseline, but if these were standard errors, rather than standard deviations (despite what was reported in the text), the standard deviations would be rather large for the intervention group (figure 11.16b , Appendix K). The assumption of a standard error gave a large significant mean difference of –39.66 (95% CI –43.91 to –35.41), favouring the intervention group.

Please refer to table 11.7 for the GRADE evidence summary for cholinesterase inhibitors.

Table 11.7: GRADE evidence summary:Cholinesterase inhibitors vs usual care

Outcome/ Author (year)	*Meta-analysis details*	*Summary statistics*	*Comments*	*GRADE details*	*GRADE Comments*
Incidence of delirium (Moretti 2004)	1trial; 230 patients; from RCT-Indirect [Community]	RR=0.65 (95%CI 0.5,0.85)	Significantly lower incidence of delirium in the rivastigmine group compared with usual care at endpoint 2 years	• Study quality:serious limitation-allocation conceal unclear • Directness: Indirect Setting- Minor, community • Imprecision: CI crosses appreciable harm/benefit • Inconsistency: consistent • Reporting bias: Adequate	Allocation concealment and blinding unclear
GRADE evidence rating: Very low					
Duration of delirium (Moretti 2004)	1trial; 230 patients; from RCT-Indirect [Community]	MD=-3.86 (95%CI -4.45, -3.27)	Duration of delirium was significantly shorter in the rivastigmine group compared with usual care	• Study quality: serious limitation some confounding • Directness: Indirect Setting- Minor, community • Imprecision: Number of patients < 400 • Inconsistency: consistent • Reporting bias: Adequate	MID:1 day;Differential use of rescue medication may have led to some confounding; Duration of delirium (unclear if patients with delirium or a mean across all patients). Results for mean across all patients presented here. Allocation concealment & blinding unclear
GRADE evidence rating: Very low					
Cognitive impairment (Moretti 2004)	1trial; 230 patients; from RCT-Indirect [Community]	MD=-0.21 (95%CI -0.98, 0.56)	No significant difference in global performance on Clinical Dementia Rating Scale (0-3)	• Study quality: serious limitations-some confounding • Directness: Indirect Setting- Minor, community • Imprecision: CI crosses MID • Inconsistency: consistent • Reporting bias: Adequate	MID: 0.6; Allocation concealment & blinding unclear; Differential use of rescue medication may have led to some confounding;
GRADE evidence rating: Very low					
Behavioural disturbance (Moretti 2004)	1trial; 230 patients; from RCT-Indirect [Community]	MD=-39.66 (95%CI -43.91, -35.41)	A significant difference in global performance on the BEHAVE-AD scale favouring the rivastigmine group	• Study quality: serious limitations-some confounding • Directness: Indirect Setting- Minor, community • Imprecision: CI crosses MID • Inconsistency: consistent • Reporting bias: Adequate	Allocation concealment & blinding unclear; Differential use of rescue medication may have led to some confounding;
GRADE evidence rating: Very low					

Overall summary of results for pharmacological prevention of delirium in long term care setting are reported in table 11.8.

Table 11.8: summary of results: pharmacological prevention of delirium in long-term care setting

Outcomes [Summary statistic]	Aceytlcholinesterase vs placebo
Incidence of delirium RR(95% CI)	1.11 (0.69 to1.79)
Duration of delirium MD (95% CI)	All patients: -3.86 (-4.45 to -3.27)
Duration of delirium MD (95% CI)	Assuming mean across patients with delirium: -3.86 (-4.66 to -3.06)
Cognitive impairment MD (95% CI)	-0.21 (-0.98 to 0.56)
Behavioural disturbance MD (95% CI)	Change scores: -39.66 (-40.06 to -39.26)
Behavioural disturbance MD (95% CI)	Overall change: -39.66 (-43.91 to -35.41)

11.12 Health economic evidence

No relevant health economic papers were identified.

11.13 Clinical evidence statements

- 1 RCT of very low quality comparing rivastigmine with usual care (indirect evidence) showed that the rivastigmine group had significantly:
  - lower incidence of delirium (endpoint 2 years).
  - fewer days of delirium.
  - lower behaviour disturbances (change score at 1 year).
  - However at 1 year there was no significant difference between the groups for change in cognitive impairment from baseline.

11.14 From evidence to recommendations

There was one very low quality study in an indirect population for the long term care setting (the community). The GDG were not confident in the evidence to make a recommendation on the basis of this study.

11.15 Recommendations

There are no recommendations for this section. In light of the evidence the GDG did not wish to make recommendations.

12 Treatment of delirium: non-pharmacological (hospital setting)

CLINICAL QUESTION: What are the most clinical and cost effective multicomponent interventions for treating people with delirium in hospital?

12.1 Clinical introduction

Despite the advances in medical science over the last three decades, mortality and morbidity from delirium have remained unchanged and health costs for this syndrome remain high. Current management of delirium relies on early recognition, elimination or correction of underlying causal factors and general symptomatic and supportive measures. However, there is much uncertainty about the effectiveness of various interventions.

Early recognition and investigation of delirium is challenging and studies have repeatedly shown that delirium is missed in two-thirds of patients in hospitals. Moreover, delirium often has multi-factorial causes and multiple potential consequences. This has led to suggestions that multicomponent interventions, including non-pharmacological interventions might be appropriate for the treatment of delirium, and several such interventions have been investigated.

12.2 Description of studies

Details of included, excluded papers together with study design are reported in table 12.1.

Table 12.1: study inclusion, exclusion and design

Papers	Comments	Study
N= 9 evaluated for inclusion		
N= 3 excluded	Reasons for exclusion are reported in Appendix G.	
N= 0 identified in update searches	None Identified	
N= 7 reports of 6 studies included*	Study designs 3 RCTs	Cole 1994; Cole 2002; Pitkala 2006; Pitkala 2008
	3 prospective studies with historical control groups	Milisen 2001; Naughton 2005; Rahkonen 2001

One study (Pitkala 2006) had more than one report (Pitkala 2006 and Pitkala 2008); hereafter these studies are referred to by the first name reports, but separately in the results section.

12.2.1 Study Design

None of the studies were conducted in the UK. Information on study sizes, geographical location and funding are described in table 12.2.

Table 12.2: study characteristics

Study	Size (N)	Geographical location	Funding
Cole 1994	88	Canada	Non-industry sources
Cole 2002	227	Canada	Non-industry sources
Milisen 2001	120	Belgium	Non-industry sources
Naughton 2005	374	USA	Non-industry sources
Pitkala 2006	174	Finland	Non-industry sources
Rahkonen 2001	102	Finland	Not stated

The unit of randomisation in the RCTs was at patient level. In one of the historical controlled trial (Naughton 2005), eligible patients were enrolled at two different time periods. The Naughton (2005) study considered three groups of patients: those studied in the pre-intervention and two groups after the intervention had ceased – these patients were studied 4 and 9 months after the initial education phase of the intervention was completed.

12.2.2 Population

All studies took place in a hospital setting; the intervention in the Rahkonen (2001) study continued after discharge from hospital as it involved support for the patient over 3 years; Patients were all admitted to medical wards, with the exception of one study (Milisen 2001). Patients were included in each of the studies if they had delirium: this was based on screening with CAM, apart from the Rahkonen (2001) study which specified that the diagnosis was based on DSM-III-R but did not specify that CAM was used. In the Pitkala (2006) study, patients found to be positive on CAM screening had their diagnosis confirmed by a physician using DSM-IV criteria.

The Naughton (2005) study reported that for all patients admitted to the Acute Geriatric Unit (AGU) one criterion for admission was cognitive impairment (score less than 25 on the MMSE).

Some patients had dementia in the studies, (Cole 1994; Cole 1992; Pitkala 2006) ranging from 10% to 58% of participants, except in the Rahkonen (2001) study, where patients with dementia were excluded.

Method of assessment of dementia varied and the following methods were reported:

- SPSMQ; scale scores range from: 0 to 10, from no impairment to severe; score of 5 or more indicative of moderate to severe cognitive impairment) (Cole 1994)

- IQCODE (Cole 2002);

- Medical record data for the diagnosis of preexisting dementia (Milisen 2001)

- Clinical Dementia Rating Scale (CDR; scale scores range from 0.5 to 3, from very mild to severe dementia), DSM-IV criteria for dementia or diagnosis by specialist using standard diagnostic tests (no further details were given) (Pitkala 2006).

The mean age across the studies was 81 to 85.5 years; the studies had a mixed gender population with a majority of females (Cole 1994: 65%; Cole 2002: 54%; Milisen 2001: 81%; Naughton 2005: 63%; Pitkala 2006: 74%; Rahkonen 2001: 90%). Ethnicity was not reported in any of the studies.

12.2.3 Interventions

The included studies investigated multicomponent interventions in a hospital (or hospital plus community in the case of Rahkonen 2001) setting for the treatment of delirium (see table 12.3, Appendix D).

Nursing intervention protocol (Cole 1994, Cole 2002),

This intervention comprised of a multidisciplinary team consisting of geriatricians and liaison nurse.

- consultation by a geriatrician or geriatric psychiatrist (completed within 24 hours after referral)

- follow-up by a liaison nurse

    - follow up included daily visits during the patients' stay (up to a maximum of 8 weeks), liaising with family members, recording information on patient's metal status and discuss management with the patient's nurses with the use of the protocol

  - assess compliance with consultant recommendations. Where appropriate, the nurse discussed management problems with the geriatrician or geriatric psychiatrist and where necessary patient was reassessed by the specialists.

- the intervention protocol targeted the following risk factors:

  - environment (not having excessive, inadequate or ambiguous sensory input, medication not interrupting sleep, presenting one stimulus or task at a time);

  - orientation (room should have a clock, calendar, and chart of the day's schedule; evaluate need for glasses, hearing aid, interpreter)

  - familiarity (objects from home, same staff, family members staying with patient, discussion of familiar areas of interest),

  - communication (clear, slow, simple, repetitive, facing patient, warm, firm kindness, address patient by name, identify self, encourage verbal expression)

  - activities (avoid physical restraint, allow movement, encourage self care and personal activities).

The intervention in the later trial (Cole 2002) was described as more intensive than in the earlier study (Cole 1994) and the following components were added to the intervention:

- consultant not only assessed initially but also followed up the patients;

- the study nurse visited the patient 5 days per week;

- the intervention team (2 geriatric psychiatrists, 2 geriatric internists and the study nurse) met after every 8 to 10 patients were enrolled to discuss delirium management problems; and

- the study investigator met the nurse weekly to discuss problems of diagnosis, enrolment and interventions.

Multicomponent geriatric intervention (Pitkala 2006)

Patients received a comprehensive geriatric assessment, which included history taking, interview with caregiver, physical examination, assessment of cognition and physical functioning, screening for depression, nutrition, and medication review.
Other aspects of the intervention included:

- recognising delirium and any underlying conditions

- orientation (with calendars, clocks, photographs)

- physiotherapy
- general geriatric interventions (calcium and vitamin D supplements; nutritional supplements for those at risk of malnutrition or malnourished; hip protectors)
- comprehensive discharge planning (including consultation of a social worker, occupational therapist's home visit, involvement of caregivers).
- medical management (avoiding neuroleptics; administering atypical antipsychotics for hyperactive/psychotic symptoms; use of cholinesterase inhibitors if patient's cognition did not improve to MMSE score above 23).

The intervention group received significantly more atypical antipsychotic drugs than the control group (69.0% versus 29.9%, $p<0.001$), more acetylcholinesterase inhibitors (58.6% versus 9.2%, $p<0.001$), vitamin D and calcium supplements (77.0% versus 9.2%, $p<0.001$), nutritional supplements (92.0% versus 0.0%, $p<0.001$) and fewer conventional neuroleptics (8.0% versus 23.0%, $p=0.006$).

Nurse-led interdisciplinary intervention (Milisen 2001)

This intervention involved nurse education to identify high-risk patients which included:

- education: a poster was developed to educate all nurses on the essential aspects of delirium, depression and dementia. This poster included the core symptoms of delirium according to the CAM criteria, comparative features and differences between delirium, dementia and depression and the relevance of correct and early recognition of delirium;
- systematic screening of cognitive function using the NEECHAM Confusion Scale following training;
- pain management: scheduled pain medication to provide effective post-operative pain control; and
- consultative service: access to a resource nurses who were given training in identifying patients by a geriatric nurse specialist in the identification and management of older hip-fracture patients. If necessary, the resource nurses could consult with a geriatric nurse specialist or psycho geriatrician; resource nursed to help the primary nurses in implementing appropriate antidelirium interventions.
- the nurses were provided with 'A nursing guide for the evaluation of causes of delirium in elderly hospitalised patients' (as reported in Milisen 1998). The guide advised a nurse to report to the attending physician of

any changes in patient's status on the following: medication, pain, hypoxemia, dehydration, electrolyte and metabolic disturbances, and infection. The interventions are briefly described below:

- medication: to be vigilant of polypharmacy, especially anticholinergics, antiparkinsonian drugs, histamine H_2-receptor antagonists;
- pain: inquire systematically about pain; observe verbal and nonverbal expressions; use of as many possible analgesics based on nonopiod drug (e.g. paracetamol) and where required minimum dose of opioids combined with non opioid drug;
- hypoxemia: monitor abnormalities in rate, depth and quality of respiration, cyanosis, $PO_2 \leq 32$; administer oxygen as ordered; determine source of hypoxia; low respiration (<10 l/min) due to opioid intoxication; consult attending physician for treatment with naloxone as antidote; in patients undergoing surgery: monitor hypothermia and postoperative shivering; maintain optimal patient temperature by applying warming [fluids and blood; gowns and blankets; humidified oxygen]; be alert for nocturnal desaturation during the first 3 days postoperatively and especially in obese patients; administer 2 l of O_2 (unless contraindicated);
- dehydration: encourage patient to drink water regularly and when necessary prepare for blood or fluid replacement;
- electrolyte and metabolic disturbances: monitor abnormalities of blood and urine chemistry; give frequent small meals and add nutritional supplements, such as calorie/protein rich drink;
- infection: be alert for urinary tract, respiratory, mouth and feet infections; stimulate patient for adequate water intake (2 l/day) (unless contraindicated); observe for abrupt onset for fever (rectal temperature >100°F) and apply cooling techniques as needed.

Systematic intervention (Rahkonen 2001.

The intervention consisted of a case manger (nurse specialist) and an annual one-week rehabilitation period at a Brain Research and Rehabilitation Centre. Patient's rehabilitation team included the study physician, the nurse specialist, physiotherapist, neuropsychologist and occupation therapist.

- a nurse specialist trained in geriatrics and care of the elderly acted as the case manager. Patients received continuous and systematic support provided by the case manager with responsibility in supporting the patients during community care through out the 3 year follow-up acting as a counsellor and advocate and in the rehabilitation unit (as the primary care nurse);

- care in the community: arranged in consultation with relatives and health and social care services, and continuity of care was achieved with regular follow-ups, including in-home visits and ‘phone calls by the case manager. Study physician was also available for consultation and medical care throughout the follow up; and

- rehabilitation period: individually structured physiotherapy once or twice daily; mobility and other special aides for daily living (e.g. hearing aids and special shoes) were arranged when needed; patients were encouraged to participate in occupational therapy and free-time events.

Education and management intervention (Naughton 2005)

The intervention was designed to improve the recognition of delirium in medically ill older adults evaluated in the emergency department [ED triaged these patients with delirium specifically to the acute geriatric unit (AGU)]. This was achieved by addressing the following factors:

- education:
    - The charting procedures in ED were changed and physicians were reminded to evaluate adults aged 75 years and older for cognitive impairment and delirium and direct the admission to the AGU. Nurses and physicians were trained to triage patients using yes/no answers to four questions from the history and mental status examination. A study nurse periodically reported the proportion of older adults correctly admitted to the AGU from the ED.

- the education component for the AGU nurses (provided by geriatricians and geriatric nurse) involved:
    - educating on prevalence and outcome of delirium;
    - sensitivity training on cognitive impairment;
    - training on methods of mental status assessment;
    - guidelines on medication management of cognitive impairment and delirium.
    - small group consensus process used to develop assessment and charting procedures; and
    - AGU physicians were provided with information on cognitive impairment and delirium in the elderly, recommended metal status assessment procedures, and review of the intervention guidelines.

- treating underlying medical factors;
- treating precipitating factors (removing precipitating medications; addressing immobility);
- providing family support;
- using non-pharmacological support for: physically non aggressive behaviour and episodes triggered with ADL care;
- medication management: reduce the use of psychotropic medications (benzodiazepines and anticholinergics); consider using synergistic agents such as neuroleptics or antidepressants that supplement behaviour treatment; sleep medication: trazadone 50 to 100 mg; zolpidem: 5 mg;
- fewer patients in the AGU received benzodiazepines (22.6% compared with 30.9% at baseline); antihistamines (6% compared with 15.5%; $p<0.02$); increased use of antidepressants (22.7% compared with 10% at baseline; $p<0.02$); and neuroleptics (27.4% compared with 10.9% at baseline; $p<.01$)
- simplifying pain regimen (minimise p.r.n.); and
- environmental stimuli: addressing problems with environmental stimuli for example, noise, sleep disruption, disruptive room mate,
- None of the studies included more than two study arms, and the comparator in all studies was ‘usual medical care’ (no further details given).

12.2.4 Comparisons

The following comparison was carried out:

- Multicomponent intervention versus usual care.
    - Two RCTs followed patients up to 8 weeks (Cole 1994, Cole 2002) and one followed patients up to 1 year (Pitkala 2006). Of the non-RCTs, one study followed patients up to 12 days (Milisen 2001), 2 months (Naughton 2005) and 3 years (Rahkonen 2001).

Two studies (Naughton 2005; Pitkala 2006) reported concurrent medications:

- opiates (42.7%); benzodiazepines (30.9%); antihistamines (15.5%); antidepressants (10.0%); neuroleptics (10.9%)
- conventional neuroleptics (22%); atypical antipsychotics (14%) and cholinesterase inhibitors (6%) (Pitkala 2006).

12.2.5 Outcome measures

The following primary and secondary outcome measures were reported:

- primary outcomes:
  - complete response (Pitkala 2006 RCT; Naughton 2005 non RCT)
  - duration of delirium (Milisen 2001 non RCT)

- secondary outcomes:
  - cognitive impairment (Cole 1994; Pitkala 2006)
  - length of stay (Cole 1994; Cole 2002)
  - health related quality of life (Pitkala 2008)
  - discharge (higher dependency: Cole 1994: Cole 2002; long-term care: Pitkala 2006)
  - days in new long-term care (non RCT: Rahkonen 2001)
  - mortality (RCTs: Cole 1994; Cole 2002; Pitkala 2006; non RCT: Rahkonen 2001)

12.3 Methodological quality

12.3.1 RCTs

The method of sequence generation was adequate in two RCTs in which a computer-generated sequence was employed (Cole 2002, Pitkala 2006), and was not stated in one RCT (Cole 1994).

One RCT reported adequate allocation concealment - central randomisation with details of a retained schedule (Pitkala 2006). One RCT was partially adequate (with independent allocation but no further details, Cole 2002). In the third RCT, allocation concealment was not stated (Cole 1994).

Outcome assessors were stated to be blinded in two RCTs (Cole 1994, Cole 2002) and this was not stated in the other RCT (Pitkala 2006). Patients were not blinded in any of the RCTs.

Two RCTs reported an *a priori* sample size calculation. One RCT (Cole 1994) reported that a sample of 30 or more was required for 80% power to detect a difference of at least 1SD in the change in the measures used (p=0.05). One RCT (Pitkala 2006) reported that 58 to 91 patients per group were needed to show a 20% difference in the combined endpoint (discharge to permanent institutional care or death) with 80% power (p=0.05). The third RCT did not report a sample size calculation (Cole 2002).

All three RCTs included in the review demonstrated baseline comparability of the groups on measures such as age, gender and baseline scores measuring delirium or mental state.

All RCTs used an intention to treat analysis for at least some outcome measures. One RCT reported no missing data in either group (Pitkala 2006). In one RCT (Cole 2002), 7 patients withdrew in the intervention group (6.2%) versus 2 (1.8%) in the control group. In the third RCT (Cole 1994), 33% of patients died in the intervention group versus 37% in the control group; mean scores for some of the outcome measures SPMSQ and Crichton Geriatric Behavioural Rating Scale [CGBRS] were given for surviving patients only (i.e. fewer than 70% of the number randomised), although all patients were included in some outcome measures (length of stay, discharge to new long-term care, mortality).

Overall, one RCT was considered to have the potential for bias (Cole 1994). This study did not state randomisation or allocation concealment methods, and some outcome measures had missing data due to patients who had died (Cole 1994). This study was considered in sensitivity analyses.

12.3.2 Non-RCTs

In the Rakhonen (2001) study, the control group was formed by matching pairs of patients on age and gender from patients fulfilling the inclusion criteria from the earlier time period; in the remaining two studies patients were not individually matched but the groups were comparable on age and gender. The Milisen (2001) study reported that the non intervention cohort had significantly greater comorbid conditions (e.g. cardiac, vascular and abdominal problems).

One study reported that the investigator was blinded to the data of the main outcome measure of the study in the control patients (Rahkonen 2001: information was collected from registers for the control patients) and unclear in the other two studies.

One study (Rahkonen 2001) reported not all eligible patients were included (10%) and it was unclear in the other two studies.

Overall, we considered the three non-RCT studies to be of low quality because of the study design.

12.4 Results

12.4.1 Multicomponent intervention versus usual care

Primary outcomes of the review:

<u>Duration of delirium</u>
Only one study reported the duration of delirium (Milisen 2001). This was significantly shorter in the intervention cohort (median = 1 day, interquartile range [IQR] = 1) compared with the non-intervention cohort (median = 4 days, IQR = 5.5, p=0.03, Mann-Whitney U test).

<u>Number of patients recovered from delirium (complete response)</u>
Two RCTs (Cole 2002; Pitkala 2005) reported complete response. The Pitkala (2006) study defined the response rate as a permanent improvement of at least 4 points on the MDAS (severity of delirium scored 0 to 30, with 30 being the worst) at 8 days; although no data or references were supplied to justify the use of this score as the measure for improvement, and the GDG considered this to be a poor measure of complete response.

Cole (2002) reported the number of patients with an improvement in cognitive status, as defined by the MMSE, during the hospital stay (mean length of stay 19 days). "Improvement" was defined as an increase in MMSE of 2 or more points; with no decrease below baseline plus 2 points thereafter. If the MMSE score at baseline was 27 or more, improvement was no decrease below 27; MMSE ranges from 0=poor to 30=excellent; a score of 23 or less indicates cognitive impairment) or 'not improved'. The GDG decided that 'the number improved' was an unsatisfactory definition of recovery from delirium, so the study was not included in the analysis for this outcome.

In the Pitkala (2006) study, the intervention significantly increased the number of patients who had recovered from delirium at 8 days after admission (RR 2.00, 95% CI 1.30 to 3.08) This corresponds to a number needed to treat of 5 (95% CI 3 to 10); figure 12.1(Appendix K). The GDG debated whether a change of 4 points on the MDAS scale would clearly show improvement and considered that any conclusions drawn from the Pitkala (2006) study should be treated with caution.

Secondary outcomes of the review:

<u>Cognitive impairment</u>
Three studies (Cole 1994; Milisen 2001; Pitkala 2006) reported cognitive impairment.

The Cole (1994) study reported scores for the SPSMQ, a 10-item questionnaire that evaluates orientation, memory and concentration (0=no impairment to 10=severe impairment) at 8 weeks. There was no difference between the intervention and usual care groups (figure 12.2, Appendix K), although the result is imprecise (MD -1.10 (95% CI -4.95 to 2.75).

The Pitkala (2006) study measured cognitive impairment with the MMSE at 6 months (Pitkala 2006). The study reported a mean score of 18.4 in the intervention group versus 15.8 in the usual care group, but no standard deviations were given (p=0.047 for repeated measures analysis of variance (ANOVA); baseline scores used as covariates). This was just significant.

The Milisen (2001) study reported the mean MMSE scores for the delirious patients in the intervention group and the non intervention group (mean MMSE scores: intervention group (delirious): 15.5; non intervention group (delirious): 9.5); the study reported that although the intervention group showed a higher overall cognitive function this difference was not statistically significant; p values or standard deviations were not reported.

Length of stay

Length of hospital stay was reported by all three RCTs (Cole 1994; Cole 2002; Pitkala 2006). The result for the Pitkala (2006) study is presented as a subgroup as the intervention differed from the other two studies (Cole 1994; Cole 2002).

The Cole (1994) study did not report standard deviations, so the study's contribution to the meta-analysis of the two studies was not estimable. There was no significant difference between intervention and usual care groups in Cole (2002), although the result is imprecise [MD 0.60 (95% CI -3.90 to 5.10)] (figure 12.3, Appendix K).

In the Pitkala (2006), length of stay appeared shorter in the usual care group. We note that the distribution of lengths of stay was skewed (median 21 days in the intervention group, range 2 to 110 days; median 16 in the usual care group, range 1 to 90 days; mean 29.3 days, SD 25.6 in intervention group and mean 22.4 days, SD 18.4 in control group; means are less than twice SD so data likely to be skewed). The result is imprecise [MD 6.90 (95% CI 0.28 to 13.52)] .

Two non RCTs also reported length of hospital stay (Milisen 2001; Naughton 2005). The Milisen (2001) study reported a median of 13.5 days (IQR 3.75 days) for the intervention cohort and 14 days (IQR 5 days, p=0.6) for the non-intervention cohort. The Naughton (2005) study reported that following intervention, a mean of 3.3 days was saved in length of stay following each episode of delirium.

Discharge to long-term care

All three RCTs reported discharge of patients who had become more dependent since their admission. Two studies reported that patients were discharged at a greater level of dependency: Cole (1994) reported the percentage of patients discharged required more care (numbers were calculated as the proportion of patients remaining alive at the end of the study); Cole (2002) reported that living arrangements were arranged hierarchically from least dependent (e.g. home alone) to most dependent (e.g. nursing home); living arrangements at discharge were compared with those at admission and were rated as more dependent, same, or less dependent.

The Pitkala (2006) study reported the number of patients discharged to permanent institutional care, and these represented new admissions to such care as patients already in permanent institutional care at admission were excluded from the study.

The results are presented as subgroups in figure 12.4 (Appendix K). There was no significant difference in effect of the intervention on discharge to higher care [OR 1.04 (95% CI 0.19 to 5.65)], a more dependent living arrangement at discharge [OR 0.77 (95% CI 0.31 to 1.92)] or to new long-term care [OR 0.69 (95% CI 0.38 to 1.26)], although the results for all three studies are imprecise.

Excluding the Cole (1994) study due to its possible bias did not materially alter the results (a forest plot showing sensitivity analysis is not presented).

The Rahkonen (2001) study reported the duration of long-term care in the three years of the study. This was a mean of 441 days (SD 366) in the intervention group compared with 535 days (SD 308) in the control group [MD -94 days (95% CI -225.28 to 37.28)] (figure 12.5, Appendix K). The mean age was comparable (82.1 years in both groups) and the study excluded patients with confirmed or suspected dementia, however, individuals with mild cognitive impairment were included.

Health related quality of life (HRQoL)

One report (Pitkala 2008) of the Pitkala (2006) study reported health related quality of life along the following dimensions: mobility, vision, hearing, breathing, sleeping, eating, speech, elimination, usual activities, mental function, discomfort and symptoms, depression, distress, and vitality. Patients were assessed with the 15D questionnaire at baseline and discharge [range 0 (poor HRQoL) to 1 (excellent HRQoL)].

There was a small significantly higher HRQoL for the intervention group (MD 0.06 (95% CI 0.02 to 0.10); figure 12.6 (Appendix K). The study reported that there were significant differences for the intervention and usual care group on the following dimensions on the 15D questionnaire: mental function corresponding to cognition and alertness ($p<0.001$), usual activities corresponding to functioning in activities of daily living ($p<0.001$), vitality ($p= 0.004$), depression ($p=0.044$), and speech ($p=0.024$).

Mortality

Three RCTs (Cole 1994; Cole 2002; Pitkala 2006) and one non-RCT (Rahkonen 2001) evaluated the number of patients who died: two RCTs at 8 weeks (Cole 1994; Cole 2002) and the other RCT at 1 year (Pitkala 2006) and the non-RCT at 3 years (Rahkonen 2001).

The Cole (1994) study reported that overall 35% (31/88) patients died in 8 weeks (33% [14/42] and 37% [17/46] deaths occurring in the intervention and control groups, respectively) [OR 0.90 (95% CI 0.51 to 1.60)]; the causes of death were not given.

The Cole (2002) study reported that overall 21% (47/227) of patients died (22% [25/113] and 19% [22/114] deaths occurring in the intervention and control groups, respectively) [RR 1.15 (95% CI 0.69 to 1.91)]; and the Pitkala (2006) study reported that overall 32% (56/174) patients died over 1 year (34% [30/87] and 30% [26/87] deaths occurring in the intervention and control groups, respectively) [RR 1.15 (95% CI 0.75 to 1.78)]; the causes of death were not reported in either study.

There was no significant difference between the interventions and usual care in the mortality rates, but the results were very imprecise (figure 12.7, Appendix K).

The non-RCT study (Rahkonen2001) reported that during the three-year follow up, a total of 42% (43/102) patients died, the causes of death were not reported [RR 0.87 (95% CI 0.55 to 1.37)] (figure 12.8, Appendix K).

Overall summary

Summary of results for the multicomponent intervention is reported in table 12.4.

Table 12.4: summary of results: multicomponent non pharmacological treatment of delirium in hospital setting.

	Complete response (RR (95% CI))	Duration of delirium Median (IQR)	Cognitive impairment	Length of stay MD (95% CI) unless otherwise stated	Discharge OR (95% CI) unless otherwise stated	HRQoL MD (95% CI)	Mortality RR (95% CI)
Nursing intervention protocol (Cole 1994)			MD -1.10 (95% CI -4.95 to 2.75)	Mean number of days: 25.3 vs 22.7 for intervention vs usual care groups respectively; SD or p alues not reported	Discharge to greater care: 1.04 (0.19 to 5.65)		0.90 (0.51 to 1.60)
Nursing intervention protocol + follow up by consultant (Cole 2002)	'The number improved' deemed unsatisfactory definition of recovery from delirium			0.60 (-3.90 to 5.10)	Discharge to a more dependent living arrangement: 0.77 (0.31 to 1.92)		1.15 (0.69 to 1.91)
Nurse-led interdisciplinary intervention (Milisen 2001)		1 day (1) vs 4 days(5.5) for intervention vs usual care groups respectively; p=0.03	Mean score on MMSE: 15.5 vs 9.5 for intervention vs usual care groups respectively; SD or p values not reported	Median (IQR): 13.5 days (3.75) vs 14 days (5) for intervention vs usual care groups respectively; p=0.06	Discharge to a new long-term care: 0.69 (0.38 to 1.26)		1.15 (0.75 to 1.78)
Education and management intervention (Naughton 2002)				Following intervention a mean o f 3.3 days saved in length of stay following each episode of delirium			
Multicomponent geriatric intervention (Pitkala 2005)	2.00 (1.30 to 3.08)		Mean score on MMSE: 18.4 vs 15.8 for intervention vs usual care groups respectively; p=0.047	6.90 (0.28 to 13.52)		0.06 (0.02 to 0.10)	
Systematic intervention (Rahkonen 2001)					Duration of long-term care in 3 years of the study: MD -		3 year follow up: 0.87 (0.55 to 1.37)

	Complete response (RR (95% CI))	Duration of delirium Median (IQR)	Cognitive impairment	Length of stay MD (95% CI) unless otherwise stated	Discharge OR (95% CI) unless otherwise stated	HRQoL MD (95% CI)	Mortality RR (95% CI)
					94 days (-222.28 to 37.28)		

12.5 Health economic evidence

12.5.1 Multicomponent interventions for the treatment of delirium in a hospital setting

One economic evaluation study was included as evidence (Pitkala 2008). This was a Finnish RCT of 174 consecutive delirium patients aged above 69 years who were admitted to the general medicine services and whose life expectancy was predicted to be above 6 months. The study aimed at assessing the effects of multicomponent geriatric treatment on costs of care and HRQoL in delirious in-patients. Patients in the intervention group received a comprehensive geriatric assessment at baseline for good detection of delirium, as well as careful diagnosis of the underlying etiological conditions. They received atypical antipsychotics if necessary and effective general treatments were implemented for all patients. After the acute phase of delirium, all patients not recovering from impaired cognition underwent detailed diagnostics for dementia and thereafter, received acetyl cholinesterase inhibitors. Patients in the comparator arm received usual care and this was not exactly described.

The average cost per patient in the intervention arm was €19,737 while the average cost per patient in the usual care arm was €19,557.The extra cost attributable to intervention was €446 per patient. This included the cost of atypical antipsychotics, acetycholinesterase inhibitors, vitamin D-calcium supplements, hip protectors, and nutritional supplements. Average unit costs in Finland were used. Health related quality of life was measured using the 15D questionnaire but the question on sexual activity was omitted. Subjective health was assessed using an ordinal scale at discharge. An unadjusted mortality rate of 35% and 30% were reported in the intervention and usual care groups respectively. The patient's measure of health status was 0.68 and 0.62 in the intervention and control groups respectively. The dimensions of HRQoL showing significant differences favouring intervention were mental function, usual activities, vitality, depression and speech.

The results of this study could be used to estimate the cost per unit of improvement in health status of delirium patients. However, patient's measure of health status was based on 15D which elicited health status scores from a Finnish general population. It was reported only at the point of discharge from hospitalisation for delirium and quality-adjusted life years were not reported. Furthermore, there was no sensitivity analysis to test the effect of the uncertainties surrounding the cost and health outcome measures. Costs were not assessed from a UK NHS and PSS perspective. The results of this study were judged to be not directly applicable to this guideline.

12.6 Clinical evidence statements

- There is very low quality evidence which showed that a multicomponent intervention targeting six modifiable risk factors (orientation, sleep, sensory impairment improvement, early mobilisation, environmental, medication) following a consultation with a geriatrician or geriatric psychiatrist and follow up by a liaison nurse showed no significant difference in:
    - cognitive impairment (measured at 8 weeks). However, there is much uncertainty around this result
    - the number of patients discharged with a greater level of dependency; there is much uncertainty around this result
    - mortality rates at 8 weeks; there is uncertainty around this result

Additional follow up assessment by geriatirican and a liaison nurse showed no significant difference.

- There is very low quality evidence to show that a multicomponent intervention targeting three modifiable risk factors (dehydration/nutrition, pain management, medication management) with training showed:
    - significantly shorter duration of delirium in patients in the intervention group
    - no significant difference in the median length of stay in hospital

- There is moderate quality evidence to show that a multicomponent geriatric intervention based on targeting four modifiable risk factors (orientation, dehydration/nutrition, early mobilisation, medication management) with comprehensive geriatric assessment showed a:
    - significant number of patients recovered from delirium (at 8 days) in the intervention group; however, there is much uncertainty around this result
    - significant difference showing a decreased length of stay in the usual care group;
    - small significant improvement in the health related quality of life (mental function, daily functioning, depression, vitality, and speech) for the intervention group at discharge

- borderline significant difference showing a lower level of cognitive impairment at 6 months for the intervention group
- non significant difference in the number of patients discharged to long-term care; there is much uncertainty around this result.

- There is very low quality evidence to show a multicomponent intervention targeting two modifiable risk factors (orientation, early mobilisation) with training, continuous nursing support and annual one-week visits to a rehabilitation unit showed no significant difference in:
  - long-term care stay over the duration of the study (3 years); there is much uncertainty around this result
  - mortality rates at 3 years

12.7 From evidence to recommendations

The evidence suggested that enhanced treatment strategies for people with delirium are more effective than usual care however the GDG did not feel confident in recommending a particular multicomponent intervention because of the low quality evidence. Instead the GDG drew on the principles of the multicomponent interventions and their clinical expertise to inform the recommendations. The GDG agreed treatment of delirium should comprise the following:

- initial management for all people with delirium,
- second line management for those who are distressed or are considered a risk to themselves or others,
- management for people whose symptoms do not resolve, either following initial or second line management.

Initial management

The multicomponent treatment review showed some indication of clinical effectiveness in one study (Pitkala 2006). The GDG considered the measure of delirium to be too unreliable to support this in economic modeling or recommendations. The GDG did draw on the components comprising the multicomponent interventions, and used them, together with information from the risk factors review to make a consensus recommendation on treating possible underlying causes of delirium (recommendation 1.6.1). The GDG recognised that sometimes there was more than one underlying cause.

The GDG also considered evidence from the non-pharmacological risk factors review and the patient information review, and drew on their clinical experience.

The GDG recognised the importance of listening and talking to the person experiencing delirium. The GDG specifically took into account the messages conveyed by the GDG patient representatives describing how difficult it was for them to tell relatives and staff about their changes in cognition.

Evidence from the multicomponent treatment review and the GDG's clinical expertise highlighted the importance of reorientation in people diagnosed with delirium. The GDG felt that reorientation could be addressed by communicating the role of the healthcare professional, who the person is and the day, date, time and place. Familiar faces of family, friends and carers may also help with orientation (evidence underpinning this came from the patient information review, chapter 15).

Hospital environments, artificial lighting and time loss through disturbed sleep patterns and periods of unconsciousness can easily lead to disorientation with the potential to aggravate delirium. The GDG therefore also considered that an important part of reorientation included maintaining a suitable care environment (recommendation 1.3.1) for people diagnosed with delirium, and therefore included this within the recommendation.

Recommendation 1.6.2 should be carried out for all people diagnosed with delirium.

Distressed people

The GDG discussed the care of people who are distressed or those considered at a risk to themselves or others and this was informed by the NICE Violence guideline (clinical guideline25) which provides information on how to calm down an escalating situation. The GDG recognised that the NICE violence guideline was restricted to short-term management of violent and disturbed behaviour in psychiatric settings and emergency departments; however, they agreed that the principles of effective communication and de-escalation techniques could be extrapolated to this guideline. The GDG considered that non-pharmacological de-escalation approaches should be tried before resorting to pharmacological treatment (recommendation 1.6.3). This was partly on the basis of their clinical experience and partly in view of their reservations about the evidence on medication (section 13.8). The GDG also noted that identifying distress in people who had hypoactive delirium can be more difficult than in people with hyperactive delirium. Although they often appear to be calm, they may be distressed by psychotic symptoms and this may not be intuitive. The GDG decided to add a statement to the recommendation to this effect.

The GDG recommended that when de-escalation techniques had not worked pharmacological interventions should be considered (see recommendation 1.6.4 in the pharmacological treatment chapter 13).

The GDG also made a recommendation for people whose symptoms remain unresolved following first or second line treatment. Persisting delirium could be due to underlying causes remaining unaddressed. Alternatively the person might have dementia rather than delirium. The GDG made a recommendation to capture this and cross-referred to the 'Dementia' guideline (NICE clinical guideline 42) for advice on diagnosing and managing dementia (see recommendation 1.6.6 in the pharmacological treatment chapter 13). This recommendation was consistent with the GDG's strategy for patients in whom it was difficult to distinguish between delirium, dementia and delirium on dementia. Failure to resolve delirium should lead the health care professional to consider dementia.

The GDG agreed that there was limited evidence relating to health care professional effective communication skills when caring for people with delirium. In addition, the recognition and recording of delirium are important factors. The GDG agreed a future research recommendation for staff education which include effective communication strategies (see below and Appendix H):

Future research recommendation:

Does an education programme for staff improve the recovery from delirium in patients in hospital compared with an education leaflet or usual care?

12.8 Recommendations

Initial management

In people diagnosed with delirium, identify and manage the possible underlying cause or combination of causes. **[1.6.1]**

Ensure effective communication and reorientation (for example explaining where the person is, who they are, and what your role is) and provide reassurance for people diagnosed with delirium. Consider involving family, friends and carers to help with this. Provide a suitable care environment (see recommendation 1.3.1). **[1.6.2]**

Distressed people

If a person with delirium is distressed or considered a risk to themselves or others, first use verbal and non-verbal techniques to de-escalate the situation. For more information on de-escalation techniques, see 'Violence' (NICE clinical guideline 25). Distress may be less evident in people with hypoactive delirium, who can still become distressed by, for example, psychotic symptoms. **[1.6.3]**

13 Treatment of delirium: pharmacological

CLINICAL QUESTIONS:

What are the most clinical and cost effective and safe pharmacological interventions for treating people with delirium in hospital?

What are the most clinical and cost effective and safe pharmacological interventions for treating people with delirium in long-term care?

13.1 Clinical introduction

Delirium is characterised by a range of symptoms that can cause distress, behaviour disturbance and place people at risk. Medications are used in clinical practice to manage these symptoms though the evidence base remains limited. Pharmacological agents that alter the course of delirium or control particular symptoms will need to demonstrate safety as well as effectiveness but would be a valuable development in treatment.

The pathophysiology of delirium is complex and people with delirium may have serious physical illness that complicates the use of drug treatment. Should drugs be given routinely or for selected symptoms? If selected symptoms then for which symptoms? Does the clinical context alter decisions about drug treatments? Would all people receive them or those at risk? These are questions for which answers are needed.

13.2 Description of studies

Details of included and excluded papers together with study design are reported in table 13.1.

Table 13.1: study inclusion, exclusion and design

Papers	Comments	Study
N= 23 evaluated for inclusion		
N= 18 excluded	Reasons for exclusion are reported in Appendix G.	
N= 1 identified in	Cochrane review identified and updated	Lonergen 2009;

update searches		
N= 5 included	Study designs 2 RCTs	Hu 2006; Lee 2005
	1 quasi randomised	Skrobik 2004
	Non randomised	Liu 2004; Miyaji 2007
	1 additional Cochrane review was identified and updated within this review	Overshott 2008

The Lonegren (2009) review examined the use of antipsychotics for delirium and identified three studies (Han2004; Hu 2004; Kalisvaart 2005) and the Oveshott (2008) review examined cholinesterase inhibitors for delirium and identified one study (Liptzin 2005).

One study identified in the Cochrane reviews, which did not meet our inclusion criteria was excluded (Han 2004). Of the remaining studies, studies examining pharmacological prevention of delirium (Kalisvaart 2005; Liptzin 2005) have been reported in Chapter 11 and one study (Hu 2004) relevant to treatment of delirium has been reported in this chapter.

Two non-randomised comparative studies (Liu 2004; Miyaji 2007) comparing typical and atypical antipsychotics were also included initially, because their comparator for haloperidol was risperidone, rather than olanzapine (which was used in the RCTs). Both had retrospective comparative designs, in which patients were selected from records). In the Liu (2004) study, patients were treated at the clinician's choice; in the other (Miyaji 2007), allocation was presumed to be by clinician choice but this was not stated. In the Liu (2004) study, there was a large difference in age between the risperidone and haloperidol groups (risperidone 68 years, range 40–85 years; haloperidol 50 years, range 15–77 years). In the Miyaji (2007) study, the participants in the injection haloperidol group were significantly younger than those in the other two groups (median 69 years versus 73 years).

In view of these methodological limitations, the GDG decided to exclude these two studies from the review, and to rely on the class effect for the comparison between typical and atypical antipsychotics. Therefore these two non-randomised studies were not considered further except for the adverse effects review (Chapter 14).

Thus the efficacy review focuses on three studies (Hu 2006; Lee 2005; Skrobik 2004).

13.2.1 Study Design

None of the studies were conducted in the UK. Information on study sizes, geographical location and funding are described in table 13.2.

Table 13.2: study characteristics

Reference/Study	N	Geographical location	Funding
Hu 2006	180	China	Not stated
Lee 2005	40	Korea	Not stated
Skrobik 2004	77	Canada	Eli Lilly

13.2.2 Population

One study (Skrobik 2004) was in an ICU, in which the patients were mostly surgical (48 elective operations; 21 urgent operations; 4 medical patients), and patients were treated within 2 hours of the diagnosis of delirium.

The two other studies had patients in a non-ICU hospital setting. In the Hu (2006) study, the type of ward was not stated, but the patients had 'senile delirium' due to metabolic (n = 68), toxic (n = 47), structural (n = 25) or infectious (n = 35) causes; the duration of delirium was reported to be between 30 minutes and 17 days. In the Lee (2005) study, patients had been referred to a psychiatric consultation service from departments of neurosurgery, internal medicine, neurology and rehabilitation medicine: those who had immediately recovered from a major operation were excluded.

Different methods were used to diagnose delirium, however, all the studies used the DSM-IV criteria in some form: in the ICU study (Skrobik 2004), patients were screened using the ICU-Delirium Screening Checklist (ICU-DSC), then if they scored 4 or more (or had a clinical diagnosis of delirium); this was confirmed using DSM-IV criteria. In the Hu (2006) study, patients were assessed using the DSM-IV criteria. They also had to have a total score on the Delirium Rating Scale (DRS) of 12 or more, and a clinical global impression scale: severity of illness (CGI-SI) score of 4 or more. In the Lee (2005) study, patients meeting the criteria for delirium were diagnosed using the DSM-IV criteria and evaluated using the Delirium Rating Scale-Revised-98 (DRS-R-98). This includes a 16-item scale to diagnose delirium and the 13-item severity subscale.

None of the studies reported whether the patients had dementia or cognitive impairment, although the Lee (2005) study excluded patients who had a previous history of a 'psychiatric disorder'.

The age range across the studies was 42 to 99 years, with the mean age ranging from 61 to 74 years. All studies included men and women. Ethnicity was not reported.

13.2.3 Interventions

The three included studies investigated the following drugs: typical antipsychotics (haloperidol) and atypical antipsychotics (amisulpride, olanzapine, and quetiapine) in the treatment of delirium in a hospital setting. The interventions were:

- Haloperidol

  - orally or by enteral tube: given within 2 h of the diagnosis of delirium, initially 2.5–5 mg every 8 hours (patients over 60 years 0.5–1 mg) then titrated based on clinical judgement for up to 5 days (Skrobik 2004)
  - intramuscular injection 2.5–10 mg per day, depending on response; the effect was observed for one week; delirium had occurred from 30 min to 17 days (Hu 2006)
- Olanzapine
  - orally or by enteral tube: given within 2 h of the diagnosis of delirium, initially 5 mg per day (patients over 60 years 2.5 mg) then titrated based on clinical judgement for up to 5 days (Skrobik 2004)
  - orally or sublingually initial dose 1.25–2.5 mg then adjusted, depending on response, to 1.25–20 mg per day; the effect was observed for one week; delirium had occurred from 30 min to 17 days (Hu 2006)
- Amisulpride
  - 50–800 mg/day (initial dose mean: 156.4 mg/day (SD 97.5)); the dose was flexible according to clinicians preferences and experience; it was unclear when the drug was administered following the diagnosis of delirium; treatment was administered until the CGI score reached 2 or less; mean duration of stabilisation was 6.3 (SD 4.4) days (Lee 2005)
- Quetiapine
  - 50–300 mg/day (initial dose mean: 113 mg/day (SD 85.5)); the dose was flexible according to clinicians preferences and experience; it was unclear when the drug was administered following the diagnosis of delirium; treatment was administered until the CGI score reached 2 or less; mean duration of stabilisation was 7.4 (SD 4.1) days (Lee 2005)

13.2.4 Comparisons

The following comparisons were carried out:

- Typical antipsychotic (haloperidol) versus no treatment (Hu 2006)
  - all patients also had ‘somatic treatment aiming at delirium’
- Atypical antipsychotic (olanzapine) versus no treatment (Hu 2006)
  - all patients also had ‘somatic treatment aiming at delirium’
- Comparison of two drugs in the same class (atypical antipsychotics)

    - Amisulpride versus Quetiapine (Lee 2005)

- Comparison of two drug classes
    - Typical antipsychotic (haloperidol) versus atypical antipsychotic (olanzapine) (Hu 2006; Skrobik 2004)
    - all patients in Hu (2006) also had 'somatic treatment aiming at delirium'

One study (Skrobik 2004) reported that the patients received concurrent benzodiazepines and fentanyl for analgesia; some patients also received other sedatives; there was no significant difference between interventions for these concurrent drugs or in the amount of rescue IV haloperidol used. The Hu (2006) study reported that all patients received 'somatic treatment for delirium'; and the Lee (2005) study reported that other antipsychotics or benzodiazepines were not allowed.

13.3 Methodological quality

13.3.1 Randomised and quasi-randomised studies

The method of sequence generation was inadequate in the quasi-randomised study (Skrobik 2004), in which the patients were allocated on an even/odd day basis, and allocation concealment was also judged inadequate because the sequence was likely to be known in advance. The methods of sequence generation and allocation concealment were not stated in either of the two RCTs (Hu 2006; Lee 2005).

In the Skrobik (2004) study, outcomes were assessed by a clinician or research nurse blinded to the allocation; it was unclear whether patients were blinded, but this was unlikely because the frequency of dosing was different. In the other two studies, it was unclear whether assessors were blinded, and it was also unclear if the patients in the Lee (2005) study were blinded. In Hu (2006) it was unlikely that the patients were blinded because of the nature of the interventions (no placebos and different routes of administration for the active drugs).

None of the studies reported an *a priori* sample size calculation.

In the Skrobik (2004) study, patients were comparable on gender, weight and APACHE score, but those on haloperidol were significantly younger. In the Lee (2005) study, there were no significant differences between the groups on age, gender, baseline DRS-R-98 and CGI scores. In the Hu (2006) study, there were no significant differences between the groups on age, gender, pre-treatment severity of mental symptoms or causes of delirium.

Two studies had less than 20% missing data in either group (Hu 2006; Skrobik 2004). One study (Lee 2005) had more than 20% missing data: 5/20 (25%)

dropped out from the quietiapine group and 4/20 (20%) from the amisulpride group; only patients who completed the study were included in the analysis. In the Skrobik (2004) study, patients were analysed according to their allocation group; and the Hu (2006) study, carried out an available case analysis.

All the studies used an adequate method of delirium assessment at baseline [DSM-IV; screened with ICU-DSC and diagnosis confirmed with DSM-IV (Skrobik 2004)] and used an adequate method of assessment to evaluate delirium following treatment (Hu 2006: DRS; Lee 2005: DRS-R-98, administered by a trained psychologist; Skrobik 2004: Delirium Index (DI) scale administered by a trained clinician).

Two studies (Hu 2006; Lee 2005) also used the CGI scale to evaluate treatment effects. The GDG noted that the CGI scale is not a direct measure of delirium and needs to be interpreted accordingly.

Overall, the Skrobik (2004) study was considered to be at high risk of bias because there was inadequate allocation concealment, the patients were not blinded and there was a significant difference in patient age. In addition, the patients received rescue medication which may have confounded the outcome measures. The other two studies also had some potential for bias because the patients were unlikely to be blinded (Hu 2006), and because of more than 20% missing data in one group (Lee 2005).

13.4 Results

A. Typical antipsychotics versus placebo/no treatment

One RCT (Hu 2006) compared a typical antipsychotic (haloperidol) versus a no treatment control.

13.4.1 Primary Outcomes

Complete response

One study Hu (2006) in 101 patients reported the measure of recovery from delirium as 'symptoms alleviated or disappeared completely' on the global improvement item of the CGI (CGI-GI) at 7 days. The analysis showed a significant improvement of delirium in the haloperidol group compared to the control group, although the result is imprecise (figure 13.1, Appendix K); RR 3.95 (95% CI 1.75 to 8.90). This corresponds to an NNT of 2 (95% CI 2 to 3) for a control group rate of 17%.

Duration of delirium

The Hu (2006) study reported the 'time to take effect', the mean number of days for the drug to take into effect, in responders only. The GDG considered that these results were potentially biased and did not consider 'time to take effect' was an adequate proxy/surrogate outcome for duration of delirium. Therefore the results are not reported.

13.4.2 Secondary Outcomes

Severity of delirium

The Hu (2006) study reported scores on the DRS (0 to 32 scale) following treatment. The severity of delirium assessed at the seventh day of treatment was significantly lower in the haloperidol group; MD: −10.40 (95% CI −13.95 to −6.85) for a control group severity score of 17.6 (figure 13.2, Appendix K).

This study also reported the scores on the CGI-SI. These were 1.79 (SD 1.12) for haloperidol and 3.97 (SD 1.76) for the control group. The GDG stated this scale is not a direct measure of delirium and should be interpreted accordingly.

Please refer to table 13.3 for the GRADE evidence summary for typical antipsychotics.

Table 13.3: GRADE evidence summary: Typical antipsychotics vs placebo /no treatment

Outcome/ Author (year)	*analysis details*	*Summary statistics*	*Comments*	*GRADE details*	*GRADE Comments*
Complete response (Hu 2006)	1trial; 101 patients; from RCT	RR=3.95 (95%CI 1.75, 8.9)	Statistically significant improvement of delirium in the haloperidol group on clinical global impression scale at 7 days	• Study quality: no serious limitations • Directness: Indirect outcome - delirium assessment method • Imprecision: Number of events < 300 • Inconsistency: consistent • Reporting bias: Adequate	It is unlikely pts blinded because of nature of the intervention (IM vs control); Clinical global impression scale- indirect method of assessment of delirium; Both groups received somatic treatment aiming at delirium. Large effect
GRADE evidence rating: Moderate					
Duration of delirium (Hu 20060	1trial; 101 patients; from RCT	MD=-1.78 (95%CI -2.86, -0.7)	Statistically significant shorter duration for the haloperidol group	• Study quality:serious limitation some confounding • Directness: Direct • Imprecision: Wide CI • Inconsistency: consistent • Reporting bias: Adequate	Reported as 'time to take effect'. Duration of delirium was given for responders so potentially biased
GRADE evidence rating: Very low					
Severity of delirium (Hu 2006)	1trial; 101 patients; from RCT	MD=-10.4 (95%CI -13.95, -6.85)	Statistically significant: severity lower in the haloperidol group on the DRS (0-32)	• Study quality: serious limitation not blinded • Directness: Direct • Imprecision: Number of patients < 400 • Inconsistency: consistent • Reporting bias: Adequate	DRS scale 0-32; MID (20% = 6.4), i.e. CI precise, but fairly small number patients. Patients not blinded. Large effect
GRADE evidence rating: Moderate					

B) Atypical antipsychotics versus placebo/no treatment

One RCT (Hu 2006) compared an atypical antipsychotic (olanzapine) versus a no treatment control.

13.4.3 Primary outcomes:

Recovery from delirium (complete response); figure 14.3

One study Hu (2006) with 103 patients reported the 'symptoms alleviated or disappeared completely' on the global improvement item of the CGI-GI scale at 7 days.

The analysis (figure 13.3) showed a significant improvement of delirium in the olanzapine group compared to the control group, but the result is imprecise; RR 3.68 (95% CI 1.63 to 8.33) (figure 13.3, Appendix K). This corresponds to an NNT of 3 (95% CI 2 to 4) for a control group rate of 17%.

Duration of delirium

The Hu (2006) study reported the 'time to take effect', in responders only, but again this outcome was considered to be biased and are not reported here.

13.4.4 Secondary outcomes

Severity of delirium

One study (Hu 2006) in 103 patients reported scores on the DRS (0 to 32 scale). There was a large significant difference between the treatments on this measure; mean difference: −11.10 (95% CI −7.69 to −14.51) for a control group severity score of 17.6 (figure 13.4, Appendix K).

This study also reported the scores on the CGI-SI. These were 2.05 (SD 0.99) for olanzapine and 3.97 (SD 1.76) for the control group. The GDG stated this scale is not a direct measure of delirium and should be interpreted accordingly.

Please refer to table 13.4 for the GRADE evidence summary for typical antipsychotics.

1

Table 13.4: GRADE evidence summary: Atypical antipsychotics vs placebo / no treatment

Outcome/ Author (year)	***Meta-analysis details***	***Summary statistics***	***Comments***	***GRADE details***	***GRADE Comments***
Complete response (Hu 2006)	1trial; 103 patients; from RCT	RR=3.68 (95%CI 1.63, 8.33)	Significant difference in favour of the olanzapine group	• Study quality:serious limitation not blinded • Directness: Indirect assessment method • Imprecision: Number of events < 300 • Inconsistency: consistent • Reporting bias: Adequate	Measured on clinical impression scale (GDG considered this indirect). All received "somatic treatment aiming at delirium". Patients not blinded; Large effect
GRADE evidence rating: Moderate					
Duration of delirium (Hu 2006)	1trial; 103 patients; from RCT	MD=-2.4 (95%CI -3.51, -1.29)	Statistically significant in favour of the olanzapine group	• Study quality:serious limitation some confounding • Directness: Direct • Imprecision: CI crosses threshold • Inconsistency: consistent • Reporting bias: Adequate	'Time to take effect' only given for responders -. likely to be confounded; All patients received somatic treatment aiming at delirium
GRADE evidence rating: Very Low					
Severity of delirium from RCT	1trial; 103 patients;	MD=-11.1 (95%CI -14.51, -7.69)	Statistically significant difference on the DRS (0-32); some uncertainty	• Study quality:serious limitation not blinded • Directness: Direct • Imprecision: Number of patients < 400 • Inconsistency: consistent • Reporting bias: Adequate	All patients received somatic treatment aiming at precise in terms of GRADE Patients not blinded; Large effect
GRADE evidence rating: Moderate					

C) Atypical antipsychotic 1 versus atypical antipsychotic 2

13.4.5 Amisulpride versus Quetiapine

One study (Lee 2005) compared two atypical antipsychotic drugs. It is noted that the study size was very small (40 patients randomised, but only 31 analysed) and that this study cannot be expected to determine a difference between two active interventions.

13.4.6 Primary outcome

The Lee (2005) study did not report results for the primary outcome complete response.

13.4.6.1 Duration of delirium

One study (Lee 2005) with 31 patients reported the mean 'duration of stabilisation', which was the time for the patients to reach recovery from delirium; there was no significant difference between groups; but the result is imprecise; MD: −1.10 days (95%CI −4.09 to 1.89), for a duration of 7.4 days for the quetiapine group (figure 13.5, Appendix K).

13.4.7 Secondary outcomes

Severity of delirium

One study (Lee 2005) with 31 patients reported scores on the DRS-R-98 (0 to 39 scale); there was no significant difference between the treatments; mean difference 0.00 (95%CI −1.48 to 1.48) for a severity score of 3.5 in the quetiapine group (figure 13.6, Appendix K).

Adverse effects

The Lee (2005) study reported that there were no serious adverse events observed, such as acute dystonia and dyskinesia, but there were very few patients in this study.

Please refer to table 13.5 for the GRADE evidence summary for comparison of two types of atypical antipsychotics.

Table 13.5: GRADE evidence summary: atypical antipsychotic1 vs atypical antipsychotic 2

Outcome/ Author (year)	*Meta-analysis details*	*Summary statistics*	*Comments*	*GRADE details*	*GRADE Comments*
Complete response (Lee 2005)	1trial; 31 patients; from RCT	No results for this outcome		• Study quality: serious limitation-incomplete follow up • Directness: Direct • Imprecision: n/a • Inconsistency: n/a • Reporting bias: Adequate	Very small study; No results reported.
Duration of delirium (Lee 2005)	1trial; 31 patients; from RCT	MD=-1.1 (95%CI -4.09, 1.89)	No significant difference between amisulpride and quetiapine groups	• Study quality: serious limitation-incomplete follow up • Directness: Direct • Imprecision: CI crosses threshold • Inconsistency: consistent • Reporting bias: Adequate	Lower CI crosses 4x MID
GRADE evidence rating: Very low					
Severity of delirium (Lee 2005)	1trial; 31 patients; from RCT	MD=0 (95%CI -1.48, 1.48)	No significant difference on the DRS-R-98 (0-39) between amisulpride and quetiapine groups	• Study quality: serious limitation-incomplete follow up • Directness: Direct • Imprecision: Number of patients < 400 • Inconsistency: consistent • Reporting bias: Adequate	Very small study; 25% missing data in 1 arm
GRADE evidence rating: Low					
Adverse events (Lee 2005)	1trial; 31 patients; from RCT	RR=1	No significant adverse events reported such as acute dystonia dyskinisea	• Study quality: serious limitation-incomplete follow up • Directness: Direct • Imprecision: Number of patients < 400 • Inconsistency: consistent • Reporting bias: Adequate	Very small study; 25% missing data in 1 arm
GRADE evidence rating: Low					

D) Typical antipsychotics versions atypical antipsychotics

One RCT (Hu 2006) and one quasi randomised study compared a typical antipsychotic (haloperidol) versus an atypical antipsychotic (olanzapine).

13.4.8 Primary outcomes

Complete response

Both randomised/quasi-randomised studies evaluated a measure of recovery from delirium, although these were reported differently and neither constituted a direct outcome measure (Hu 2006; Skrobik 2004).

Hu (2006) reported the 'symptoms alleviated or disappeared completely on the global improvement item of the clinical global impression scale' at 7 days.

Skrobik (2004) reported the numbers of patients requiring rescue IV haloperidol on day 1 (19/45 patients on haloperidol and 10/28 on olanzapine) and the numbers for subsequent days (4/45 haloperidol and 1/28 olanzapine). This was converted into the numbers not requiring rescue medication (by subtraction), i.e. 22/45 on haloperidol and 17/28 on olanzapine. This was assumed to be an approximation to a complete response to study treatment.

Meta-analysis of the two studies in 219 patients did not demonstrate a significant difference between the treatments (figure 13.7, Appendix K); RR 0.99 (95% CI 0.80 to1.21). There was insignificant heterogeneity between studies ($I^2 = 27\%$; $p = 0.24$). In the absence of the Skrobik (2004) study, which was at much higher risk of bias, there was no significant difference between interventions; RR 1.07 (95%CI 0.85 to 1.35).

Duration of delirium

The Hu (2006) study reported the ‘time to take effect’, in responders only, but again this outcome was considered to be biased and are not reported here.

13.4.9 Secondary outcomes

Severity of delirium

One study (Hu 2006) reported scores on the DRS (0 to 32 scale; figure 13.8, Appendix K), which showed no significant difference between the treatments on this measure; mean difference 0.70 (95% CI –0.45 to 1.85) for a control group severity of 6.5 units.

This study also reported the scores on the CGI-SI. These were 1.79 (SD 1.12) for haloperidol and 2.05 (SD 0.99) for olanzapine. The GDG considered the CGI scale was likely to be less specific for measuring delirium symptoms than the DRS.

The quasi-randomised study (Skrobik 2004) reported the mean daily delirium index scores on a graph. The mean daily delirium index scores at day 5 were haloperidol 4.85, olanzapine 5.40, mean difference 0.55; there was no significant difference between interventions (p=0.83). It is noted that these data were likely to be confounded by the use of rescue IV haloperidol medication, predominantly on the first day in around a third of the patients in each group.

Adverse events

The Skrobik (2004) study reported 13% (6/45) patients receiving haloperidol were noted to have low scores on extrapyramidal symptom testing and no extrapyramidal symptoms or other adverse effects were reported in patients receiving olanzapine. There was no significant difference between the interventions [RR 8.20 (95% CI 0.48, 140.09)]; the confidence interval is wide. (figure 13.9, Appendix K)

Please refer to table 13.6: for the GRADE evidence summary for comparison of typical antipsychotics with atypical antipsychotics.

Table 13.6: GRADE Evidence summary: Cross review - typical antipsychotic vs atypical antipsychotic

Outcome/ Author (year)	*Meta-analysis details*	*Summary statistics*	*Comments*	*GRADE details*	*GRADE Comments*
Complete response (Hu 2006; Skrobik 2004)	2 trials; 219 patients; from Meta analysis of RCTs	RR=0.99 (95%CI 0.8, 1.21); p=0.24; I^2=27%	No significant difference between haloperidol and olanzapine groups	• Study quality: serious limitations not blinded • Directness: Indirect outcome - delirium assessment method • Imprecision: Number of events < 300 • Inconsistency: consistent • Reporting bias: Adequate	Haloperidol vs olanzapine. One study [32.4% weight] inadeq -uate sequence generation & allocation concealment, funding, and outcome possibly inadequate. Patients unblinded in major study and indirect outcome measure
GRADE evidence rating: Low					
Duration of delirium (Hu 2006)	1trial; 146 patients; from RCT	MD=0.62 (95%CI 0.06, 1.18)	Significantly shorter time to take effect for the olanzapine group compared to the haloperidol group	• Study quality: Very serious limitations • Directness: Direct • Imprecision: Number of patients < 400 • Inconsistency: consistent • Reporting bias: Adequate	Reported as 'time to take effect' in responders only - likely to be biased
GRADE evidence rating: Very low					
Severity of delirium (Hu 2006)	1trial; 146 patients; from RCT	MD=0.7 (95%CI -0.45, 1.85)	No significant difference between the haloperidol and the olanzapine groups on the DRS (0-32)	• Study quality: serious limitations not blinded • Directness: Direct • Imprecision: Number of patients < 400 • Inconsistency: consistent • Reporting bias: Adequate	All patients received somatic treatment aiming at delirium; DRS scale 0-32, narrow CI, but fairly small trial. Patients not blinded.
GRADE evidence rating: Moderate					
Adverse event (extrapyramidal) (Skrobik 2004)	1trial; 73 patients; from Quasi RCT	RR=8.2 (95%CI 0.48, 140.09)	No significant difference	• Study quality: Very serious limitations • Directness: Direct • Imprecision: Wide CI • Inconsistency: consistent • Reporting bias: Adequate	Haloperidol vs olanzapine; quasi randomised design; wide CI. Adverse events carefully recorded; not blinded
GRADE evidence rating: Very low					

Overall summary

Summary of results for pharmacological treatment of delirium in hospital setting are presented in table 13.7.

Table 13.7: summary of results: pharmacological treatment of delirium in hospital setting

Outomes	Typical antipsychotics vs placebo	Atypical antipsychotics vs placebo	Atypical antipsychotic 1 vs Atypical antipsychotic 2	Typical antipsychotic vs atypical antipsychotic
Complete response RR (95% CI)	3.95 (1.75 to 8.90)	3.68 (1.63 to 8.33)	No results reported for this outcome	0.99 (0.80 to 1.21)
Duration of delirium MD (95% CI)	Reported 'time to take effect'. GDG considered these results were biased and time to take effect was not an adequate surrogate for duration of delirium.	Reported 'time to take effect'. GDG considered these results were biased and time to take effect was not an adequate surrogate for duration of delirium.	-1.10 (-4.09 to 1.89)	Reported 'time to take effect'. GDG considered these results were biased and time to take effect was not an adequate surrogate for duration of delirium.
Severity of delirium MD (95% CI)	-10.40 (-13.96 to -6.85)	-11.10 (-14.51 TO -7.69)	0.00 (-1.48 to 1.48)	0.70 (-0.45 to 1.85)
Adverse events RR (95% CI)				8.20 (0.48 to 1.40.09)

13.5 Health economic evidence

The health economic model (chapter 16) assessed the cost-effectiveness of haloperidol and olanzapine in the hospital setting.

13.6 Clinical evidence statements

13.6.1 Typical antipsychotics versus placebo/no treatment

- There is moderate quality evidence from one RCT showing a:
    - significant improvement in delirium and a significantly lower severity of delirium (an indirect measure of delirium was used) in the haloperidol group compared with no treatment at 7 days. There is some uncertainty around this result.

13.6.2 Atypical antipsychotics versus placebo/no treatment

- There is moderate quality evidence from one RCT showing a:
  - significant recovery from delirium in favour of the olanzapine group compared with no treatment at 7 days. There is much uncertainty with this result.
  - significantly lower severity of delirium in the olanzapine group compared with no treatment (an indirect measure of delirium was used).

13.6.3 Comparison of two atypical antipsychotics

There is very low quality evidence and low quality evidence from one RCT showing no significant difference in the duration of delirium (there is some uncertainty with this result) and severity of delirium, respectively between amisulpride and quetiapine groups.

13.6.4 Typical antipsychotics versus atypical antipsychotics

- There is low quality evidence from a meta-analysis of two studies [one RCT and one quasi-RCT] showing no significant difference in recovery from delirium between the haloperidol and olanzapine groups.
- There is moderate quality evidence from one RCT showing no significant difference in the severity of delirium between the haloperidol and the olanzapine groups (an indirect measure of delirium was used).

13.7 Health economic evidence statements

The results of the economic model (chapter 16) showed the following:

- The use of haloperidol and olanzapine was cost-effective in the treatment of delirium in the hospital. This finding was robust as the interventions remained cost-effective after a series of sensitivity analyses were conducted.

- Haloperidol was more cost-effective than olanzapine in the treatment of delirium in the hospital. However, there was a wide uncertainty around the incremental cost-effectiveness of haloperidol compared to olanzapine.

13.8 From evidence to recommendations

There was little evidence for the use of pharmacological agents for the treatment of delirium. The GDG observed that there was evidence from one moderate quality RCT, but did not wish to make a recommendation on the basis of a single study which had a risk of bias (Hu 2006).

Economic evidence was obtained by modelling the treatment pathway for the two pharmacological interventions investigated in this study, and was also informed by the review on the consequences of delirium. The model also incorporated evidence on cost, quality of life and baseline risks.

The health economic analysis showed that haloperidol and olanzapine were cost effective compared with placebo for treating delirium, but the uncertainty around the cost effectiveness estimates precluded recommending one drug over another. The GDG took into consideration the possible harms of the medication, for which the evidence was largely indirect. The GDG were uncertain whether there was a risk of stroke when using these medications in the short-term treatment of delirium. Due to the limited evidence the GDG did not wish to consider a class effect and hence made recommendations for individual drugs (recommendation 1.6.4).

On balance, weighing up the effects of reduced mortality and dementia, versus possible increased risk of stroke, and taking into account the cost effectiveness analysis, the GDG decided that the benefits outweighed the risks, and that they should recommend drug treatment after other treatment interventions had been tried. In the light of the adverse events associated with these drugs for longer term use, and their uncertainty about the evidence, the GDG did not want to recommend the routine use of these drugs for everyone with delirium. The GDG therefore decided to make a cautious recommendation that healthcare professionals consider giving pharmacological treatment as short term treatment. Short-term treatment was defined as 1 week or less, based on the evidence from the Hu (2006) study and usual practice.

The GDG considered that this treatment should only be given to patients who had distressing symptoms and whose behaviour meant their safety or the safety of those around them was compromised. This was in line with the summary of product characteristics (SPC) indications for these drugs for the treatment of symptoms: 'rapid control of agitation and disturbed behaviours in patients with schizophrenia or manic episode' for olanzapine and 'As an adjunct to short term management of moderate to severe psychomotor agitation, excitement, violent or dangerously impulsive behaviour' for haloperidol' (SPCs).

The GDG were aware that antipsychotic drugs such as haloperidol and olanzapine should be used with caution or not at all for people with conditions such as Parkinson's disease and/or Lewy-body dementia. They therefore made a recommendation to this effect and cross-referred to the NICE guidelines on 'Parkinson's disease' (NICE clinical guideline 35) and 'Dementia' (NICE clinical guideline 42).

The GDG also wanted to give guidance for all people who had progressedthrough the care management and treatment pathway but whose delirium symptoms had not fully resolved. This could be due to underlying causes remaining to be addressed or could indicate that the person has dementia.

The GDG wished to investigate further the clinical and cost effectiveness of the range of pharmacological agents currently used for treating delirium and proposed a research recommendation (see below and Appendix H).

Future research recommendation:

In hospital patients with delirium, which is the most effective medication (atypical antipsychotic, typical antipsychotic, benzodiazepines) compared with placebo or each other for treating delirium?

13.9 Recommendations

Distressed people

If a person with delirium is distressed or considered a risk to themselves or others and verbal and non-verbal de-escalation techniques are ineffective or inappropriate, consider giving short-term (usually for 1 week or less) haloperidol[14] or olanzapine[14]. Start at the lowest clinically appropriate dose and titrate cautiously according to symptoms. **[1.6.4]**

Use antipsychotic drugs with caution or not at all for people with conditions such as Parkinson's disease or dementia with Lewybodies[15]. **[1.6.5]**

If delirium does not resolve

For people in whom delirium does not resolve:

- Re-evaluate for underlying causes.
- Follow up and assess for possible dementia[16]. **[1.6.6]**

---

[14] Haloperidol and olanzapine do not have UK marketing authorisation for this indication.

[15] For more information on the use of antipsychotics for these conditions, see'Parkinson's disease' (NICE clinical guideline 35) and'Dementia' (NICE clinical guideline 42)

[16] For more information on dementia, see 'Dementia' (NICE clinical guideline 42).

14 Adverse effects

CLINICAL QUESTIONS:

What are the most clinical and cost effective and safe pharmacological interventions for the prevention of delirium in people in hospital?

What are the most clinical and cost effective and safe pharmacological interventions for the prevention of delirium in people in long-term care?

What are the most clinical and cost effective and safe pharmacological interventions for treating people with delirium in hospital?

What are the most clinical and cost effective and safe pharmacological interventions for treating people with delirium in long-term care?

14.1 Background

A wide variety of pharmacological interventions are available for the prevention and treatment of delirium. The drugs have varying pharmacological actions, and patients may potentially be troubled by a wide spectrum of adverse effects depending on the agent administered.

In making rational treatment choices, healthcare professionals need to carefully weigh up evidence on the anticipated benefits against that of any relevant concerns about the safety and tolerability. There are two important aspects in a review of adverse effects data for drugs in delirium:

- Evaluation of comparative data among different drugs can help healthcare professionals arrive at a treatment decision for a particular agent based on whether the safety profile (nature and frequency of adverse effects) is more, or less, acceptable than the other available agents.

- Healthcare professionals should be aware of the most important adverse effects that can arise after giving the therapy so that they can take appropriate measures to detect and minimize the risk from adverse effects

In most illnesses, patients are given adverse effects information to guide their choice of treatment and to enable them to seek medical attention for any untoward symptoms. However, patients receiving treatment for delirium may have little say in the matter, and have to rely on the actions of the healthcare professional. As such the onus is on the healthcare professional to make the appropriate decisions and to institute relevant monitoring and precautionary measures.

While some details on adverse effects have been covered in the parallel efficacy reviews of delirium, there is limited information on the specific adverse effects. It is also unclear whether the classes of drugs differ in their safety and tolerability profile.

14.1.1 Objective:

To determine what specific adverse effects may arise from drug therapy for prevention or treatment of delirium.

14.2 Selection criteria

The selection criteria described in the general methodology section (section 2.3.1) were used, but some were specific to the evaluation of adverse effects and are reported in the following sections.

14.2.1 Types of studies

We did not apply any specific inclusion criteria based on study design; however, we aimed to exclude:

- Published case reports and case series of specific adverse events, as there is a large degree of publication bias stemming from authors' and editors' decisions favouring manuscripts covering esoteric or interesting patients. Such cases are unlikely to be representative of the general patient population
- Cross-over studies, as it is impossible to discriminate between events that arise as a complication of the first (previous) treatment, or as events resulting from the present therapy (carry-over).
- Small studies with fewer than 20 patients exposed to the intervention of interest, as such studies are unlikely to be able to detect any important adverse effects, and may lead to falsely reassuring findings that no safety problems were identified.

14.2.2 Types of participants

- Adults (18 years and over)

- Patients requiring treatment for delirium or being given treatment to prevent delirium

- Not end-of-life patients or patients with primarily psychiatric disorders such as schizophrenia, bipolar disorder or other psychoses.

Following GDG advice and post-hoc evidence from an indirect population was included in order to investigate stroke as an adverse event. The GDG extended the population to include older patients and those with dementia.

14.2.3 Interventions of interest

- Typical antipsychotics: haloperidol

- Atypical antipsychotics: risperidone, olanzapine, quetiapine, amisulpride

- Benzodiazepines: diazepam, flunitrazepam

- Cholinesterase inhibitors: donepezil, rivastigmine

- 5-HT3 antagonists: ondansetron

Duration of intervention: any

14.2.4 Comparators

For controlled studies, we accepted comparisons of any of the above agents versus placebo or no treatment. We also included studies that directly compared two or more agents from the above list of interventions. However, we excluded studies if the relevant intervention was tested against an active comparator that was not on the list of included drugs, as it would then be impossible to draw any valid conclusions on the relative safety profile of the agent of interest (safer or more harmful than an intervention of unknown effect).

14.2.5 Outcomes

All outcomes reported within the categories of 'adverse effects, side effects, adverse events, complications, safety, or tolerability'.

14.2.6 Assessment of Validity of Adverse Effects Data

The methods for assessing validity were based on recommendations of chapter 14 of the Cochrane Handbook of Systematic Reviews. This focuses on two major factors:

- How thorough were the methods used in monitoring adverse effects?

- How complete or detailed was the reporting?

In view of this, the following parameters were recorded:

- What methods (if any) did the studies stipulate for the specific assessment of adverse effects?
- Did the investigators prespecify any possible adverse events that they were particularly looking out for?
- What categories of adverse effects were reported?

14.3 Identification of studies

Articles that had already been retrieved for the efficacy reviews were considered and reference lists were checked to identify specific articles on adverse effects.

A total of 170 full text articles were screened, with 16 studies fulfilling the inclusion criteria.

However, we had to make further exclusions due to no adverse effects data being extractable. Three eligible studies failed to mention anything about adverse effects and were not evaluated any further. (Hu 2006: olanzapine, haloperidol and control; Liu 2004: risperidone; Moretti 2004: rivastigmine).

Adverse effects data were extracted from 13 included papers (Aizawa 2002; Bayindir 2000; Breitbart 2002; Kalisvaart 2005; Kaneko 1999; Kim 2001; Lee 2004; Liptzin 2005; Miyaji 2007; Pae 2004; Parellada 2004; Prakanratta 2007; Skrobik 2004).

Following GDG advice, indirect evidence was obtained from three further studies (Douglas 2008; Gill 2005; Hermann 2004).

14.3.1 Study Design

The following types of studies were included in the adverse effects analysis:

- Direct head to head comparison of two antipsychotic agents: 1 RCT (Lee 2005), 1 quasi-randomised study (Skrobik 2004), 1 prospective cohort study (Gill 2005; with retrospective elements), and 2 retrospective cohort studies (Herrmann 2004; Miyaji 2007).
- Typical antipsychotic: haloperidol, 2 placebo controlled RCTs (Kalisvaart 2005; Kaneko 1999); typical antipsychotics generally, 1 retrospective cohort study (Douglas 2008)

- Atypical antipsychotics: 6 studies consisting of 1 RCT (Prakanratta 2007), 3 open trials without control arms (Breitbart 2002; Kim 2001; Pae 2004), and 3 observational studies (Douglas 2008; Parellada 2004).

- Benzodiazepines: diazepam, flunitrazepam: no studies met the eligibility criteria. One study (Aizawa 2002) that was included in the efficacy review had to be excluded as the intervention involved three agents – diazepam, flunitrazepam and pethidine, and it would not have been possible to tell if any adverse effects were due to the benzodiazepine or the pethidine.

- Cholinesterase inhibitors: donepezil, rivastigmine. One double blind placebo controlled RCT (Liptzin 2005).

- 5-HT3 antagonists: ondansetron – one open trial without control arm (Bayindir 2000)

14.3.2 Population

The studies looked at a wide range of participants, but for the most part were in patients undergoing surgery or admission to intensive care. Three of the studies (Douglas 2008; Gill 2005; Hermann 2004) reported on stroke adverse events associated with antipsychotics in older patients, who were likely to be at risk of delirium. These studies were on an indirect population.

14.3.3 Intervention and Comparisons

There was a diverse range of interventions, and associated comparator agents across the trials.

14.3.4 Assessment and Reporting of Adverse Effects

A diverse range of methods were used, with the most well-defined ones being scales for assessing extrapyramidal signs and symptoms. It is not clear though how useful such scales are in postoperative or intensive care patients, in contrast to ambulant psychiatric patients.

14.4 Results

The interventions, comparators and populations were extremely varied, as was the reporting of adverse effects outcomes. Descriptive summaries are given in Appendix D.

14.4.1 Direct comparison of active agents

Five studies (Gill 2005; Herrmann 2004; Lee 2004; Miyaji 2007; Skrobik 2004) reported direct comparisons between two antipsychotic agents.

Extrapyramidal adverse effects were the main focus of three studies (Lee 2004; Miyaji 2007; Skrobik 2004), with one study (Skrobik 2004) describing specific efforts to "carefully record" such events. Two studies reported specifically on stroke as an adverse event (Gill 2005; Herrmann 2004). One study was in older adults (mean age 81.7 years) (Herrmann 2004) and one study was in older adults with dementia (mean age 82.6 years) (Gill 2005).

No extrapyramidal events were found in the Lee (2004) study, but both Miyaji (2007) and Skrobik (2004) studies described a higher incidence of extrapyramidal effects with haloperidol as compared to quetiapine, and olanzapine respectively. However the Miyaji (2007) study was retrospective while Skrobik (2004) was quasi-randomised, and neither study had any blinding and are thus subject to bias from investigators who may favour the new atypical antipsychotics when recording the extrapyramidal effects.

While the ascertainment of mortality is less subjective, the baseline differences in populations receiving the interventions in the Miyaji (2007) study makes it difficult to draw any reliable conclusions, simply because the more severely ill patients may have received parenteral haloperidol.

Two studies carried out multivariate analyses (Gill 2005; Herrmann 2004). The Gill (2005) study did not take into account confounders such as smoking history, presence and severity of hypertension, lipid status and specific valvular heart conditions. Similarly the Herrman (2004) study did not take into consideration smoking or obesity. The most commonly prescribed antipsychotic was risperidone in both studies (Gill 2005: 76%; Herrmann 2004: 61%)

The Gill (2005) study reported that in older patients with dementia there is no significant difference in the effects of atypical antipsychotics compared with typical antipsychotics.

The Herrmann (2004) study reported results separately for olanzapine and risperidone compared with typical antipsychotics. There was no significant effect of olanzapine [RR 1.1 (95% CI 0.4 to 2.3)] or risperidone [RR 1.4 (95% CI 0.7 to 2.8)] on the incidence of stroke. A head to head comparison (risperidone versus olanzapine) showed no difference in effect [RR 1.3 (95% CI 0.8 to 2.2); (figure 14.1, Appendix K).

14.4.2 Results of specific classes of interventions versus no treatment or placebo

Typical and atypical antipsychotics

One retrospective cohort study (Douglas 2008) was an intra-patient study comparing periods of antipsychotic use and non-use in older adults (indirect population). Median age when first exposed to any antipsychotic drug was 80 years. The study reported on the risk of stroke in patients presenting with first ever stroke (at least 12 months after initial registration on the UK general practice database). The most commonly prescribed atypical antipsychotic was risperidone (81%), followed by olanzapine (18%), amisulpride and quetiapine (4% in each group).

Exposure to any of the antipsychotics was a significant risk factor for stroke [RR 1.73 (95% CI 1.60 to 1.87)]. When typical and atypical antipsychotics were analysed separately, a significant effect was observed [typical antipsychotics: RR 1.28 (95% CI 1.18 to 1.40); atypical antipsychotics: RR 2.32 (95% CI 1.73 to 3.11)]. (figure 14.2, Appendix K).

Haloperidol

There were two included RCTs, both covering the use of haloperidol in postoperative patients. (Kalisvaart 2005, Kaneko 1999)

Both trials reported on active measures to detect adverse effects, with frequent clinical assessments. Haloperidol use in this setting appeared to be relatively safe with no excess of withdrawals from adverse events compared to control, and no extrapyramidal effects.

Atypical antipsychotics

For risperidone, we identified one RCT (Prakanratta 2007) and one observational study (Parellada 2004). There were two open uncontrolled trials of olanzapine (Breitbart 2002, Kim 2001), and one of quetiapine (Pae 2004).

Both the risperidone studies looked for specific adverse effects but did not show any clear trend for harm.

One olanzapine study (Breitbart 2002) used clinical methods to evaluate adverse effects, and this showed sedation to be a problem necessitating dosage reduction.

The remaining two studies (Kim 2001, Pae 2004) did not mention any specific monitoring for adverse effects, and data were sparse.

Cholinesterase inhibitors

One study (Liptzin 2005) which was a randomised double-blind controlled trial of donepezil was identified. Despite methodological strengths elsewhere, this study did not describe any specific monitoring of adverse effects, and did not provide numerical data, even though there was a statement about equivalent rates of adverse effects between drug and placebo. Moreover, adherence to treatment was poor, and as such, no conclusions can be drawn on the relative safety of donepezil.

5-HT3 antagonists

One study (Bayindir 2000) which was an open-label uncontrolled evaluation of ondansetron in postoperative patients was identified. The authors did not state what, if any monitoring was used for detecting adverse effects, and it is difficult to have any confidence in their conclusions that the therapy was safe, without any apparent side effects.

1

Table 14.1: GRADE evidence summary - Adverse Events

Outcome/ Author (year)	Meta-analysis details	Summary statistics	Comments	GRADE details	GRADE Comments
Typical antipsychotic vs placebo (prevention)					
Adverse event (extrapyramidal) (Kalisvaart 2005)	1trial; 430 patients; from RCT	RR=1	Study reported no extrapyramidal events	• Study quality:no serious limitation • Directness: Direct • Imprecision: Number of events < 300 • Inconsistency: consistent • Reporting bias: Adequate	Placebo comparison Adverse events data-prevention trial
GRADE evidence rating: Low					
Adverse events (sedation) (Kalisvaart 2005)	1trial; 430 patients; from RCT	RR=1	Study reported no sedation events in either group	• Study quality:no serious limitation • Directness: Direct • Imprecision: Number of events < 300 • Inconsistency: consistent • Reporting bias: Adequate	Placebo comparison Adverse events data-prevention trial
GRADE evidence rating: Low					
Adverse events (tachycardia) (Kaneko 1999)	1trial; 78 patients; from RCT	RR=3.15 (95%CI 0.13, 75.12)	No significant difference	• Study quality: serious limitation-method of assessment of delirium unclear • Directness: Direct • Imprecision: CI crosses appreciable harm/benefit • Inconsistency: consistent • Reporting bias: Adequate	Placebo comparison Adverse events data-prevention trial
GRADE evidence rating: Very low					
Atypical antipsychotic vs placebo (prevention)					
Outcome/ Author (year)	*Meta-analysis details*	*Summary statistics*	*Comments*	*GRADE details*	*GRADE Comments*
Adverse event (cardiovascular instability) (Prakanrattana 2007)	1trial; 126 patients; from RCT	RR=0.76 (95%CI 0.18, 3.27)	Not significant	• Study quality:no serious limitation • Directness: Direct • Imprecision: Wide CI • Inconsistency: consistent • Reporting bias: Adequate	Placebo comparison Adverse events data-prevention trial
GRADE evidence rating: Low					
Atypical antipsychotic 1 vs atypical antipsychotic 2 (prevention)					
Outcome/ Author (year)	*Meta-analysis details*	*Summary statistics*	*Comments*	*GRADE details*	*GRADE Comments*
Adverse events (Lee 2005)	1trial; 31 patients; from RCT	RR=1	No significant adverse events reported such as acute dystonia dyskinisea	• Study quality: serious limitation-incomplete follow up • Directness: Direct • Imprecision: Number of patients < 400 • Inconsistency: consistent • Reporting bias: Adequate	Adverse events data-prevention trial Very small study; 25% missing data in 1 arm
GRADE evidence rating: Low					
Adverse events (sedation) (Brietbart 2002)	1trial; 79 patients; from Cohort	Proportion (%)30	High proportion of patients with sedation events	• Study quality: very serious limitation • Directness: Indirect patients - minor, comorbidity • Imprecision: • Inconsistency: • Reporting bias:	Olanzapine; Hospitalised cancer patients;clinical examination for adverse events
GRADE evidence rating: Low					
Cross review: typical antipsychotic vs atypical antipsychotic 2 (treatment)					
Outcome/ Author (year)	*Meta-analysis details*	*Summary statistics*	*Comments*	*GRADE details*	*GRADE Comments*
Adverse event (extrapyramidal) (Skrobik 2004)	1trial; 73 patients; from Quasi RCT	RR=8.2 (95%CI 0.48, 140.09)	No significant difference	• Study quality: very serious limitation • Directness: Direct • Imprecision: Wide CI • Inconsistency: consistent • Reporting bias: Adequate	Haloperidol vs olanzapine quasi randomised Adverse events data-treatment trial Adverse events carefully recorded; Not blinded.
GRADE evidence rating: Very low					

Overall summary

Results for stroke as an adverse effect for different types of antipyschotics in an indirect population is presented in table 14.2.

Table 14.2: summary of results for stroke as an adverse effect

	All antipsychotics vs placebo	Atypical antipsychotics vs no treatment	Typical antipsychotics vs no treatment	Atypical antipsychotics vs typical antipsychotics vs placebo	Atypical antipsychotic 1 vs Atypical antipsychotic 2
Stroke	1.73 (1.60 to 1.87)	2.32 (1.73 to 3.11)	1.28 (1.18 to 1.40)	1.01 (0.81 to 1.26)	1.3 (0.8 to 2.2)
				1.1 (0.4 to 2.3)	
				1.4 (0.70 to 2.8)	

14.5 Health economic evidence

No relevant health economic papers were identified.

14.6 Clinical evidence statements

- There is moderate quality evidence in a large:
  - retrospective cohort study that antipsychotics have a significant effect on the incidence of stroke in patients who have a median exposure time of 3 to 4 months. This is indirect evidence for patients who receive antipsychotics for delirium, who will have the drugs for much shorter periods. (Douglas 2008)
  - mixed prospective-retrospective cohort study in patients with dementia to suggest there is no significant difference in the effects of typical relative to atypical antipsychotics compared with each other. (Gill 2005)
  - retrospective cohort study to suggest that there is no significant difference between risperidone and olanzapine as risk factors for stroke in patients who received drugs for at least 30 days. (Herrmann 2004)

14.7 From evidence to recommendations

The GDG recognized the paucity of the reported adverse effects data is a major limitation. Most of the investigators appear to have focused on extrapyramidal effects, and omitted to consider or discuss the possibility of other adverse events. Another important limitation is that patients with delirium are unable to accurately describe of any untoward symptoms, and thus adverse events may have been missed by the clinicians. The heterogeneous data on haloperidol are of interest here, as this may possibly reflect susceptibility to bias in the unblinded studies that found an excess of extrapyramidal symptoms, when compared to newer atypical agents. The data on extrapyramidal effects and mortality should be judged cautiously, given that higher quality RCTs with thorough adverse effects monitoring have failed to replicate these findings.

All three studies (Douglas 2004; Gill 2005; Herrmann 2004) reporting on the incidence of stroke and antipsychotic use attempted to take into account known confounders, but each had limitations; the Gill (2005) study may have been higher quality because it was prospective but was solely in patients with dementia and the results may therefore not be generalisable.

The indirect evidence from the severe adverse events (stroke) of pharmacological interventions was incorporated into the health economic model. The GDG also took into consideration other direct evidence on adverse events (such as extrapyramidal symptoms) when making recommendations.

The GDG weighed up the benefits and harms when making their recommendations on pharmacological treatment (see section 13.8 and GRADE tables, section 13.4).

14.8 Recommendations

See recommendation 1.6.4.

15 What information is useful for people with delirium and their carers?

CLINICAL QUESTION: What information should be given to people at risk of developing delirium, or people with delirium, and their families or carers?

15.1 Clinical Introduction

Delirium can be a distressing experience for affected individuals, family caregivers and professionals. The symptoms can be complex and full or partial recall after the episode has resolved is common. Sometimes this can result in unpleasant "flashback" episodes. Information and education to improve understanding of delirium and its effects might help to improve outcomes from the condition.

15.2 Description of studies

Twenty four studies were ordered for this review. Fourteen studies were excluded.

One included study was UK based. There were four Swedish studies, two studies conducted in the USA and one each in Australia, Canada and Finland.

One non randomised control trial was reviewed and nine qualitative studies were critically appraised using the NICE qualitative methodology checklist. These studies were evaluated on the basis of six parameters which include: theoretical approach, study design, data collection method, validity, analysis and ethics.

Two of the included studies used a phenomenologic approach, one study used a hermaneutic approach and another study used a combined phenomenologic-hermaneutic approach. There were three studies which employed content analysis to elicit categories and themes based upon patient interviews and one further study which used an interview questionnaire to obtain subjective responses of family carers to the experience of delirium.

One of the included studies described an information giving intervention in a hospice setting. Although people receiving end-of-life care are excluded from the guideline, this was the only comparative study identified and the only study which assessed the actual development and implementation of a delirium educational tool for family caregivers. It was considered that the information in this study could be imputed to other settings.

15.3 Results

Owens & Hutelmyer (1981) conducted a non randomised control trial among 64 adults having cardiac surgery. The study tested the hypothesis that patients who are educated pre-operatively about the possibility of unusual sensory or cognitive experiences will not have such experiences postoperatively or will feel comfortable or in control of the experiences if they occur. Patients were assigned on a consecutive admission basis to either the intervention or control group. The staff did not discuss the psychological aspects of postoperative care with any participants. The investigator discussed the possibility of memory loss, inability to concentrate, inability to recognise familiar objects or persons and the possibility of seeing or hearing things that could not be explained or were not really there with the experimental group only. Post-operative interviews were conducted on days 4-8. Of the 32 patients in the control group, 25 reported at least one unusual experience. In the experimental group, 19 patients reported such experiences. The difference was not statistically significant. When the groups were compared as to whether they felt comfortable or in control during an unusual experience, the experimental group was significantly ($p<.05$) more comfortable.

Margarey and McCutcheon (2005) interviewed eight patients who had experienced hallucinations during an ICU admission. Most of these patients remembered the nurses talking to them even if they did not recall the ICU environment. Reassurance and comfort from the nurses was important to patients, particularly reassurance that the experience of delirium is common and that they were not going mad. The presence of family members was associated with the beginning of recovery. The authors of this study suggest that post ICU clinics to allow patients to discuss the experience of delirium and post ICU visits so that patients can put their experiences into context may be useful.

Duppils and Wikblad (2006) interviewed 15 patients who had undergone hip-related surgery and experienced delirium during their hospital stay. Difficulty in communication was identified as one of the risk factors in delirium. Patients complained that the nurses talked ‘about’ them, not ‘to’ them. Nurses were encouraged to try to understand the patients thought and experiences in order to communicate information in a therapeutic manner.

Nineteen patients who had been ventilated and stayed at least 36 hours in the ICU were interviewed by Granberg et al (1998) about one week after discharge and again 4-8 weeks after their discharge from the ICU. Patients described their first feelings and memories after delirium. Relatives provided a lifeline to reality. Patients were very sensitive to the attitude and behaviour of staff. They also reported the effort to regain control over their bodies. Patient reaction to the equipment of ICU which is unfamiliar and uncomfortable and limits mobility resulted in fear and tension. Caring nurses could provide rest from a state of prolonged tension and engender a feeling of security by helping with orientation to the surroundings and providing a sense of ‘We are with you.’ It was important for patients to know that unreal experiences are common and that their intellectual capacity would not be impaired. They appreciated nurses who would explain equipment and procedures and who understood that they needed help to regain control over their bodies.

Heleena Laitinen (1996) conducted a study of 10 postoperative intensive care coronary artery by-pass patients. Implications for nursing practice were highlighted, particularly understanding and acceptance. Being aware of space and time gives patients more confidence for coping with being in the ICU. Consciousness of space and time presumes that events and stimuli in the environment are constantly being explained to the patient in a sensitive manner.

Ewa Stenwall et al (2008) interviewed seven geriatric patients who had experienced acute confusional state (acute confusional state; delirium). Patients stated that gaining knowledge about what was happening and what was planned evoked feelings of safety.

Good communication occurs through the senses. Relatives can inform carers which sense the patient prefers and which sense is less efficient.

Another study by Stenwell et al (2008) explored the experience of relatives of patients with delirium. The conclusions of this study with regard to information giving were as follows:

- Relatives need information about acute confusional state (delirium) to alleviate their insecurity about interactions with the patient and to aid their understanding of the patient's behaviour which will allow trust to develop. It is necessary to inform relatives of the short term nature of acute confusional state and the need to have support and advice from professionals on how to communicate.

- Relative's knowledge of the patient should be used to inform the communication style of carers with that individual. Communication must be responsive to the individual encounter.

Fourteen elders participated in a phenomenologic study describing the experience of delirium patients (McCurren & Cronin, 2003). Three themes were identified:

- Being in the confusion event
- Responding to confusion
- Dealing with confusion

The latter theme involved the responses of family, staff and the patient. Among the interventions which helped with delirium included explanations from nurses which helped to reassure patients and families. Anticipatory explanations for surgical patients were also identified as helpful.

Another interpretative phenomenological analysis of nine patients (Harding 2008) aimed to understand the delirium experience of older people after reparative hip surgery. Semi structured interviews were conducted and two primary themes were identified:

- Struggling to understand the experience of delirium
- Strategies used in discussing delirium

Based upon an in-depth analysis of the experiences and concerns of the participants the authors suggested the following:

- Providing information for patients and relatives (e.g. in a leaflet) to help them understand delirium
- Training healthcare staff to help facilitate open discussions with patients about their delirious symptoms and supervision to help staff better understand and manage their own anxieties.

A psycho-educational intervention was implemented in a palliative care hospice to help family caregivers cope with delirium and eventually to contribute to early detection (Gagnon et al, 2002). Phase 1 of this study aimed to develop the framework of an optimal psycho-educational intervention about delirium through focus group discussion. Phase 2 was the development of a brochure to be used as part of the psycho-educational intervention and Phase 3 included the implementation and evaluation of the intervention by comparing 58 family who received 'usual care' and 66 families who received explanations by nursing staff and a brochure on delirium. The delirium brochure included the symptoms of delirium, the cause of delirium, staff actions when a patient has delirium and how to behave with a patient with delirium.

Those who received the intervention felt more competent in making decisions than those in the usual care group ($p=0.006$) and the majority felt that all family caregivers should be informed on the risk of delirium ($p<0.009$).

15.4 Health economic evidence

No relevant health economic papers were identified.

15.5 Clinical evidence statements

Overall, the studies on giving information to patients employ a variety of qualitative methods, with typically small numbers of participants in each study. Papers on information giving address the needs of patients, professional staff and family carers and identify needs throughout the delirium continuum from pre-delirium, to the delirium experience itself and finally to the post-delirium state. The following themes for information sharing appear in the literature:

- Patients need insight into the experience of delirium to promote their understanding and to decrease fear. Pre-op information and a visit to the ICU are recommended.

- Nurses require insight into the patient experience in order to promote empathy.

- As relatives provide a link with reality and can facilitate communication, they require anticipatory information about the risk for delirium.

- Post-delirium patients appreciate to offer the opportunity of discussing their experience and provide reassurance. Visiting the ICU following extubation may help a patient understand his/her experience.

15.6 From evidence to recommendations

There was qualitative and quantitative evidence from the patient information review, which informed GDG discussions.

The GDG discussed who should be given information about delirium and at what stage(s) in the patient pathway. It was decided that it was impractical to give every person that presented in hospital or long-term care information about delirium and it may unduly worry those who were not at risk. Information would be most useful to people in hospital or long-term care at two stages in their care pathway: those who had been assessed and found to be at risk of delirium, and at a later stage to people diagnosed with delirium.

The GDG also noted from the evidence that it was important for information to be given to the relatives and carers of people both at risk of delirium and diagnosed with delirium.

The evidence review and experience of the patient representatives added to the patient information recommendations. The GDG considered that information about delirium could easily be incorporated into existing material for patients and relatives.

The GDG agreed a recommendation about patient information in accordance with equalities legislation and NICE's equality scheme. The information given should be accessible to people with additional needs such as physical, sensory or learning difficulties, and to people who do not speak or read English. Standard information delivery may not be applicable / beneficial to people with different cultural, linguistic, cognitive and literacy needs.

The evidence from the patient information review was also used to help inform recommendation 1.6.2 in the non-pharmacological treatment review.

The GDG also made two recommendations for future research (see below and Appendix H)

Future research recommendations:

Does giving information about delirium to people in a UK hospital or long-term care, who are at risk of delirium, increase their ability to cope if delirium subsequently occurs, and does the information decrease the duration of delirium?

In people with dementia, does an education programme in delirium for carers improve the recognition of acute confusion and reduce the severity and duration of delirium, compared to an education leaflet or usual care?

15.7 Recommendations

Information and support

Offer information to people who are at risk of delirium or who have delirium, and their family and/or carers, which:

- informs them that delirium is common and usually temporary
- describes people's experience of delirium
- encourages people at risk and their families and/or carers to tell their healthcare team about any sudden changes or fluctuations in behaviour
- encourages the person who has had delirium to share their experience of delirium with the healthcare professional during recovery
- advises the person of any support groups. **[1.7.1]**

Ensure that information provided meets the cultural, cognitive and language needs of the person. **[1.7.2]**

16 Health economic models: cost-effectiveness analyses of delirium prevention and pharmacological treatment

16.1 Introduction

The occurrence of delirium has been shown in a systematic review to result in adverse consequences (chapter 9). The adverse consequences could lead to a reduction in patients' health-related quality of life, HRQoL, and the expenditure of the resources of the NHS or PSS. It will therefore be useful to know the cost-effectiveness of prevention and treatment interventions for delirium.

We searched the literature for existing cost-effectiveness results that could reliably inform the guideline recommendations and we identified four papers (Bracco 2007; Pitkala 2008; Rizzo 200; Robinson 2002). However, none of them were felt to be directly applicable to the guideline population. It therefore became necessary to develop an original economic evaluation model to determine the cost-effectiveness of strategies for the prevention and treatment of delirium in different care settings. As described above and in the general cost-effectiveness methods section (section 2.6), the model was constructed for prevention and treatment interventions in hospital care setting.

16.1.1 Interventions

There were a number of interventions strategies included in the systematic review of prevention and treatment interventions (chapters 10, 11, 12 and 13). However, after considering the existing evidence, the GDG wanted more information on the cost-effectiveness of two multicomponent prevention interventions and two pharmacological treatment interventions. They advised that these should be evaluated in the economic model. The two multicomponent prevention interventions were those included in the Inouye (1999) study and Marcantonio (2001) study. The two pharmacological treatment interventions were those in Hu (2006). These studies have been described fully (chapters 11 and 13).

Study participants in the Inouye (1999) study were consecutive patients admitted to the general medicine service in the non-intensive care section between March 1995 and March 1998. Patients were at least 70 years old, had no delirium at the time of admission, and were at intermediate or high risk for delirium at base line. There were 852 patients in the study and half of the sample received the multicomponent targeted intervention, Elder Life Program. They received standard protocols for the management of six risk factors for delirium namely, cognitive impairment, sleep deprivation, immobility, visual impairment, hearing impairment, and dehydration. Geriatric nursing assessment and interdisciplinary rounds were other program interventions targeted towards the risk factors. Patients in the usual care group received standard hospital services in the general-medicine unit.

Study participants in the Marcantonio (2001) study were 65 years old or older patients and were admitted non-electively for surgical repair of hip fracture. Patients in the intervention group received proactive geriatric consultation, which began preoperatively or within 24 hours of surgery. They received targeted recommendations based on a structured protocol from the geriatrician during the period of hospitalization. Patients in the control group received usual care. They received management by the orthopaedics team, including internal medicine consultants or geriatricians on a reactive rather than proactive basis.

The study participants in the Hu (2006) study were elderly inpatients with senile delirium selected from September 2001 to September 2003. The enrolled patients were divided into three groups including two treatment groups and a control group. Each of the two treatment groups received somatic treatment in addition to either haloperidol or olanzapine. The control group received only somatic treatment only.

16.1.2 Population

The model was developed for patients in hospital settings. The two multicomponent interventions were targeted at patients with specific risk factors for delirium while the treatment interventions were indicated for patients with delirium. For the prevention interventions, we chose to model the cost-effectiveness in the trial population rather than extrapolate to other populations as the patients were selected on the basis of specific risk factors and the intervention was targeted at modifying those specific risk factors. Therefore the GDG felt that the efficacy may not translate to other populations. The starting age used in the model was 79 years. This was based on the mean age reported in the largest of the three studies above (Inouye 1999).

16.1.3 Outcomes

The outcomes of interest for the model were the incremental cost and the incremental quality-adjusted life years (QALY) gained. These were used to calculate the incremental cost effectiveness ratio (ICER) and the incremental net monetary benefit (INMB).

16.2 The prevention model

16.2.1 The model structure for the prevention interventions

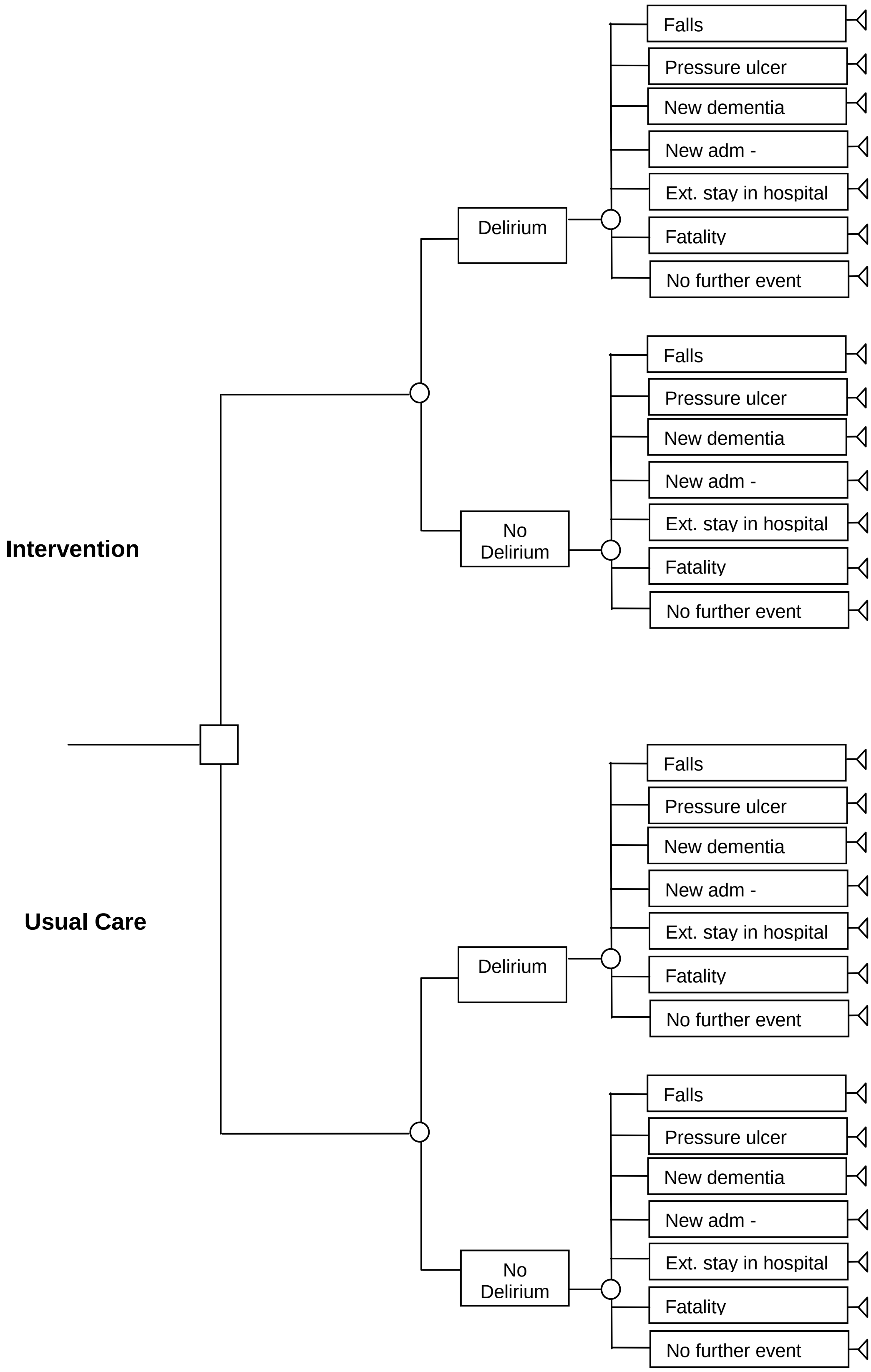
Decision Tree

The cost-effectiveness model consists of a simple decision tree which captures the outcomes of economic importance. The outcomes at the end of each branch of the tree include the adverse consequences of delirium. These outcomes will negatively impact on patient's health status and will lead to the expenditure of the resources of the NHS and PSS. The GDG advised that the adverse consequences to be used in the economic model should include falls, pressure ulcer, new dementia, new admission to institution, extended stay in the hospital and fatality. The decision tree was applied to each strategy and was used to estimate the impact of each strategy on the expected number of delirium cases, cost and QALYs associated with the adverse consequences. The decision tree is as shown below in figure 16.1.

Some members of a hypothetical cohort receiving each intervention strategy will become delirious and others will not. In the usual care strategy, the number that will become delirious will depend on the baseline risk of delirium in the care setting. The baseline risk of delirium is the risk of becoming delirious under no intervention conditions. In the intervention strategy, the number will depend on the baseline risk as well as the relative risk of becoming delirious if exposed to the intervention. The relative risk measure here is a measure of the efficacy of the intervention strategy. It is a ratio of the risk of becoming delirious among members of a population exposed to an intervention compared with a similar population that is not exposed to the intervention.

In non-delirious patients, the number of cases of the adverse consequences will depend on the baseline risk of the adverse consequence. In delirious patients, it will depend on the baseline risk as well as the relative risk of experiencing the adverse consequences if exposed to delirium. The end point of each branch of the tree implies a particular cost and a particular QALY. The total number of cases of delirium and the adverse consequences, the associated total cost and QALYs are summed up for each strategy.

Figure 16.1: decision tree for prevention intervention strategies



16.2.2 Baseline Risk

Hospital (intervention in general medicine services)

The baseline risk of delirium in the hospital was taken from a matched controlled trial in the USA (Inouye 1999). The study has been described in the review of prevention interventions (section 10.19). Study participants were consecutive patients admitted to the general medicine service in the non-intensive care section. Patients were at least 70 years old, had no delirium at the time of admission, and were at intermediate or high risk for delirium at base line. Half of the sample received the multicomponent targeted intervention while the other half received usual care. Usual care was defined as standard hospital services in the general-medicine unit. Patients were screened and baseline assessments were completed within 48 hours after admission. They were subsequently evaluated daily until discharge with a structured interview consisting of the Digit Span Test, Mini-Mental State Examination, and Confusion Assessment Method rating. Their medical records were reviewed after discharge for evidence of delirium, final diagnosis, medications, laboratory results, and destination after discharge. The primary outcome of the study was delirium defined according to the Confusion Assessment Method criteria. The median lengths of stay in the intervention and usual care groups were 7.0 and 6.5 days respectively. The incidence of delirium in the usual care group was 15% and this was used in the model as the probability of delirium in this group of hospitalized patients. In a sensitivity analysis, we used a lower incidence of delirium of 12.5%, which was the lower range of incidence reported in the needs assessment review for general medical patients (chapter 5).

Hospital (intervention in hip fracture surgery)

The baseline risk of delirium in the hospital for this patient group was taken from a randomised trial in the USA (Marcantonio 2001). The trial has been described elsewhere (section 10.19). Study patients were 65 years old or older and were admitted non-electively for surgical repair of hip fracture. Patients in the intervention group received proactive geriatric consultation, which began preoperatively or within 24 hours of surgery. Patients in the control group received usual care. The median length of stay in both groups was 5 days and the cumulative incidence during acute hospitalization was reported as 50% in the usual care group. This estimate was used as the probability of delirium in this patient group. In a sensitivity analysis, we used the lower estimate (15%) reported above for patients in general medicine services.

Dementia

The baseline risk of dementia was taken from a Canadian prospective cohort study (Rockwood 1999). It has been described in the section on the review of delirium consequences (chapter 9). Study patients were 65 years old or older and were consecutively admitted to the general medicine services of a tertiary-care hospital. A study cohort of 203 patients was followed up between June 1994 and August 1995, and dementia incidence as well as death was the primary outcome. Dementia diagnosis was done to conform to the Canadian Study of Health and Ageing dementia protocol. Dementia status was evaluated using the Informant Questionnaire on Cognitive Decline in the Elderly. Interview was obtained from proxy informants. A screening interview was also done to evaluate cognition and function. Cognition was done with the Blessed dementia rating scale while function was done with the Barthel index and the Physical Self-Maintenance Scale. The incidence of dementia in patients without cognitive delirium at baseline was reported as 5.6% per year. This baseline probability was used in the economic model.

Pressure Ulcer

The baseline risk of pressure ulcer was taken from a study that focussed on reporting the incidence of pressure sores across a NHS Trust hospital (Clark & Watts 1994). The number of patients admitted to the wards over 52 weeks were recorded alongside the number of those developing pressure sores. The severity and anatomical locations of pressure sores were also recorded. The incidence was monitored across four medical, three surgical and two orthopaedic wards and a record form was completed weekly. This enabled the identification of all patients that developed sores during the preceding seven days. The form also contained details of admissions and discharges from each ward and the details were obtained weekly. The number of people admitted in the wards as in-patients between December 1990 and November 1991 was 8935 and 360 patients developed pressure sores. This is equivalent to an incidence of 4.03% which we used as the baseline probability of pressure ulcer in the model. Some of the patients may have had delirium and as such 4.03% could be an over-estimate. We therefore used 1.68% in a sensitivity analysis. The latter estimate was reported in the O'Keeffe and Lavan study (1997) where two out of 119 non-delirious hospitalised patients acquired pressure sores. The latter study is briefly described in the next paragraph.

Falls

The baseline risk of falls was taken from a prospective cohort study in Ireland (O'Keeffe & Lavan, 1997). The study has been described in the section on review of delirium consequences (chapter 9) and it aimed to determine whether delirium is an independent predictor of adverse outcomes of hospitalization in older patients. The study population was 225 people admitted as an emergency over an 18-month period to an acute geriatric unit in a university teaching hospital. Only those on first admission within the study period were included in the study. Patients were excluded if they were not admitted to the geriatric unit on the day of admission, if they were admitted electively for investigations, rehabilitation, or respite care. Those that had severe aphasia or deafness, those that expected to remain in hospital for less than 48 hours, and those not assessed by a study doctor within 48 hours of admission were excluded. Patients were interviewed using the Delirium Assessment scale to elicit the presence and severity of individual DSM-III (Diagnostic and Statistical Manual, 3rd Edition) criteria for delirium. An initial assessment was done which included administration of an adapted Folstein Mini-Mental State Examination (MMSE) validated for use in an Irish population. All study patients were reviewed regularly and discussed with nursing and residential medical staff. The delirium status of patients was discussed at the multidisciplinary team meetings, and members of the team other than the study physicians were not aware of the underlying hypothesis of the study. Cases of falls, pressure sores, and urinary incontinence were recorded as hospital-acquired complications according to standardized criteria and were identified on the basis of interviews with the nursing staff. Pressure sore corresponds to grade 2 of Shea's classification. The number of patients studied was 225 and 42% had delirium defined by the DSM-3 criteria. The mean age of those with and without delirium was 82 years. Sixty eight percent of those without delirium were female and 16% of those without delirium were admitted from long-term care. Nine (7%) of the 131 non-delirious patients had falls, and we have used 7% as the baseline risk of falls in the economic model.

New admission to institution

We took the baseline risk of new admission to long-term care (LTC) from a prospective cohort study and it has been described in the section on the review of delirium consequences, chapter 9 (Bourdel-Marchasson 2004). The study was carried out in France with the aim of assessing the effects of delirium on the institutionalization rate in older patients hospitalized in an acute care geriatric unit, taking into account other components of frailty. Study participants were those older than 75 years old who were admitted between July 2000 and June 2001. Patients were excluded from the analyses if they spent less than 3 days in hospital, died before discharge or were usually living in an institution. The assessment of delirium was done with CAM within 24 hours following admission and then every three days during the hospital stay. The outcome considered for the analyses of study results was admission to a geriatric institution. There were 230 patients who were reported to be symptom free and 40 (17%) of these were discharged to geriatric institutions. We used 17% as the baseline risk of new admission to institute.

Mortality (in hospital)

The baseline risk of in hospital mortality was taken from the O'Keeffe and Lavan (1997) study described above. It was reported in the study that five percent of patients without delirium died during hospitalization, and we used this estimate as the baseline risk of mortality.

Mortality or new admission to institution

We have assumed that the adverse outcomes on the decision tree are mutually exclusive. This could potentially lead to double counting and over-estimation of costs and QALYs as some patients will experience more than one outcome at a time. The consequences review reported data on the relative risk of "mortality or new admission to nursing home" in delirious patients and we used this composite outcome rather than the single outcomes "mortality" and "nursing home admission" in a sensitivity analysis. This should reduce the double-counting and over-estimation of costs and QALYs associated with using the single outcomes in the model. We explored the effect of this sensitivity analysis on the cost-effectiveness result. This analysis requires an estimate of baseline risk for this composite outcome.

The baseline risk of mortality or new admission to institution was taken from a prospective cohort study in the USA (Marcantonio 2000). The study has also been described in the section on consequences review (chapter 9). The aim was to evaluate the role of delirium in the natural history of functional recovery after hip fracture surgery, independent of pre-fracture status. The study data were collected as part of a randomised trial to test whether proactive acute geriatrics consultation could prevent delirium after hip fracture repair. The effect of the intervention could have potentially affected the relationship between delirium and functional recovery but it was reported that the effect size of the associations did not differ between the two groups. Study participants were patients aged 65 years or older who were admitted to an academic tertiary medical centre for primary surgical repair of hip fracture. Patients with metastatic cancer or other co-morbid illnesses likely to reduce life expectancy to less than six months were excluded from the study. Study participants were interviewed daily during the duration of the hospitalization, including the Mini-Mental State Examination and Delirium Symptom Interview, and delirium was diagnosed using the Confusion Assessment Methods algorithms. They or their proxies were further contacted one and six months after fracture. They underwent interviews similar to those at enrolment to determine death, persistent delirium, decline in Activities of Daily Living function, decline in ambulation, or new nursing home placement. It reported the percentage of non delirious patients who died or were admitted to nursing home institute one month after hip fracture to be 12% and we have used this as the baseline risk of this outcome. This estimate is not compatible with the estimates reported above for new admission to nursing home and mortality but we recognise that these estimates were generated from studies carried out in different settings.

Mortality is defined in the model to be associated with zero cost. The number of people experiencing "new admission to institution" alone among the number of people experiencing "mortality or new admission to institution" was estimated by multiplying the total number of patients that died or were admitted to institute by 9%. This estimate was taken from the Marcantonio et al study (2000) which reported that, after one month, only three people died in a sample of 33 people that either died or had new nursing home placement. This was done to obtain an accurate cost and QALY estimate for this composite outcome.

Life Expectancy of delirious and non-delirious persons after discharge

The starting age in the model was 79 years. The survival of non-delirious patients post-discharge was different from that of delirious patients. We took account of this in the model by using the Kaplan-Meier survival curve reported in the Rockwood (1999) study. Of the delirious patients that were followed up for a median time of 32.5 months, 21% were alive, while 57% of the non-delirious patients were alive at follow-up. The median survival time was significantly shorter for those with delirium than for those without. An adjusted hazard ratio of occurrence of death of 1.71 was reported after adjusting for potential confounders on the risk of death. We used the data from the survival curve, fitted an exponential survival function to the data and estimated a baseline hazard of mortality of 0.007. In the three years after discharge, we applied these estimates to capture the different survival expectations in the three years after discharge for patients who have or haven't experienced delirium during admission. We then applied the same general population mortality rates (Interim Life Tables for England and Wales, 2005 - 07) to both groups up to age 100. We estimated a life expectancy of 3.6 years for patients with delirium and 5.4 years for patients without delirium.

Life expectancies applied in the model for patients in nursing homes and patients with new dementia

Patients staying in nursing home

The data on length of stay in long-term care attributable to delirium was taken from the results of two large-scale surveys of residential and nursing home residents in England (Netten 2001). They were a longitudinal survey of eighteen English local authorities and a cross-sectional survey conducted for the most part in the same authorities as the longitudinal survey. Information about the circumstances of 2,544 permanent publicly funded admissions from the authorities to residential and nursing home care was obtained in the longitudinal survey during a period from mid-October 1995 to mid-January 1996. In the cross-sectional survey, information about 11,900 residents in the homes was returned during the autumn of 1996. Cognitive impairment was identified using items from the Minimum Data Set. This allowed the compilation of the Minimum Data Set Cognitive Performance Scale. We assumed that the extra time a delirium patient spends in the long-term care after being transferred from the hospital will be equivalent to the time a patient with mild cognitive impairment spends in long-term care. The median length of stay for people with mild cognitive impairment was 18.9 months and we have assumed in our model that this is the survival time of patients that stay in long-term care.

New dementia

We took data on the life expectancy of a dementia patient from the study on the costs of dementia in England and Wales in the 21st century (McNamee 2001).The McNamee et al study (2001) was a Medical Research Council Cognitive Function and Ageing Study as well as a Resource Implications study. It provides estimates of formal care cost of dementia based on a population subgroup identified as cognitively impaired. The diagnosis of dementia was done using the Geriatric Mental State, and age- and gender-specific prevalence rates were estimated using data collected in a multi-centre study of four areas of England and one area in Wales. A sample of 2500 individuals was randomly selected from Family Health Services Authority or general practice files in the five centres. This included individuals in long-term hospital care. Life expectancy with dementia was estimated by applying age- and gender-specific prevalence rates for dementia to life tables. Cohort specific expectation of life with dementia was reported for the age groups, 65-69, 70-74, 75-79, 80-84, and 85+ for men and women. The specific life expectancies in years in the respective age groups for men were 0.7, 0.7, 0.9, 0.9 and 0.8 respectively. It was 1.5, 1.4, 1.8, 1.8 and 1.3 for the respective age groups in women. The population sizes in these cohorts were reported and we used in the base case analysis a weighted mean of 1.2 years as the length of time a dementia patient will live. The GDG suggested that this is rather an underestimate and suggested that the median estimate in the Dementia UK report (Dementia UK, Full report, 2007; Fitzpatrick et al 2005) should be used in a sensitivity analysis. The median life expectancies for individuals with Alzheimer's disease, vascular dementia and mixed dementia were reported as 7.1, 3.9 and 5.4 years respectively. The estimates were based on a US cohort study that examined mortality in 3602 participants who were evaluated for dementia incidence between 1992 and 1999 and followed for 6.5 years. The study was a subset of a larger Cardiovascular Health Study which recruited participants from Medicare eligibility lists in four US communities. Participants were to have completed a magnetic resonance imaging and three Mini-Mental State Exams in order to be eligible for the study. Dementia status was ascertained using data already collected in the Cardiovascular Health Study but supplemented with additional data on cognitive measures. The mean age of those with Alzheimer's disease, vascular dementia and mixed dementia were 80.1, 78.3 and 79.8 years respectively. We used a life-expectancy of 1.2 years for patients with dementia in the base case which is less than the modelled life-expectancy for patients without dementia. But in a sensitivity analysis we assumed that there is no increased risk of mortality due to dementia and therefore applied the life-expectancy for patients without dementia but taking into account the effect of delirium on life-expectancy.

16.2.3 Relative Risk of the adverse consequences of delirium

The relative risk estimate of adverse consequences of delirium was taken from the review of those consequences in chapter 9 and the estimates we used are listed in table 16.1 below.

The risk of new dementia was taken from the study by Rockwood et al (1999). This was the only study with a moderate quality that was included in the review for this outcome. It reported an adjusted odds ratio of 5.97 for new dementia which was assessed over a period of three years. We used relative risk estimates in the model and converted the reported odds ratio to a relative risk estimate using the formula,

RR = (OR) / [(1-Po) + (Po X OR)] (Zhang & Kai 1998)

Where RR is relative risk; OR, the odds ratio; and Po, the incidence rate in the unexposed population. The annual incidence of dementia among people without cognitive impairment at baseline was reported as 5.6% per year. We estimated a relative risk of 4.67 which we used in our economic model.

We used a similar method to estimate the relative risk of 2.05 for new admission to institution using an adjusted odds ratio of 2.64 (Bourdel-Marchasson et al 2004). There was a range of studies that reported the risk for this outcome but the odds ratio of 2.64 was chosen as it used incident delirium to estimate new admission to long-term care at the point of discharge.

The risk of falls and pressure ulcer was available from only one study (O'Keeffe and Lavan, 1997). The study reported an adjusted odd ratio of 2.3 for developing hospital-acquired complications which included falls and pressure ulcer. The relative risk of 2.18 for falls and pressure ulcer was estimated using the combined rate in the non-delirious group for falls and pressure ulcer.

The adjusted odds ratio of 2.6 for mortality in delirium patients in the hospital was taken from the O'Keeffe and Lavan (1997) study. We estimated a relative risk of 2.41 which we used in our model. There were other studies that reported the risk of in-hospital mortality but the GDG advised that it is best to use a UK study for this outcome. The way we have treated post-discharge mortality has already been described above.

Delirium extends hospital length of stay and the additional length of stay used in the model was estimated from a Kaplan-Meier plot reported in the Holmes and House (2000) study. This study was chosen because it was a UK study and was judged as being a high quality study for this outcome. We fitted a Weibull function using a lambda of 0.08 and gamma of 0.87 that were estimated from the Kaplan-Meier plot on the proportion of people in hospital at different times of discharge. This was for the people that were reported to be without a psychiatric diagnosis. The study also reported the result of a Cox Proportional Hazards model which showed that delirium is associated with a hazard ratio of 0.53 for hospital discharge. We applied this adjusted estimate to fit a Weibull function for the delirious group and estimated the difference in the area between the two fitted functions. This difference was 16.83 days and was treated in the model as the additional hospital length of stay due to delirium.

The adjusted odds ratio for the composite outcome of “mortality or new nursing home placement" after one month was reported as 3.0 (Marcantonio 2000). We converted this to a relative risk estimate of 2.41 which was used in a sensitivity analysis in the economic model.

Table 16.1: the baseline and relative risks of the adverse consequences of delirium

<table>
<tr><th>Adverse consequences</th><th>Baseline risk</th><th>Source</th><th>Odds ratio (95% CI)</th><th>Estimated relative risk (95% CI)</th><th>Source</th></tr>
<tr><td>New dementia</td><td>5.6%</td><td>Rockwood 1999</td><td>5.97 (1.83, 19.54)</td><td>4.67 (1.43, 15.29)</td><td>Rockwood 1999</td></tr>
<tr><td>New admission to institution</td><td>17.4%</td><td>Bourdel-Marchasson 2004</td><td>2.64 (0.83, 8.45)</td><td>2.05 (0.65, 6.57)</td><td>Bourdel-Marchasson 2004</td></tr>
<tr><td>Pressure ulcer</td><td>4.0%</td><td>Clark & Watts 1994</td><td rowspan="2">2.30 (1.7, 5.0)</td><td rowspan="2">2.18 (1.61, 4.73)</td><td rowspan="2">O'Keeffe & Lavan 1997</td></tr>
<tr><td>Falls</td><td>6.9%</td><td>O'Keeffe & Lavan 1997</td></tr>
<tr><td>Mortality</td><td>5.0%</td><td>O'Keeffe & Lavan 1997</td><td>2.60 (0.7, 6.2)</td><td>2.41 (0.65, 5.74)</td><td>O'Keeffe & Lavan 1997</td></tr>
<tr><td>Mortality or new admission to institution</td><td>12.2%</td><td>Marcantonio 2000</td><td>3.00 (1.1, 8.4)</td><td>2.41 (0.88, 6.76)</td><td>Marcantonio 2000</td></tr>
</table>

16.2.4 Efficacy of Interventions

The efficacy of the different intervention strategies has been reported in the review of multicomponent prevention interventions (section 10.19). It was reported that the use of these interventions by older general medical patients, who were at intermediate or high risk of delirium, was associated with a relative risk of delirium of 0.66 (Inouye et al 1999). The use of these interventions in older patients that underwent hip fracture surgery was reported to result in a relative risk of delirium of 0.65 (Marcantonio et al 2000). We have applied these estimates in our economic model.

16.2.5 Cost of Adverse Consequences of Delirium

Falls (cost)

The cost of falls data came from a randomised, controlled study of the prevention of fractures in the UK primary care. (Iglesias 2008). Eligible study participants were women aged 70 years and above with one or more risk factors for hip fracture and a total of 3,314 women were recruited into the study. The intervention group received daily oral supplementation using 1000mg calcium with 800 IU cholecalciferol and information leaflet on dietary calcium intake and prevention of falls (Porthouse 2005). The control group received leaflet only. Data on fracture and fall incidence, in additional to data on HRQoL and fear of falling, were collected at baseline and every 6 months after that for a minimum of 2 years and maximum of 42 months.

A fall and fracture questionnaire was used for resource use data collection and was administered to 1190 women participating in the prevention study and who had previously indicated to be willing to be contacted in the future for research purposes. Participants were asked if they had experienced a fall and / or fracture in the last 12 months, the number of times they had seen a doctor, GP or consultant and whether they had been hospitalised for reasons other than a fall or fracture and for how long, in the same period. Those that had experienced a fall or a fracture were further asked whether they had been hospitalised and how long they spent in hospital, the number of times they had seen a doctor or nurse, whether they had changed residence because of their fall and / or fracture and for how long. They were asked to describe any treatments that were specifically prescribed for their fall or fracture over the same period. Resource use was valued using unit costs from NHS reference cost data, Personal and Social Services Research Unit (PSSRU) data, as well as the Chartered Institute of Public Finance and Accountancy (CIPFA) data base. The NHS reference cost data was used to cost hospital inpatient length of stay as well as the cost of surgery following hip, wrist, arm and vertebral fractures. The CIPFA database was used to cost specialist contact visits, and the PSSRU data was used to cost GP and nurse visits, residential accommodation and the cost of home help.

The response rate to the questionnaire was 93% and 302 out of 1110 respondents reported falls in the previous 12 months and 62 of those that fell reported that their fall resulted in a fracture. Falls that did not result in fractures were generally associated with less resource use. There were 243 falls events that did not result in fractures and the mean cost was reported as £1,088. The number of falls that led to fractures was 10 for hip fracture, 7 for wrist fracture, 10 for arm fracture and 2 for vertebral fracture. The cost of falls leading to a fracture was reported as £15,133; £2,753; £1,863; £1,331; and £3,498 for hip, wrist, arm, vertebral, and other fractures respectively. We used a weighted estimate of £1875 in our economic model

Pressure Ulcer (cost)

The cost of pressure ulcer used in our model was taken from a cost study that aimed to estimate the annual cost of treating pressure ulcers in the UK (Bennett 2004). Treatment protocols which reflect good clinical practice for treating pressure ulcers of different grades were developed and costs for the daily resources defined in the protocol were assigned using representative UK NHS unit costs at 2000 prices. It was assumed that care is provided in a hospital or long-term care setting and that pressure ulcer patients are not admitted solely for the care of pressure ulcer. Resources to be used for care include nurse time, dressings, antibiotics, diagnostic tests, support surfaces and inpatient days where appropriate. Pressure ulcer was classified in four grades with grade 1 as the least severe and grade 4, the most severe. The daily costs for the ulcer grades were estimated for patients whose ulcer would heal normally as well as for patients whose ulcers were associated with critical colonisation, cellulitis and osteomyelitis. We assumed that pressure ulcers resulting from delirium are grade 1 pressure ulcers, would heal normally and are not associated with further complications. This assumption is conservative and is based on the finding that more complicated pressure ulcers are less common and represent less than 5% of all cases (Clark 1994). The cost per day for a grade 1 ulcer that heals normally is £38 and it will take 4.06 weeks on average for this class of ulcer to heal. The mean time to heal was taken from the same Bennett (2004) study and this estimate was reported to have come from a review of clinical literature. We therefore used a cost estimate of £1,064, up rated it to a 2007 estimate of £1364 (£1228.09 to £1499.86) using the inflation indices reported in PSSRU. The up rated estimate was applied in the model. The GDG suggested that some of the pressure ulcer cases due to delirium will be grade 4 pressure ulcers that will heal normally. They advised that the impact of this on the cost-effectiveness estimates should be investigated. We carried out a deterministic sensitivity analysis using the cost of grade 4 ulcer that heals normally. This was equivalent to a 2007 estimate of £9934.99.

Stay in long-term care (cost)

The cost of long-term care used in the model was estimated from the unit cost of stay in private nursing homes, private residential care, voluntary residential care and local authority residential care facility for older people. The care package costs per permanent residential week in private nursing homes were reported as £687 (PSSRU 2007). In private, voluntary and local authority residential care these were reported as £483, £480 and £858 respectively.

These unit costs have been estimated to include cost for external services such as community nursing, GP services as well as personal living expenses. They also include capital costs for the local authority residential care, and fees for the private and voluntary residential care. We subtracted £9.20, the cost of personal living expenses per week, from each unit cost and estimated £655.66, the weighted average of £677.80, £473.80, £470.80 and £848.80, to be the unit cost of long-term care. The weighting was based on the distribution of residents, 65 years and older, in care homes in 1996. It was reported that in nursing homes, local authority, private and voluntary residential homes the number of residents were 5746, 5476, 2791 and 3664 respectively (Netten 1998).

The NHS does not pay towards long-term care for all patients. It was suggested that only two percent of residents were funded by the NHS and overall, about 70% of the care home population were publicly funded (Netten 1998). We will consider the effect of this on the cost-effectiveness result by assuming in a sensitivity analysis that only 70% of the costs of long-term care will be borne by the NHS and PSS. The length of time a patient spends in the long-term care has been assumed to be 18.9 months and the source of this estimate is described above.

Hospital stay (Unit Cost)

We have used the unit cost estimates per excess day associated with complex elderly patients. This was reported as unit cost per day for days exceeding the trim point. We took all the HRG unit costs reported for all Complex Elderly patients (Hospital Episode Statistics for England. Inpatient statistics, 2007 – 08) and found a weighted mean of £152. There will be no additional costs on the basis of inpatient rehabilitation services as the GDG advised that, if at all, only a small number of delirium patients will need such services.

New Dementia (Cost)

Our cost estimate for dementia was taken from a report of the prevalence and cost of dementia prepared by the PSSRU and the Institute of Psychiatry (Dementia UK, The full report, 2007). The cost estimate was based on an interview of 132 dementia patients and dementia carers, who were referred to psychiatric services between January 1997 and June 1999. Service use was measured with a version of the Client Service Receipt Inventory and study participants were asked for details of accommodation and services during the past three months. Medication, inpatient and outpatient care, day hospitals, day centres, community health services, social care and respite care were the services included in the costing framework. Resource use for the services was valued using unit cost and estimated costs were inflated to reflect 2005/6 price levels. Cost of accommodation was based on a weighted average of unit costs for supported accommodation. Costs were based on only 114 definite cases of dementia, the study sample was London-based and an adjustment was made to reflect the UK as a whole. The cost of informal care was also included but we have excluded such costs here as the cost of informal care is outside the remit of NICE. The annual cost of late onset dementia per person was reported to be £25,472. Of this, accommodation accounted for 41%, NHS care services 8%, social care services 15%, and informal care services 36%. We subtracted the cost of informal care services and arrived at a cost estimate of £16,302 which was used as the annual cost of new dementia in our economic model. In a sensitivity analysis, we assumed that the cost of accommodation has been accounted for in the model, and have also subtracted the cost of accommodation. We estimated the cost of dementia to include only the cost of NHS services and social care services and arrived at a cost of £5,859. In the base case analysis, we have assumed that the life expectancy of a delirium patient is 1.2 years, and we have increased this in sensitivity analysis. The sources of the life expectancy estimates are described above.

Mortality (Cost)

We have not accounted for any additional cost resulting from mortality in our model. We have assumed that the cost associated with mortality has been incurred in the period up to the point of death, and that this has been captured in the model in the cost of adverse consequences that would eventually lead to death.

16.2.6 Utility of Adverse Consequences of Delirium

Falls (Utility)

The utility estimate for falls used in the economic model was taken from a Dutch randomised controlled trial (Hendriks 2008). It was an economic evaluation that aimed to assess whether a multidisciplinary intervention program would be preferable to usual care in the Netherlands. The study participants were those 65 years of age or over, and who had visited the accident and emergency department or general practice cooperative for the consequences of a fall. The exclusion criteria were inability to speak or understand Dutch, inability to complete questionnaires or interviews by telephone, cognitive impairment, admission for more than 4 weeks to a hospital or other institution, being permanently wheelchair-dependent or bedridden. Follow-up time was 12 months after baseline. The intervention included medical and occupational-therapy assessment that aimed to assess and address potential risk factors for fall. In usual care, medical risks and other risk factors were not systematically recorded and addressed by hospital physicians, specialists or GPs. Participants responded to the standard Dutch version of the EQ-5D in self-administered questionnaires at baseline and after 4 and 12 months. Utility scores for the EQ-5D responses were estimated using UK based social tariff. The mean age of the 167 participants in the usual care arm of the trial was 75.2 years. The mean utility at 4 and 12 months was reported as 0.72 and 0.71 respectively. The QALYs at the end of the follow-up was reported as 0.71.

In order to estimate the expected lifetime QALY gains for patients who experience falls we applied a utility multiplier in the first year of a falls' patient's life. The utility multiplier was estimated as the ratio of the utility of 0.71 reported at the end of the study follow-up and 0.74, the utility of a person aged 75.2 years old in the UK population. The utility of the population varies by age and the population utility was derived from an algorithm that was produced after a re-analysis of data from Kind (1998) in Ward (2007) study. In the model, the starting age is 79 years and the utility multiplier, 0.96 was used to adjust 0.72, the utility of an average British person aged 79. The QALY gains for the rest of the patient's life expectancy were estimated from a Markov survival model from the Life Table. In our estimates, we took account of the three year differences in survival chances of delirious and non-delirious patients (see section on mortality after hospital discharge).

Pressure Ulcer (Utility)

We did not identify any useful utility data on the HRQoL impact of pressure ulcer. The life-time expected QALY gain for a person who has experienced a pressure ulcer was assumed to be equal to the QALY gain of a person without any adverse consequence of delirium. This was estimated from a Markov survival analysis from the Life Table and we accounted for the three year differences in the survival chances of delirious and non-delirious patients (see section on mortality after hospital discharge). We estimated the expected lifetime QALY gain of a delirious person as 2.13 and the expected lifetime QALY gain of a non-delirious person as 3.09.

Long-term care (utility)

We could not identify a useful study that measured the utility of patients in long-term care. The GDG advised that the utility of a delirium in long-term care should be assumed to be equivalent to 0.25, the utility of a patient with severe dementia (Ekman 2007). The Ekman (2007) study aimed to obtain primary data on community-based health utilities in different stages of mild cognitive impairment and dementia from a general population sample. It was a cross-sectional study of subjects aged 45 – 84 years who were randomly selected in Sweden. A questionnaire was sent to a sample of 1,800 subjects and a description of the health conditions and how to value them was given. Four vignettes describing health conditions involving cognitive impairments typical for the progressive stages of dementia were made using the Clinical Dementia Rating scale. Mild cognitive impairment was defined as an overall Clinical Dementia Rating score of 0.5. Valuation of the perceived quality of life in theses stages was carried out using the time trade-off techniques. Respondents were reported as fairly representative of the general population in terms of age, gender, and employment. The mean age of women and men were 66.4 and 67.1 years respectively and 54.4% of the study sample was women. The mean utility score for severe dementia was reported as 0.25. This was used as a utility multiplier in the model. The mean age in the model is 79 years and the utility multiplier, 0.25 was multiplied with 0.72, the utility of an average British person aged 79. The adjusted utility of 0.18 was used to estimate the expected lifetime QALY gains after admission to long-term care.

Hospital stay (Utility)

We would expect some utility changes for staying in the hospital but the associated QALY gain will be small because of the short length of stay in hospital. We have therefore not included the impact of utility changes resulting from hospital care in our economic model.

New Dementia (Utility)

The utility score for new dementia was taken from the report by Ekman (2007). This study has been described above in the section on the utility of patients in long-term care. The mean utility score for mild, moderate and severe dementia were reported as 0.62, 0.40 and 0.25 respectively. The GDG advised that we use the utility score reported for moderate dementia. We applied this as a utility multiplier in the model and estimated a utility of 0.28 which was used to estimate the expected lifetime QALY gains for this outcome. The life expectancy used in the base case was 1.2 years and in the sensitivity analysis we used 3.6 years for dementia patients who experienced delirium and 5.4 years for those who did not experience delirium.

Mortality (Utility)

We have used zero QALY gain in the event of mortality.

16.2.7 Cost of multicomponent Targeted Intervention

The use of multicomponent targeted intervention in older patients admitted non-electively for surgical repair of hip fracture

The costing of multicomponent targeted intervention in patients admitted for surgical repair of hip fracture is based on the intervention protocol of a randomised controlled trial in an orthopaedic surgery service (Marcantonio 2001). The trial has been described in the section on the use of multicomponent interventions for delirium prevention (section 10.19). The trial aimed to determine whether proactive geriatrics consultations can reduce delirium after hip fracture repair. It was carried out in US patients, 65 years or older, who were admitted non-electively for surgical repair of hip fracture. All study patients had an intake assessment that included a patient interview, a proxy interview, and a review of the medical record. Patients in the intervention group received proactive geriatric consultation, which began preoperatively or within 24 hours of surgery. They received targeted recommendations based on a structured protocol from the geriatrician during the period of hospitalization. Patients in the control group received usual care. They received management by the orthopaedics team, including internal medicine consultants or geriatricians on a reactive rather than proactive basis.

The structured protocol used for the recommendations included 10 modules with each containing two to five specific recommendations (Appendix J). Recommendations were prioritized and limited to no more than five after the initial visit by the geriatrician and no more than three after follow-up visits. This was done to improve adherence. The GDG suggested that the geriatrician and other NHS personnel would be needed to apply this intervention on patients. It was suggested that modules one to four, eight, and 10 would be delivered by doctors. This will require additional 15 minutes of geriatrician's time per patient per week. The duration of application of this intervention was taken to be equivalent to the median length of stay of patients with fracture of neck of femur which was reported as 16 days (HES Online, 2007 – 2008). It will therefore cost an additional £100 to apply the four modules. The application of modules five to seven, and module nine were assumed to be part of the routine work for nurses on pay Band 5. However, additional work and NHS resources would be expected for applying module 6a and 7b. The additional time for applying module 6a was suggested to be ten minutes thrice daily per patient while module 7b would require ten minutes four times daily per patient. The hourly cost of a nurse pay Band 5, including cost of qualification, is £22 [PSSRU 2007]. The application of module 6a would cost £11 per patient daily and module 7a would cost £15 per patient daily. This is equivalent to £176 and £235 respectively over 16 days. The total cost of applying multicomponent targeted intervention to older patients admitted non-electively for surgical repair of hip fracture would therefore amount to £511.

The use of multicomponent targeted intervention in consecutive older patients at intermediate or high risk of delirium who were admitted to the general medicine service

The cost estimate for using multicomponent targeted intervention in older patients at intermediate or high risk of delirium who were admitted to the general medicine service was based on a trial of patients aged 70 years or older who were consecutively admitted to the general medicine service of a hospital (Inuoye 1999). This trial has been described in the section on the use of multicomponent interventions for delirium prevention (section 10.19). At the point of admission, the patients in the trial showed no evidence of patients having delirium, but they were assessed to be at immediate or high risk of developing delirium. The study sample was 852 people, including 426 matched pairs of intervention and control, enrolled in the clinical trial in a hospital between March 1995 and March 1998. The trial had three aims namely, to compare the effectiveness of a multicomponent strategy for reducing the risk of delirium with that of a usual plan of care for hospitalized older patients, to determine the level of adherence to the intervention protocol, and to measure the effect of the intervention on the targeted risk factors. Eligible study patients underwent screening and base line assessments which were completed within 48 hours after admission. Patients in the intervention group received standard protocols for the management of six risk factors for delirium namely, cognitive impairment, sleep deprivation, immobility, visual impairment, hearing impairment, and dehydration (Appendix J). Geriatric nursing assessment and interdisciplinary rounds were other program interventions targeted towards the risk factors. The intervention, the Hospital Elder Life Program, was implemented by a trained team, which consisted of a geriatric nurse-specialist, two specially trained Elder Life specialists, a certified therapeutic-recreation specialist, a physical-therapy consultant, a geriatrician, and trained volunteers. Patients in the usual care group received standard hospital services provided by physicians, nurses, and support staff. The study reported the total cost of intervention to be $139,506. The number of people in the intervention group was 426 and the average cost of intervention was reported as $327 per patient. This included staff time spent in intervention activities, equipment, supplies and consultant costs.

It was recommended that the staff required to implement the Hospital Elder Life Program in 200 to 250 patients per year are one full-time Elder Life Specialist who also serves as Volunteer Coordinator, one half-time Geriatric Nurse Specialist, and 0.10 to 0.20 of a full time equivalent geriatrician, who also acts as a Program Director (Inouye, 2000). We used this time equivalence in our cost estimation. A description of the duties of each staff is given in Appendix J. Volunteers play a critical role in the implementation of the program and the tasks of a volunteer would be carried out by NHS personnel. It was suggested that a minimum of 21 Volunteers would be required to operate a program of 200 to 250 patients. Each was to serve one shift per week and 3 to 4 hours per shift. The GDG advised that the pay band for the geriatric nurse specialist would be Band 6; Elder Life specialist would be Band 5; Geriatrician would be the annual salary equivalent of an NHS Medical Consultant and the Volunteer would be Band 2. We applied the Agenda for Change salaries and used the April 2006 scale mid-point. These were used to estimate the unit cost for the Elder Life Program Staff. We estimated that the personnel cost per patient would be £370. We assumed that each of the 21 volunteers would work four hours per week, geriatricians would work 0.15 Full Time Equivalence and the number of patients that received intervention would be 225 patients.

Equipment such as computers, telephone and photocopying machines that would be needed to implement the program are assumed to be available and would not need to be purchased additionally by the NHS. Some of the materials needed for implementing the intervention protocol described in the study by Inuoye (1999) are already available to the NHS patient and are used during usual care. The additional materials that would need to be purchased are listed in Appendix J. They include standard word games and relaxation tapes or music. We have assumed that cost of providing instructions by the intervention staff will be accounted for through the salary paid to them by the NHS. We have not added any additional cost of providing instructions.

We could not find cost data on what the NHS pays for a standard word game or relaxation tapes. We have assumed the cost to be £50 each and life expectancies of the materials to be 0.5 and 1 year respectively. We have also assumed that 10 pieces of relaxation tapes will be required for a multicomponent targeted intervention program for 225 patients over a year. We assumed that 20 pieces of standard word game will be required for the same number of patients over the same time period. The additional cost of the materials was estimated at £7 per patient.

We have estimated the cost of using multicomponent targeted intervention in older patients at intermediate or high risk of delirium who were admitted to the general medicine service in the NHS as £377. This does not include additional training cost as we have assumed that this has already been included as part of the time resources required by the Program staff to implement the program. We also did not include the cost associated with screening and base line assessment at the beginning of the intervention for the same reason. In a sensitivity analysis, we assumed that the Geriatric nurse specialist will be on band 7 and the Elder Life Specialist, on band 6. This increased the total cost of personnel to £404. This was to account for possible additional work load for these two roles.

A summary of the data inputs used in the model is given below. The baseline and relative risk estimates of the adverse consequences have been given above in table 16.2.

Table 16.2: other inputs used in base case analysis in the economic model

Model input	Point Estimate (95% CI)	Source
Baseline risk		
Delirium in hospital (general medicine services)	15.0%	Inouye 1999
Delirium in hospital (hip fracture surgery)	50.0%	Marcantonio 2000
Unit cost		
New dementia (per year)	£16,302	Dementia UK, The full report, 2007
Stay in long-term care (per week)	£656	PSSRU 2007, Netten 1998
Pressure ulcer	£1,364 (£1,228 to £1,500)*	Bennett 2004
Falls	£1,875	Iglesias 2008
Utility		
New dementia	0.29	Ekman 2007 (reported 0.4 for moderate dementia)
New admission to institution	0.18	Ekman 2007 (reported 0.25 for moderate dementia, GDG suggested it should be used to estimate utility for this outcome)
Falls	0.69	Hendriks 2008 (reported 0.71 after 12 months)
Duration		
Stay in long-term care (months)	18.9	Netten 2001
Extended hospital stay (days)	16.83 (9.36, 25.34)	Holmes & House 2000
Life with dementia (years)	1.2	McNamee 2001
Intervention Efficacy		
MTI (general medical services)	0.66 (0.46, 0.95)	Inouye 1999
MTI (hip fracture surgery)	0.65 (0.42, 1)	Marcantonio I 2000
Intervention Cost		
MTI (medical services)	£377	Based on study protocol in Inouye 1999

Model input	Point Estimate (95% CI)	Source
MTI (hip fracture surgery)	£511	Based on study protocol in Marcantonio 2000

*Reported as mean (+ and – 10%)

16.2.8 Sensitivity Analyses

Deterministic Sensitivity Analyses

In the deterministic analysis we estimated the point estimate for cost, QALYs gained, ICER and INMB using the base case model structure and point estimates for model input parameters. We have carried out a series of deterministic sensitivity analyses (DSA) to explore the uncertainties that relate to the base case structure.

The first approach we have taken is to assume that not all the adverse consequences are important to the model structure. We assumed that each and only one of the six adverse consequences was the only adverse outcome associated with delirium. We estimated the INMB after assuming that new admission to nursing homes was the only adverse outcome to be associated with delirium. The same was done for mortality, new dementia, falls, pressure ulcer and extended hospital stay. In another DSA, we included nursing home admission and mortality as a composite outcome and did not include them as single model inputs. We explored the cost-effectiveness of interventions in low risk patients and used 12.5% as the baseline risk of delirium. This was the lower estimate of the range of delirium incidence reported in the needs assessment review (chapter 5) for general medical patients. We explored the effect of using this lower estimate for both populations considered by the model (elderly patients at risk of delirium who were admitted to the general medicine service and patients undergoing surgical repair of hip fracture).

In the base case analysis, we have assumed that the life expectancy of delirious patients to be shorter than that of non-delirious patients. This was due to difference in post-hospital chances of survival for the two groups. In a DSA we have assumed that the survival chances for delirious patients are equivalent to those of non-delirious patients. In another DSA we have assumed the life expectancy of dementia patients to be 3.6 years and 5.4 years for patients with and without previous delirium experience respectively.. In the base case, we used 1.2 years regardless of the previous delirium experience. We have assumed in the base case that patients in long-term care will survive for only 18.9 months. In a sensitivity analysis, we estimated lifetime QALY gains over a life expectancy of 3.6 years for those with delirium and 5.4 years for those without delirium.

The annual cost of dementia was reduced to £5,859. This was to remove potential double counting of the cost of stay in long-term care as a proportion of the cost of dementia in the base case was due to stay in long-term care. In another DSA, we included only 70% of the cost of stay in long-term care, as we assumed that 100% of this cost will not be funded by the public. Further analyses were done to explore the impact on the model results of increased cost of pressure ulcer resulting from grade 4 ulcer that heal normally, and increased cost of the multicomponent targeted interventions resulting from higher pay Band to the Geriatric Nurse Specialist and Elder Life Specialist.

Probabilistic Sensitivity Analyses

In the DSA we used point estimates for the model input parameters. However, point estimates are subject to uncertainties. We have carried out a probability sensitivity analysis, PSA, to reflect the uncertainty in the input parameters of the model. The results of the PSA show the uncertainty in the primary outcomes of the model that results from the uncertainty in the model inputs. Each of the input parameters is assigned a probability distribution which reflects the standard error of each parameter estimate.

We randomly selected from each parameter distribution in a simultaneous manner and calculated the cost, QALYs, ICERs and INMB. This was repeated 5000 times to produce 5000 estimates that reflect the uncertainties in the input parameters. An average of the estimates was found and the most cost-effective strategy is the one with the highest mean INMB. However, the one with the highest mean INMB may or may not be the most cost-effective in all the simulations. The model parameters, the type of distribution and distribution parameters are listed in the table below (table 16.3). The model input parameters that we did not vary probabilistically are life expectancy of a patient with dementia, survival length of time in long-term care, post-discharge mortality differences for delirious and non-delirious patients, and the discount rate.

Table 16.3: input parameters, type of distribution and distribution parameters used in the probabilistic sensitivity analysis

Parameter	Type of distribution	Point estimate	Distribution parameters	Source
Baseline Risk				
Delirium in Hospital (general medical services)	Beta	15.0%	$\alpha = 64$, $\beta = 362$	Inouye 1999
Delirium in Hospital (hip fracture surgery)	Beta	50.0%	$\alpha = 32$, $\beta = 32$	Marcantonio 2000
Falls	Beta	6.9%	$\alpha = 9$, $\beta = 122$	O'Keeffe & Lavan 1997
Pressure Ulcer	Beta	4.0%	$\alpha = 360$, $\beta = 8575$	Clark & Watts 1994
Dementia	Beta	5.6%	$\alpha = 7$, $\beta = 117$	Rockwood 1999
New admission to institution	Beta	17.4%	$\alpha = 40$, $\beta = 190$	Bourdel-Marchasson 2004
In hospital Mortality	Beta	5.0%	$\alpha = 7$, $\beta = 124$	O'Keeffe & Lavan 1997
Mortality or new admission to institution	Beta	12.2%	$\alpha = 9$, $\beta = 65$	Marcantonio 2000
Post-discharge survival				
Difference in mortality	Lognormal	HR = 1.71	Log (mean) = 0.54,	Rockwood 1999

Parameter	Type of distribution	Point estimate	Distribution parameters	Source
between delirious and non-delirious patients			SE = 0.26	
Relative Risk				
Falls and pressure ulcer	Lognormal	RR = 2.18	Log (mean) = 0.78, SE = 0.27	O'Keeffe & Lavan 1997
Dementia	Lognormal	RR = 4.67	Log (mean) = 1.54, SE = 0.60	Rockwood 1999
New admission to institution	Lognormal	RR = 2.05	Log (mean) = 0.72, SE = 0.59	Bourdel-Marchasson 2004
Mortality	Lognormal	RR = 2.41	Log (mean) = 0.88, SE = 0.56	O'Keeffe & Lavan 1997
Mortality or new admission to institution	Lognormal	RR = 2.41	Log (mean) = 0.88, se = 0.52	Marcantonio 2000
Cost				
Falls	Gamma	£1,875	Mean = £1,875, SE* = £239	Iglesias 2008
Pressure Ulcer	Gamma	£1,364	Mean = £1,364, SE = £69	Bennett 2004
Dementia	Gamma	£16,302	Mean = £16,302, SE *= £2079	Dementia UK, The Full Report, 2007
Extended hospital stay	Gamma	£152	Mean = £152, SE* = £19	HES England, 2007-08
Stay in long-term care	Gamma	£656	Mean = £656, SE* = £84	PSSRU 2007
MTI (general medical)	Gamma	£377	Mean = £377, SE* = £48	Based on recommended protocol and GDG advice
MTI (hip fracture surgery)	Gamma	£511	Mean = £511, SE* = £65	Based on recommended protocol and GDG advice
Utility				
Falls	Beta	0.71	α =249 , β = 102	Hendriks 2008
Dementia	Beta	0.40	α = 730, β = 1094	Kman 2007
Stay in institution	Beta	0.25	α = 293, β = 880	Ekman 2007
Population utility	Multinormial	Linear relationship with age	Age-Utility intercept: 1.06; Age-Utility gradient: -0.00	Based on a re-analysis of data from Kind 1998 in Ward 2007
Duration				
Extended hospital stay	Gamma	16.83	Mean = 16.83, SE = 4.08	Holmes and House 2000
Efficacy of MTI intervention				
Relative risk (general medicine services)	Lognormal	0.66	Log (mean) = -0.42, SE = 0.19	Inouye 1999
Relative risk (hip fracture surgery)	Lognormal	0.65	Log (mean) = -0.43, SE = 0.22	Marcantonio 2000

*Assumed that upper and lower confidence intervals will be 125% and 75% of mean estimate respectively.

16.2.9 Results

Cost-effectiveness of multicomponent targeted prevention interventions in older patients at intermediate or high risk of delirium who were admitted to the general medicine service

The table below (table 16.4) shows the cost-effectiveness model results for the use of multicomponent targeted prevention interventions in patients at immediate or high risk of delirium and who were admitted to the general medicine service. The result of the deterministic analysis suggests that this intervention is cost-effective when compared to usual care and is associated with an incremental net monetary benefit (INMB) of £2,130.

The result of the probabilistic sensitivity analysis (PSA) suggests that the usual care strategy will cost £13,200 on average whereas the prevention strategy will cost £12,690. This is the mean total cost that includes the cost of the adverse consequences and the unit cost of the intervention itself. The QALY gains associated with both strategies are 2.140 and 2.220 QALYs respectively. The prevention strategy was therefore the dominant strategy because it reduced cost and increased QALY gains when compared to the usual care strategy. It was associated with an INMB of £2,200.

Table 16.4: costs, QALYs and cost-effectiveness of multicomponent targeted intervention compared to usual care*

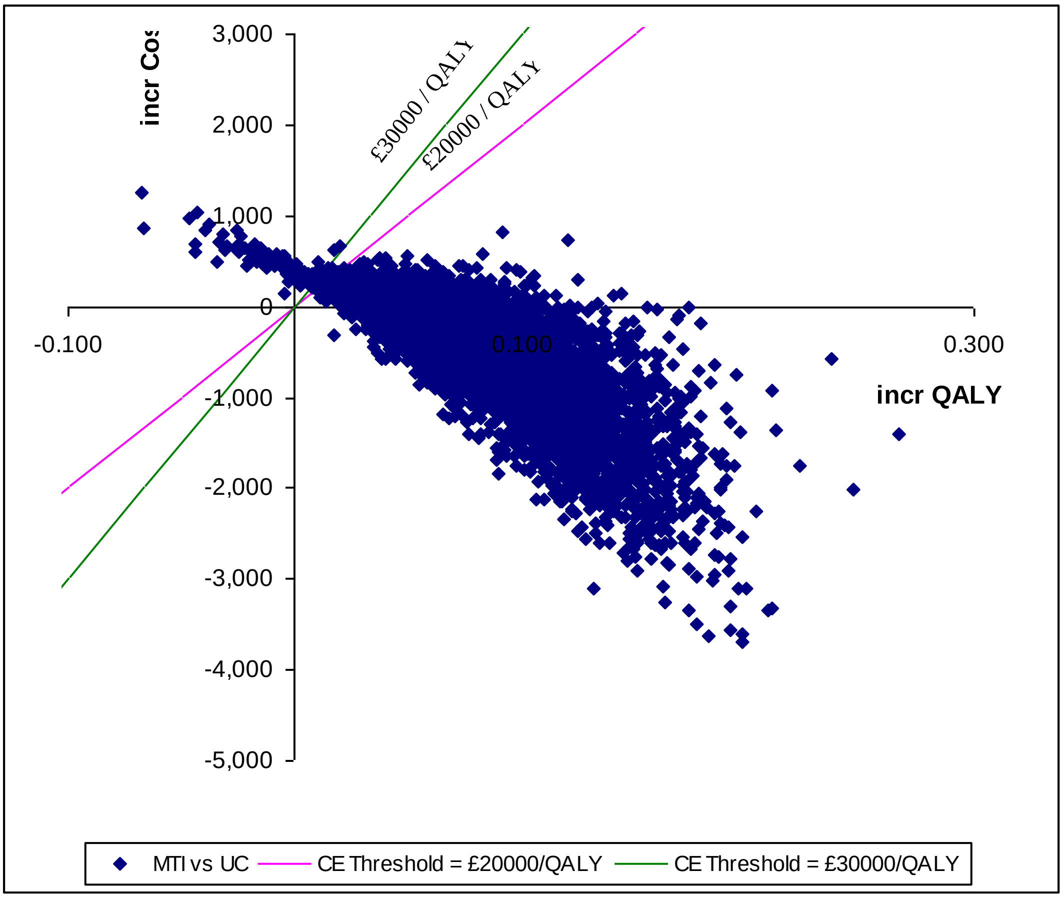
		Usual Care	MTI
Probabilistic	Mean cost	£13,200	£12,690
	Mean QALYs	2.140	2.220
	Incremental Cost	N/A	-£520
	Incremental QALYs		0.084
	Incremental Cost per QALY gained		MTI dominates
	Incremental net monetary benefit (INMB)**		£2,200
	% of simulations where strategy was most cost-effective**	3%	97%
Deterministic	Incremental net monetary benefit (INMB)**	N/A	£2,130

*Costs and QALYs are mean total costs and QALYs across 5000 PSA simulations

** Using a cost-effectiveness threshold of £20,000 per QALY gained.

At a cost-effectiveness threshold of £20,000 per QALY, the prevention strategy was associated with a higher INMB estimate and was more cost-effective in 96.8% of the simulations that were run in the PSA. In 1.5% of the simulations, the intervention strategy increased cost and reduced QALY gains (figure 16.2). The INMB was £3,040 at a cost-effectiveness threshold of £30,000 per QALY

Figure 16.2: cost-effectiveness plane for multicomponent targeted intervention compared to usual care



The results of the one-way deterministic sensitivity analyses are presented in table 16.5. The use of the prevention intervention remained cost-effective for the majority of the DSA. The only exceptions were when we assumed that pressure ulcer, falls, in-hospital mortality and extended hospital length of stay were the only adverse outcome associated with delirium. In these cases the intervention was not cost-effective. The intervention remained cost-effective when we excluded the survival difference between delirious and non-delirious cases, removed the cost of dementia attributable to stay in long-term care, increased the cost of pressure ulcer. The INMB was £2330 when the life expectancy of dementia was increased from 1.2 years to 3.6 and 5.4 years for dementia patients with and without delirium respectively, An explanation for a higher INMB even when the survival implications of dementia are less severe is that the additional cost of dementia incurred in additional life years more than off-sets the additional health benefits due to increased life expectancy. In further analyses, we used the composite outcome of new admission to institution and mortality, and assumed that the NHS and PSS would pay only 70% of the cost of stay in long-term care but the intervention remained cost-effective.

Table 16.5: other deterministic sensitivity analyses on the cost-effectiveness of multicomponent targeted intervention compared to usual care

	Incremental NMB (deterministic)
All model parameters (base case)	£2,125
Baseline risk of delirium = 12.5% (base case = 15%)	£1,710
In hospital mortality is the only consequence of delirium	-£140
New dementia is the only consequence of delirium	£440
New admission to nursing home is the only consequence of delirium	£660
Falls is the only consequence of delirium	-£210
Pressure ulcer is the only consequence of delirium	-£370
Extended hospital stay is the only consequence of delirium	-£250
Including 3-year survival difference between delirious and non-delirious patients (as the only adverse outcome in model)	£670
Excluding 3-year survival difference between delirious and non-delirious patients (but including all adverse consequences)	£2009
Excluding the cost of dementia attributable to stay in long-term care (cost of dementia = £5859) (base case = £16,302)	£1994
Life expectancy for dementia patients with previous delirium = 3.6 years, without previous delirium, 5.4 years (base case = 1.2 years)	£2,330
QALY gain for stay in long-term care over life expectancy of 3.6 years for patients with previous delirium and 5.4 years for those without	£2,110
Cost of pressure ulcer using the cost of grade 4 ulcer that heals normally	£2,150
Baseline risk of pressure ulcer = 1.68%	£2,120
Accounted for only 70% of cost of stay in long-term care	£1980
Composite outcome, mortality and new admission to institution	£1980
Increased pay band for Geriatric Nurse (Band 7) and Elder Life Specialist (Band 6)	£2090

Cost-effectiveness of multicomponent targeted prevention interventions in older patients admitted non-electively for surgical repair of hip fracture

The use of multicomponent targeted prevention interventions in older patients admitted non-electively for surgical repair of hip fracture resulted in an incremental net monetary benefit (INMB) of £8070 (table 16.6). In the probabilistic sensitivity analysis (PSA), the mean total cost of the usual care strategy and prevention strategies in this population were estimated as £19,530 and £17,040 respectively. The mean QALYs were 1.540 and 1.820 respectively. The intervention strategy reduced cost by £2,490 and increased QALY gain by 0.290. It therefore dominates the usual care strategy. The INMB for this intervention strategy compared to the usual care strategy was £8,180.

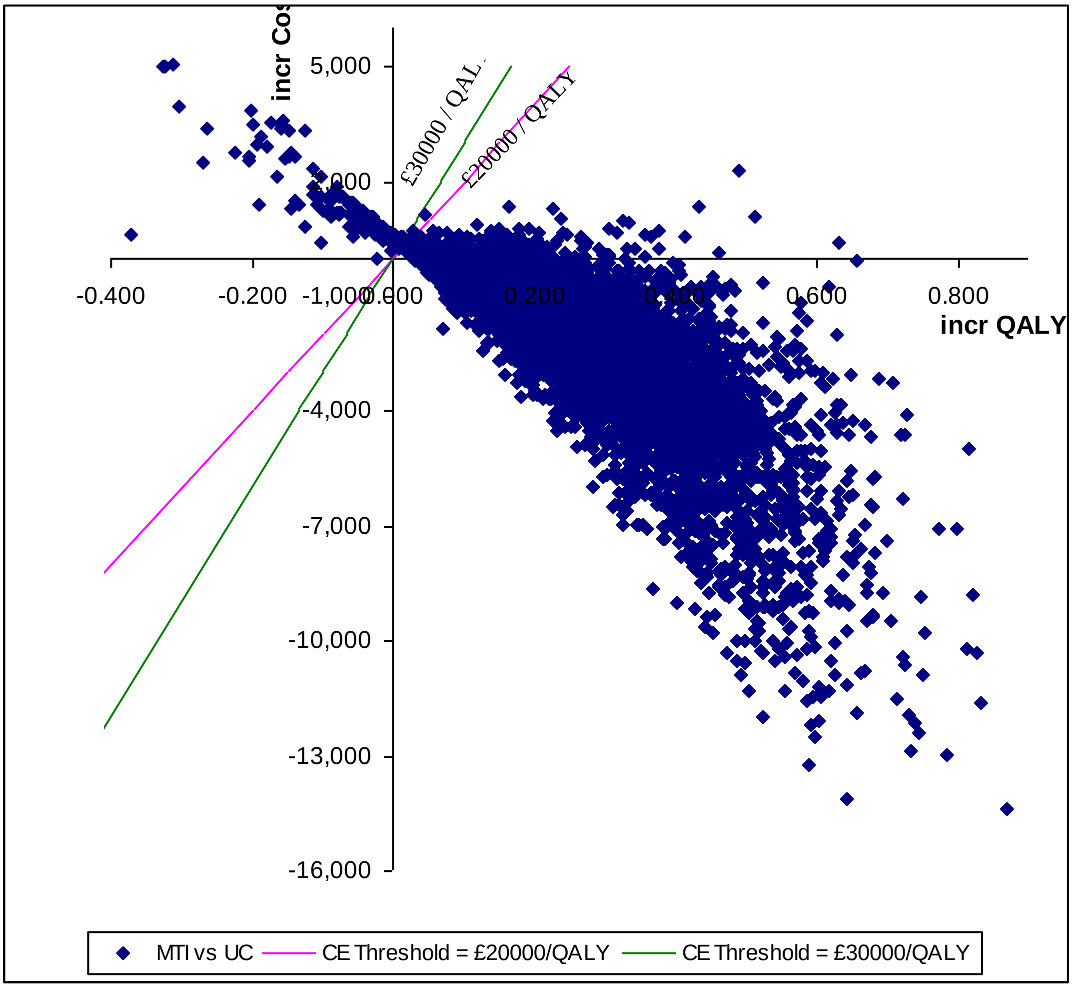
Table 16.6: costs, QALYs and cost-effectiveness of multicomponent targeted intervention compared to usual care

		Usual Care	MTI
Probabilistic	Mean cost	£19,530	£17,040
	Mean QALYs	1.540	1.820
	Incremental Cost	N/A	-£2,490
	Incremental QALYs		0.290
	Incremental Cost per QALY gained		MTI is dominant
	IIncremental net monetary benefit (INMB)*		£8,180
	% of simulations where strategy was most cost-effective*	4%	96%
Deterministic	Incremental net monetary benefit (INMB)*	N/A	£8,070

* Using a cost-effectiveness threshold of £20,000 per QALY gained.

At a cost-effectiveness threshold of £20,000 per QALY, the prevention strategy was more cost-effective in 96.4% of the simulations that were run in the PSA. The intervention strategy increased cost and reduced QALY gains in 2.8% of the simulations (figure 16.3). The INMB was £11,030 at a cost-effectiveness threshold of £30,000 per QALY

Figure 16.3: cost-effectiveness plane for multicomponent targeted intervention compared to usual care



This intervention strategy remained cost-effective in most of the DSA conducted (table 16.7). The exceptions were when we assumed that pressure ulcer and extended hospital length of stay were the only adverse outcome associated with delirium. The intervention was not cost-effective in these cases. When the life expectancy of dementia was increased to 3.6 and 5.4 years for dementia patients with and without delirium respectively, the INMB was higher than the INMB in base case. In this case, the additional cost of dementia incurred in additional life years more than off-sets the additional health benefits due to increased life expectancy.

Table 16.7: other deterministic sensitivity analyses on the cost-effectiveness of multicomponent targeted intervention compared to usual care

	Incremental NMB (deterministic)
All model parameters (base case)	£8,074
Baseline risk of delirium = 12.5% (base case = 50%)	£1,640

In hospital mortality is the only consequence of delirium	£290
New dementia is the only consequence of delirium	£2,270
New admission to nursing home is the only consequence of delirium	£3,060
Falls is the only consequence of delirium	£60
Pressure ulcer is the only consequence of delirium	-£500
Extended hospital stay is the only consequence of delirium	-£62
Including 3-year survival difference between delirious and non-delirious patients (as the only adverse outcome in model)	£3,070
Excluding 3-year survival difference between delirious and non-delirious patients (but including all adverse consequences)	£7,670
Excluding the cost of dementia attributable to stay in long-term care (cost of dementia = £5859) (base case = £16,302)	£7,630
Life expectancy for dementia patients with previous delirium = 3.6 years, without previous delirium, 5.4 years (base case = 1.2 years)	£8,760
QALY gain for stay in long-term care over life expectancy of 3.6 years for patients with previous delirium and 5.4 years for those without	£8,030
Cost of pressure ulcer using the cost of grade 4 ulcer that heals normally	£8,150
Baseline risk of pressure ulcer = 1.68%	£8,070
Accounted for only 70% of cost of stay in long-term care	£7,570
Composite outcome, mortality and new admission to institution	£7,590

16.3 The treatment model

16.3.1 The model structure for the treatment interventions

Decision Tree

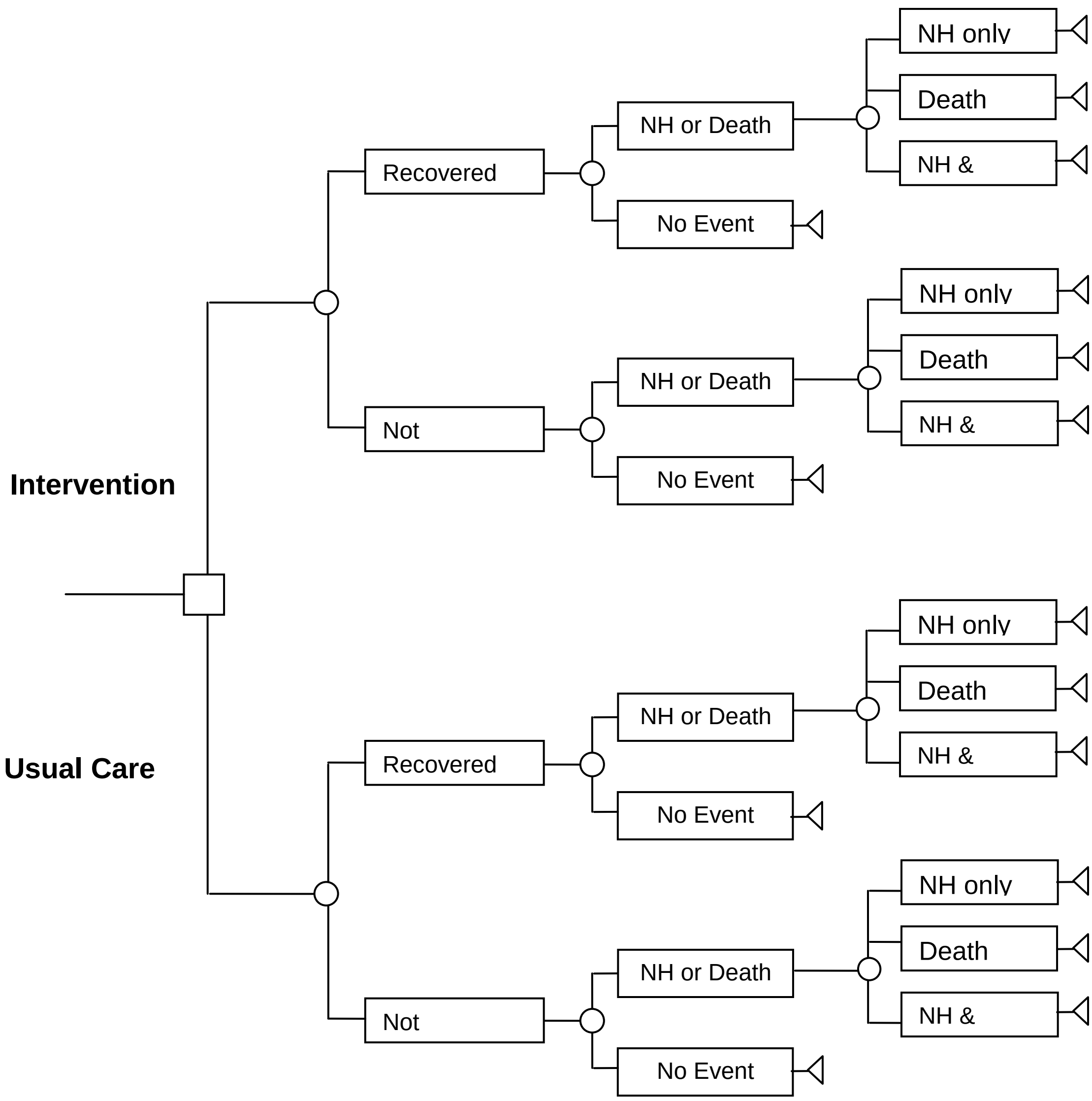
A change in the duration and severity of delirium through treatment will unlikely lead to a QALY gain. However, treatment will reduce the cost and QALY loss associated with adverse consequences that will occur in delirious patients. In the systematic review of the treatment strategies, there were no data on the direct effect of treatment on the adverse consequences used in the prevention model above. There were data on intermediate outcomes and we had to use an intermediate outcome to link the effect of treatment with adverse delirium consequences. The GDG advised that we use “complete recovery from delirium” as the intermediate outcome in the model. Data were reported in the adverse consequences review on the increased risk of nursing home admission and death for patients without complete recovery.

The treatment cost-effectiveness model consists of a decision tree (figure 16.4). In the usual care arm of the tree, the members of a cohort of patients with delirium will either recover completely or not recover at all. The number of people recovering will depend on the baseline risk of recovery in a care setting. Regardless of their recovery status some of them will have no further adverse event and others will be admitted to the nursing home or will die. Those that experience further adverse event will either experience admission to nursing home only, death only or both. The number of people that experience any of these three outcomes will depend on the baseline risk of these outcomes in the care setting. In the treatment arm, it will depend on the baseline risks as well as the relative risk of complete recovery if exposed to the treatment.

The GDG advised that we consider the impact of treatment side effects in the model. A review of the adverse effects of antipsychotic agents suggests that the only useful evidence for the existence of side effect is for stroke. It was therefore the only side effect that was considered in the model. We carried out a sensitivity analysis where stroke was included as one of the branches of the decision tree.

The end of each branch of the tree implies a particular cost and a particular QALY. The total cost and QALYs are summed up for each strategy.

Figure 16.4: decision tree for treatment intervention strategies



16.3.2 Absolute Risk Estimates

Complete recovery

The baseline risk of complete recovery was taken from the Hu (2006) study and this study has been described in details in the section on review of hospital treatment using pharmacological interventions (chapter 13). It was reported in the control arm of the study that five out of a total of 29 people experienced complete recovery. We therefore used 17.2% as the baseline risk of complete recovery.

Admission to nursing home or death

The baseline risk of "nursing home admission or death" for patients that recovered as well as those that did not recover were taken from the McAvay study (2006) which has been described in the section on adverse consequences review (chapter 9). The study compared 1-year institutionalization and mortality rates of patients who were delirious at discharge, patients whose delirium resolved by discharge, and patients who were never delirious in the hospital. Twenty one out of 31 of patients (67.7%) whose delirium resolved at discharge experienced "death or nursing home placement".This equates to a hazard rate of 1.13 using the formula $-(1/t)\ln\{s/n\}$, where "a" is the hazard rate, and "t" is time (= 1 year), "n" is the sample size at baseline (= 31) and "s" is number of those surviving to time "t" without an event (= 10). An adjusted hazard ratio of 1.73 was reported for "nursing home admission or mortality" for patients who had delirium at discharge compared to those whose delirium resolved. We used this adjusted hazard ratio to estimate the risk of "nursing home admission or mortality" for patients who had delirium at discharge by assuming that the hazard was constant over time. We got a hazard of 1.96 (1.13 X 1.73) for patients with delirium at discharge. This gave a 1-year risk of 85.9% using the formula:

$p=1-\exp\{-at\}$, where "p" is probability of nursing home admission or mortality", "a" is the hazard, and "t" is time. The McAvay (2006) study also reported data which we used to estimate the proportion of people with death only, nursing home admission only, and "nursing home admission and death" for patients whose delirium resolved as well as those whose delirium did not resolve. For those whose delirium resolved, the proportion of people with nursing home admission only, death only, and "nursing home admission and death" was estimated as 61.9%, 33.3% and 4.8% respectively. For those whose delirium did not resolve, this was estimated as 55.0%, 5.0% and 40.0% respectively.

Stroke

We took the baseline risk of stroke from Wooltorton (2002) who reported an analysis of drug manufacturer's trials involving elderly patients with dementia. Wooltorton (2002) reported that in four placebo-controlled trials lasting one to three months and involving more than 1200 patients with Alzheimer's disease or vascular dementia, cerebrovascular adverse events were twice as common in the risperidone treated group as in the placebo group. Risperidone is an atypical antipsychotic and cerebrovascular adverse events were reported to include stroke and transient ischemic attacks. In the placebo arm, it reported that seven out of 466 patients experienced this adverse event. We have therefore used 1.5% as the baseline risk of stroke in our model.

Efficacy of Treatment Interventions

The efficacy of different antipsychotic drug treatment interventions has been reviewed in chapter 13. The two drugs that were identified to be clinically effective are haloperidol and olanzapine, and we have included only these two in our model. Haloperidol and olanzapine were estimated to have relative risk of complete recovery of 3.95 and 3.68 respectively.

Relative risk of stroke as side effect of antipsychotic drugs

The relative risk of stroke following the administration of antipsychotic agents has been reviewed in chapter 14. We used the data from the Douglas and Smeeth (2008) study which reported the relative risk of stroke for all antipsychotics compared to no treatment (RR=1.73); typical antipsychotic compared to no treatment (RR=1.69); and atypical antipsychotic compared to no treatment (RR=2.32). In the base case cost-effectiveness analysis we have not included stroke as a side effect of using antipsychotic agents. In a sensitivity analysis we have included an increased risk of stroke using the relative risk for all antipsychotics compared to placebo. In a second sensitivity analysis we have used the relative risks reported specifically for haloperidol and olanzapine.

16.3.3 Cost and QALYs of Outcomes on the decision tree

Nursing home admission

The estimates of unit cost and duration of stay in long-term care are the same as the estimates used above in the prevention model. The unit cost of stay in long-term care is £656 per week and the duration of stay is 18.9 months. The expected lifetime QALY gain for this outcome has been estimated the same way it was estimated in the base case of the prevention model.

Death only

The mortality risk was taken from a study (McAvay 2006) which reported this risk in patients followed up for one year post-hospital discharge. We have assumed that the patient with this outcome will live for six months before death and we have estimated a QALY again for a 79 year old person who lived for just six months. We have also assumed that mortality will be associated with zero cost.

Nursing home admission and death

The cost of this outcome was estimated as a product of the unit cost of stay in long-term care and the duration of stay. The duration of stay was assumed to be six months only after which the patient dies. The expected lifetime QALY gain was estimated in a similar way as it was done in the prevention model. The only difference is that it was estimated over a period of six months. We used the same adjusted utility score of 0.18 and the way this was estimated has been described above.

Nil Event

For the nil event arm of the decision tree we have assumed that patients will not experience any death in the first year. Their survival from the second year was estimated to reflect the increased risk of mortality for persons with delirium. Adjusted mortality risk was estimated from data from the Rockwood (1999) study and applied in the prevention model for three years. In the treatment model, we have applied the adjusted increased mortality risk for only 2 years. The life expectancy of a patient without any event was estimated to be 5.29 years and the QALYs was estimated as 3.24.

Stroke

<u>Cost</u>

The cost of stroke in the first year was taken from a cost-effectiveness analysis that compared different models of stroke care provided in London and Copenhagen (Grieve 2000). In the Copenhagen centre, acute and rehabilitation unit were combined, and patients could be transferred from the acute hospital for further inpatient rehabilitation at a separate hospital. In the London care centre, patients were usually admitted to general wards where they are treated by general medicine specialist, but could be transferred to a rehabilitation stroke unit where geriatricians led care. Further rehabilitation as an inpatient at a separate hospital was not an option. A range of community services including further rehabilitation and support services were available in both centres.

The study participants were first-ever stroke patients and resource use was recorded one year post stroke. Measurement of resource use took a hospital and community health perspective and covered primary hospital stay, subsequent transfer to other hospital, readmissions, institutional care and use of outpatient and community health services. Data was collected on the use of diagnostic investigations, the length of stay by ward type, and doctors' and nurses' time resources. The amount of therapy each patient received was recorded as well as the length of stay in institutions.

A standard costing method was reported to have been used in costing inpatient services. The costs for institutional and community services were based on interviews undertaken with providers, and the median cost of the item concerned was used as the unit cost. The cost of a GP consultation came from PSSRU (Netten & Dennet1996) and the same methodology was applied to cost a consultation in Copenhagen. Disaggregated costs for surgery were not available for the London centre and were based on costs of surgery in Copenhagen. A factor of 0.74 was used to multiply the costs of surgery in Copenhagen to obtain surgery costs in London, and the factor was taken from the ratio of costs per hospital day between the centres. Costs were estimated in 1995 local prices but were converted into dollars using the purchasing power parity index.

In the London centre, 358 patients were included in the study but 20 were excluded from the main analysis because of missing case severity data. Most patients were admitted to a general medical ward and after an average stay in the initial area of 8 days, 26% were subsequently transferred to the rehabilitation stroke unit, and 6% were readmitted to hospital. The mean total length of all hospital stay in the year following stroke was reported as 35.3 days. On average, there were 3.9 visits to day centre, and the mean length of days spent in sheltered, residential and nursing homes were 8.1, 8.5 and 16.9 respectively. The mean cost of care in the year following stroke in London was reported as $8,825. We converted this to £5,643 using the PPP index for the year 1995 and up rated the converted estimate to £8,486 using the PSSRU pay and price indices of 166 for 1995/96 and 256.9 for 2007/08. We applied in our model £8,486 as the cost of care following first year of stroke.

For the cost of care in subsequent years we required information on the life expectancy of a stroke patient as well as the yearly cost. We took the yearly cost from the NICE Stroke guideline (Stroke: NICE clinical guideline 68, 2008). Dependent and independent stroke were reported to cost £11,292 and £876 per patient per year for subsequent years respectively. These estimates were costs of inpatient care taken from health technology assessment reports and were largely determined by calculating total length of hospital stay after stroke and multiplying by the average cost of inpatient care. We assumed that 62% of the cases will be independent, 38% will be dependent and the life expectancy of a stroke patient is 4.7 years (Stroke: NICE clinical guideline 68, 2008). We estimated the yearly cost of stroke for subsequent years to be £4827.

<u>Utility</u>

The utility data for stroke was taken from the cross-sectional study by Lindgren 2008. The primary aim of the study was to assess the utility loss among stroke survivors at different time points following the stroke. The EQ-5D questionnaire was sent to 393 patients, divided into groups with three, six, nine and 12 months having passed since the stroke. The study patients had to be above the age of 18 and below the age of 76 years. This was done to avoid patients with a high degree of co-morbidities such as dementia. Furthermore, the sampling process aimed to identify at least 50 patients with ischemic stroke in each of the four groups listed above, and as many hemorrhagic strokes as were encountered. The study was conducted among stroke patients at six different centres that reported data to the Swedish national stroke register. The recruitment of patients was done consecutively at the study centres during a one month period. The questionnaire responses were converted to utility scores using the UK social tariff that were elicited with the time trade-off methodology. The utility scores for stroke were 0.65, 0.75, 0.63 and 0.67 for patients who have had stroke for 3, 6, 9 and 12 months respectively. The mean utility score for all patients was 0.67 and mean age of study population was 64.4 years. The QALY gain due to stroke was estimated using a utility multiplier and duration of 4.7 years. We estimated the utility multiplier, 0.85, as the ratio of the utility of 0.67, the mean utility score, and 0.79, the utility of a person aged 64.4 years old in the UK population. The starting age in the model is 79 years and we have used the utility multiplier to adjust the utility of an average person aged 79 years. The utility score for stroke that we used in the model was 0.62.

16.3.4 Cost of Treatment Interventions

Haloperidol

The costing of haloperidol is based on the oral dosage, 0.5 to 1mg every eight hours for up to five days. This is based on the dosage that was reported in the review of treatment interventions (chapter 13) for patients over 60 years. We have chosen this dosage as the starting age of our model is 79 years. The net price of 28-tab pack of haloperidol 500 micrograms is 91p (BNF 57, [http://bnf.org/bnf/bnf/current/3225.htm#this] *accessed on 19/08/09*]). Using an average of 0.75 mg thrice daily for five days will require 22.5 tablets. We have therefore used £0.73 as the cost of haloperidol in our model. We did not consider additional drug administration costs. In a sensitivity analysis we used the higher dosage of 2.5 to 5mg every eight hours for five days. This dosage was meant to be for patients less than 60 years old. The net price of 28-tab pack of haloperidol 5 mg is £3.87. Using 2.5 mg thrice daily for five days will cost £2.59 and we used this estimate in a sensitivity analysis.

Olanzapine

We have estimated the cost of olanzapine based on the oral dosage, 2.5 mg daily for up to five days. This dosage was reported for the treatment of patients over 60 years (chapter 13) and we have chosen this dosage in our base case analysis as the starting age of our model is 79 years. The net price of 28-tab pack of olanzapine 2.5 mg is £33.29 (BNF 57, [http://bnf.org/bnf/bnf/current/56912.htm#this], *accessed on 19/08/09*]). Using 2.5 mg daily for five days will require only five tablets and will cost £5.94. In a sensitivity analysis, we used the dosage of five mg daily for five days. This is the dosage for those less than 60 years old. This will require 10 tablets and will cost £11.89.

A summary of the input parameter estimates used in the model is in table 16.8 below.

Table 16.8: other inputs used in base case analysis in the economic model

Model input	Point Estimate (95% CI)	Source
Baseline risk		
Complete recovery	17.2%	Hu 2006
Stroke	1.5%	Wooltorton 2002
Absolute risk		
NH admission or death in patients with complete recovery	67.7%	McAvay 2006
NH admission or death in patients with delirium at discharge	85.9%	
Proportion of people with death only, nursing home admission only, and “nursing home admission and death”		
NH admission only in patients with complete recovery	61.9%	
Death only in patients with complete recovery	33.3%	
NH admission and death in patients with complete recovery	4.8%	
NH admission only in patients with delirium at discharge	55.0%	
Death only in patients with delirium at discharge	5.0%	
NH admission and death in patients with delirium at discharge	40.0%	
Unit cost		
Stay in long-term care (per week)	£656	PSSRU 2007, Netten 1998
Stroke (first year)	£8486	Grieve 2000
Stroke (subsequent years)	£4827	NICE clinical guideline on Stroke, CG68 (2008). Assumed that 38% of strokes cases are dependent and 62%, independent
Utility		
Stay in long-term care	0.18	Ekman 2007 (reported 0.25 for moderate dementia, GDG suggested it should be used to estimate utility for this outcome)
Stroke	0.62	Lindgren 2008 (reported 0.67 as mean utility score)
Duration		
Stay in long-term care (months)	18.9	Netten 2001
Life expectancy for stroke (years)	4.7*	NICE clinical guideline on Stroke, CG68 (2008)
Intervention Efficacy		
Haloperidol	3.95 (1.75, 8.9)	Hu 2006
Olanzapine	3.68 (1.63, 8.33)	
Intervention Cost		

Model input	Point Estimate (95% CI)	Source
Haloperidol	£0.73	BNF 57 (dosage for people over 60 years as stated in treatment review)
Olanzapine	£5.94	BNF 57 (dosage for people over 60 years as stated in treatment review)
Relative risk of stroke as a side effect of using antipsychotic agents		
All antipsychotic agents	1.73 (1.60, 1.87)	
Haloperidol	1.69 (1.55, 1.84)	Douglas and Smeeth 2008
Olanzapine	2.32 (1.73, 3.11)	

*Life expectancy for a patient without any event is 5.3 years

16.3.5 Sensitivity Analyses

As described previously for the prevention model, we have used a DSA to explore the importance of the various model assumptions and probabilistic sensitivity analysis to explore the impact of parameter uncertainty associated with the various model inputs. In the first DSA we included the impact of stroke in the model as this was not done in the base case analysis. We used the relative risk of 1.73 for both haloperidol and olanzapine in the first sensitivity analysis. In the second analysis, we used drug specific relative risk estimates (haloperidol = 1.69, olanzapine = 2.32).

One of the adverse consequences included in the model was nursing home admission and death. In the base case, we assumed that death will occur after the patient has spent six months in long-term care. In another DSA we assumed the patient will spend 12 months in long-term care. Further analysis was carried out by assuming that only 70% of the cost of long-term care will be publicly financed. The model parameters, the type of distribution and distribution parameters used in probabilistic sensitivity analysis (PSA) are listed in the table below (table 16.9).

Table 16.9: input parameters, type of distribution and distribution parameters used in probabilistic sensitivity analysis

Parameter	Type of distribution	Point estimate	Distribution parameters	Source
Baseline Risk				
Complete recovery	Beta	17.2%	α = 5, β = 24	Hu 2006
Absolute Risk				
NH admission or death in patients with complete recovery	Beta	67.7%	α = 21, β = 10	
NH admission or death in patients with delirium at discharge	Beta	85.9%	α = 9, β = 1	
NH admission only in patients with complete recovery	Dirichlet	61.9%	α = 13	McAvay 2006
Death only in patients with		33.3%	α = 7	

Parameter	Type of distribution	Point estimate	Distribution parameters	Source
complete recovery				
NH admission and death in patients with complete recovery		4.8%	α = 1	
NH admission only in patients with delirium at discharge	Dirichlet	55.0%	α = 11	
Death only in patients with delirium at discharge		5.0%	α = 1	
NH admission and death in patients with delirium at discharge		40.0%	α = 8	
Post-discharge survival				
Difference in mortality between delirious and non-delirious patients	Lognormal	HR = 1.71	Log (mean) = 0.54, SE= 0.26	Rockwood 1999
Cost				
Stay in long-term care	Gamma	£656	Mean = £656, SE* = £84	PSSRU 2007
Haloperidol	Gamma	£0.73	Mean = £0.73, SE*= £0.09	BNF 57
Olanzapine	Gamma	£5.94	Mean = £5.94, SE*= £0.76	BNF 57
Utility				
Stay in institution	Beta	0.25	α = 293, β = 880	Ekman2007
Population utility	Multinormial	Linear relationship with age	Age-Utility intercept: 1.06; Age-Utility gradient: -0.00	Based on a re-analysis of data from Kind 1998 in Ward 2007
Efficacy of treatment interventions				
Haloperidol	Lognormal	3.95	Log (mean) = 1.37, SE= 0.41	Hu 2006
Olanzapine	Lognormal	3.68	Log (mean) = 1.30, SE= 0.42	

*Assumed that upper and lower confidence intervals will be 125% and 75% of the mean estimate respectively.

16.3.6 Results

The costs, QALYs and cost-effectiveness estimates of the treatment model are presented in the table 16.10 below. In the deterministic base case analysis haloperidol and olanzapine were both cost-effective when compared to usual care. Haloperidol and olanzapine were estimated to have incremental (INMB) of £10,340 and £9,390 respectively. In the probabilistic sensitivity analysis (PSA), the mean total cost of the three treatment strategies, usual care, haloperidol and olanzapine were £31,120, £25,630, and £26,090 respectively. The mean total QALYs were 0.615, 1.035 and 1.004 respectively. The use of haloperidol or olanzapine reduced cost and increased QALYs when compared to usual care.

The INMB estimates were £13,900 and £12,820 respectively. Haloperidol dominates olanzapine because it saves more costs and generates more QALYs.

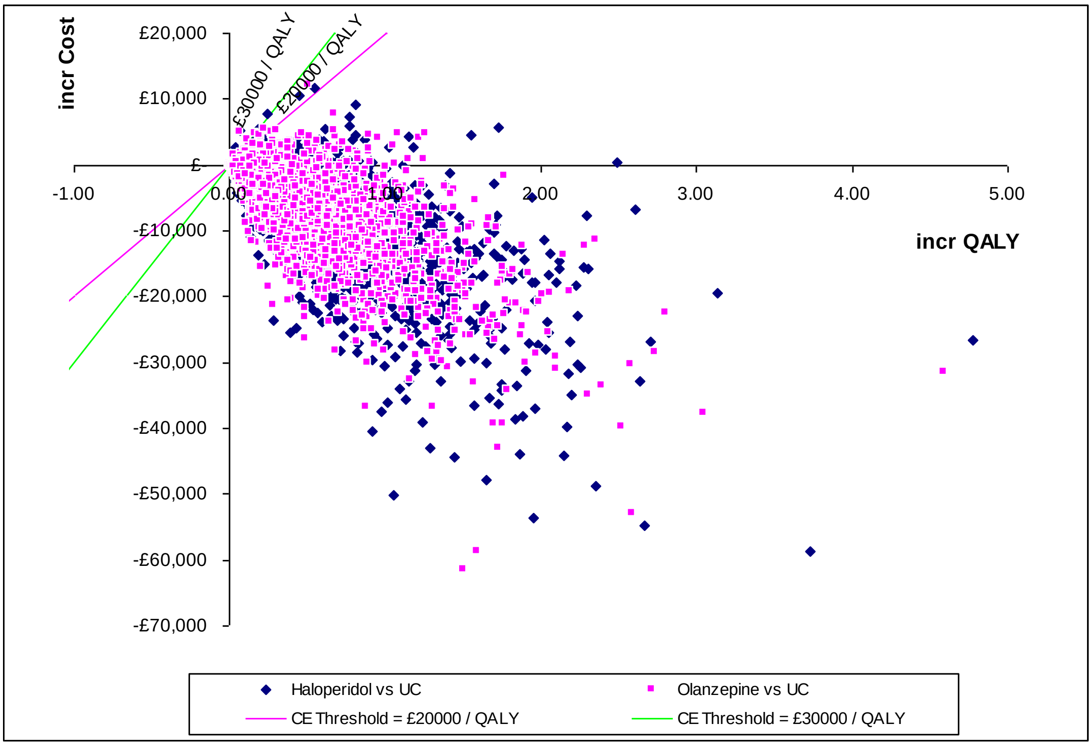
Table 16.10: costs, QALYs and cost-effectiveness of haloperidol and olanzapine treatment intervention compared to usual care

		Usual Care	Haloperidol	Olanzapine
Deterministic	Incremental net monetary benefit (INMB)*	N/A	£10,340	£9,390
Probabilistic	Mean cost	£31,120	£25,630	£26,090
	Mean QALYs	0.615	1.035	1.004
	Incremental Cost	N/A	-£5,490	-£5,030
	Incremental QALYs		0.420	0.390
	Incremental Cost per QALY gained		Haloperidol dominates	Olanzapine dominates
	Incremental net monetary benefit (INMB)*		£13,900	£12,820
	% of simulations where strategy was most cost-effective*	0%	54%	45%

* Using a cost-effectiveness threshold of £20,000 per QALY gained.

At a cost-effectiveness threshold of £20,000 per QALY, the use of haloperidol was the most cost-effective in 54.4% of the simulations that were run in the PSA (figure 16.5). The use of Olanzapine was most cost-effective in 45.4% of the simulations. Usual care was the most cost-effective strategy in only 0.3% of all simulations. Haloperidol increased cost and reduced QALYs in 0.00% of the simulations while olanzapine increased cost and reduced QALYs in 0.02% of the simulations. When compared to usual care and at a threshold of £20,000 per QALY, haloperidol was cost-effective 99.74% of all the 5000 simulations. For olanzapine, it was 99.72%. At a cost-effectiveness threshold of £30,000 per QALY, it was 99.92% and 99.90% for haloperidol and olanzapine respectively.

Figure 16.5: cost-effectiveness plane for haloperidol and olanzapine treatment interventions compared to usual care



When compared with olanzapine, haloperidol was associated with a mean cost reduction of -£460 and a mean incremental QALY of 0.031. The INMB was £1,080. However, there is wide uncertainty around the incremental cost-effectiveness of haloperidol compared to olanzapine as shown in figure 16.6. Haloperidol was more cost-effective in 54.5% of the 5000 simulations and olanzapine was more cost-effective in the rest (45.5%) of the simulations.

The two treatment intervention strategies in the model remained cost-effective in all the univariate DSA that we conducted (table 16.11). When compared with usual care, the use of the drugs resulted in higher INMB and became even more cost-effective when the time a person stays in long-term care before death was increased to 12 months. They became less cost effective when the impact of stroke side effect is included in the model. When compared to olanzapine, haloperidol was estimated to have the higher INMB for all the analyses conducted.

Figure 16.6: cost-effectiveness plane for haloperidol treatment interventions compared to olanzapine

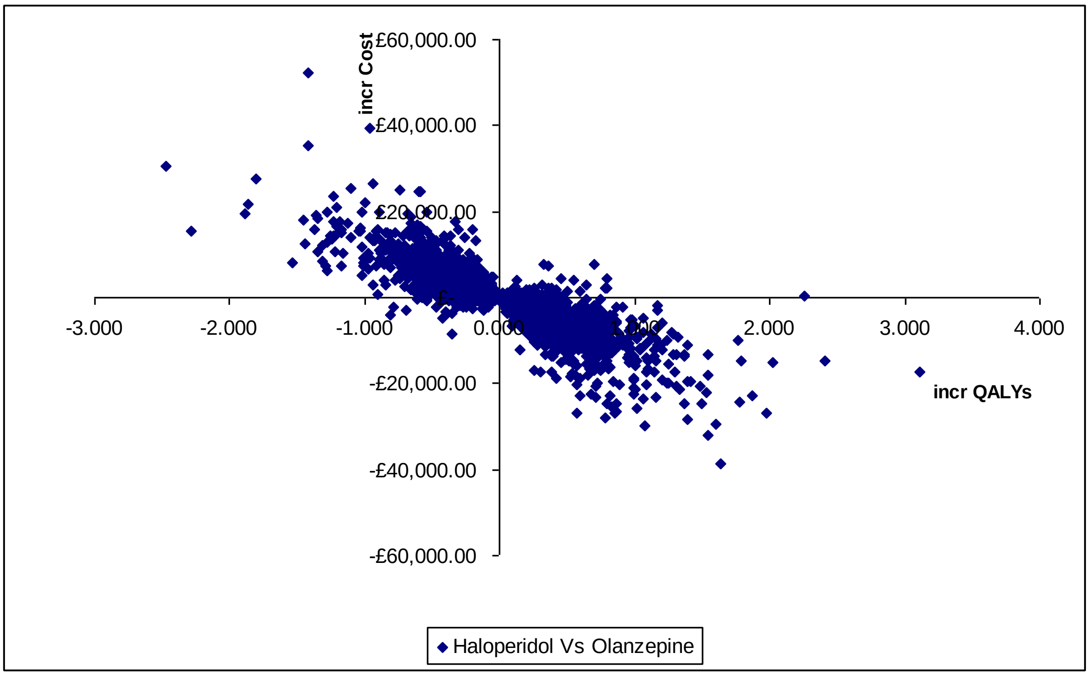


Table 16.11: other deterministic sensitivity analyses on the cost-effectiveness of haloperidol and olanzapine treatment interventions compared to usual care

	Incremental NMB (Haloperidol)	Incremental NMB (Olanzapine)
All model parameters excluding the side effect stroke (Base case)	£10,340	£9,390
All model parameters including the side effect stroke (RR for both atypical antipsychotic = 1.73)	£9,950	£9,000
Drug specific stroke relative risk (Hal=1.69, Ola=2.32)	£9,970	£8,680
Duration of stay in long-term care before death=12 months	£12,750	£11,580
Accounted for only 70% of cost of stay in long-term care	£9,100	£8,260
Increased cost of haloperidol due to increased dosage	£10,340	N/A
Increased cost of olanzapine due to increased dosage	N/A	£9,384

16.4 Summary of results of the cost-effectiveness analysis

We estimated the cost-effectiveness of prevention and treatment interventions using an original economic evaluation model. The use of multicomponent targeted interventions was found to be cost-effective in the prevention of delirium in the population groups considered in the model (elderly patients at risk of delirium who were admitted to the general medicine service and patients undergoing surgical repair of hip fracture). The use of haloperidol and olanzapine in the treatment of delirium was also cost-effective. On average, haloperidol was associated with a higher net monetary benefit but there is wide uncertainty around the incremental cost-effectiveness.

There are a number of limitations with the model findings and the GDG considered these when interpreting the results of the analyses. In the prevention model we have assumed that the adverse outcomes on the branches of the decision tree are mutually exclusive. It is possible that a patient with delirium who experiences dementia will also be admitted to a nursing home and the total cost and QALY gain for that patient might be different from the modelled estimate as the two outcomes are occurring in the same patient rather than in separate individuals We tried to test the impact of this assumption by considering that each of the six adverse outcomes was the only outcome to be associated with delirium therefore removing the risk of double counting. The results of the model were robust in that multicomponent interventions remain cost-effective.

In the prevention and treatment model, the baseline risk estimates we used for delirium in hospital, dementia, new admission to institution, complete recovery after delirium incidence and stroke were taken from studies in other countries. The baseline risk of complete recovery and efficacy of treatment interventions were taken from a study set in China (Hu 2006). The absolute risk used in the treatment model for nursing home admission, death or nursing home admission and death were taken from a US study (McAvay 2006). We could not identify suitable UK studies for these outcomes and the ones chosen were the best available in terms of study quality and applicability. We assumed that the relative risk of falls and pressure ulcer are the same. No other better studies could be identified for these outcomes. The GDG discussed the applicability of the studies that were used and considered them in the interpretation of the results.

The cost estimate used in the base case analysis for pressure ulcer in the prevention model was based on the assumption that it would be a grade 1 pressure ulcer that would heal normally. We made an alternative assumption that it would be a grade 4 ulcer. We assumed in the base case analysis for the prevention and treatment models that all the cost of long-term care will be paid by the NHS and PSS. We made an alternative assumption that only 70% of this cost will be paid by the public. The cost of dementia in the prevention model included the cost of stay in long-term care. It could be argued that the cost of long-term care has been accounted for as a different model outcome and that we have double counted cost. We made an alternative assumption and removed the proportion of cost of dementia attributable to long-term care. In all the alternative assumptions the model results suggest that the prevention and treatment interventions considered above remained cost-effective. In the treatment model we have assumed, in base case analysis, that patients who experience nursing home admission and death will spend only six months in long-term care before death. The cost-effectiveness estimate from this assumption was conservative as an increase in the duration to 12 months showed that the treatment interventions were even more cost-effective.

The point estimates used in the model were associated with some uncertainties which are normally captured in confidence intervals and ranges. We have tried to explore the effect of such uncertainties using probabilistic sensitivity analysis. The results of which did not change the findings that the use of multicomponent treatment interventions was found to be cost effective in elderly patients that had surgery for hip fracture repair, or elderly patients at intermediate or high risk of delirium who were admitted in the general medicine services. The use of haloperidol and olanzapine were also found to be cost-effective in the treatment of delirium.

17 References

The references listed below are for included studies and background papers

Cochrane Handbook for Systematic Reviews of Interventions 4.2.5 [Updated May 2005] (2007) in: Higgins J and Green S, (editors) *The Cochrane Library, Issue 3, 2005.* Chichester, UK: John Wiley & Sons, Ltd.,

Adamis D, Morrison C, Treloar A, Macdonald AJ, and Martin FC (2005) The Performance of the Clock Drawing Test in Elderly Medical Inpatients: Does It Have Utility in the Identification of Delirium?, *Journal of Geriatric Psychiatry and Neurology,* 18(3):129-33.

Agostini JV, Leo-Summers LS, and Inouye SK (2001) Cognitive and Other Adverse Effects of Diphenhydramine Use in Hospitalized Older Patients, *Archives of Internal Medicine,* 161(17):2091-7.

Aizawa K, Kanai T, Saikawa Y, Takabayashi T, Kawano Y, Miyazawa N, and Yamamoto T (2002) A Novel Approach to the Prevention of Postoperative Delirium in the Elderly After Gastrointestinal Surgery, *Surgery Today,* 32(4):310-4.

American Psychiatric Association (1980) *Diagnostic and Statistical Manual of Mental Disorders (DSM-III)*, 3rd Edition. Arlington, VA, US: American Psychiatric Publishing, Inc.

American Psychiatric Association (1987) *Diagnostic and Statistical Manual of Mental Disorders (DSM-III-R)*, Revised 3rd Edition. Arlington, VA, US: American Psychiatric Publishing, Inc.

American Psychiatric Association (1994) *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV)*, 4th Edition. Arlington, VA, US: American Psychiatric Publishing, Inc.

Andersson EM, Gustafson L, and Hallberg IR (2001) Acute Confusional State in Elderly Orthopaedic Patients: Factors of Importance for Detection in Nursing Care, *International Journal of Geriatric Psychiatry,* 16(1):7-17.

Andrew MK, Freter SH, and Rockwood K (2005) Incomplete Functional Recovery After Delirium in Elderly People: a Prospective Cohort Study, *BMC Geriatrics,* 5:5.

Andrew MK, Freter SH, and Rockwood K (2006) Prevalence and Outcomes of Delirium in Community and Non-Acute Care Settings in People Without Dementia: a Report From the Canadian Study of Health and Aging, *BMC Medicine,* 4:15.

Andrew MK, Bhat R, Clarke B, Freter SH, Rockwood MR, and Rockwood K (2009) Inter-Rater Reliability of the DRS-R-98 in Detecting Delirium in Frail Elderly Patients, *Age and Ageing,* 38(2):241-4.

Angles EM, Robinson TN, Biffl WL, Johnson J, Moss M, Tran ZV, and Moore EE (2008) Risk Factors for Delirium After Major Trauma, *American Journal of Surgery,* 196(6):864-9.

Balas MC, Deutschman CS, Sullivan-Marx EM, Strumpf NE, Alston RP, and Richmond TS (2007) Delirium in Older Patients in Surgical Intensive Care Units, *Journal of Nursing Scholarship,* 39(2):147-54.

Balas MC, Happ MB, Yang W, Chelluri L, and Richmond T (2009) Outcomes Associated With Delirium in Older Patients in Surgical ICUs, *Chest,* 135(1):18-25.

Bayindir O, Guden M, Akpinar B, Sanisoglu I, and Sagbas E (2001) Ondansetron Hydrochloride for the Treatment of Delirium After Coronary Artery Surgery, *Journal of Thoracic and Cardiovascular Surgery,* 121(1):176-7.

Beaussier M, Weickmans H, Parc Y, Delpierre E, Camus Y, Funck-Brentano C, Schiffer E, Delva E, and Lienhart A (2006) Postoperative Analgesia and Recovery Course After Major Colorectal Surgery in Elderly Patients: a Randomized Comparison Between Intrathecal Morphine and Intravenous PCA Morphine, *Regional Anesthesia and Pain Medicine,* 31(6):531-8.

Bennett G, Dealey C, and Posnett J (2004) The Cost of Pressure Ulcers in the UK, *Age and Ageing,* 33(3):230-5.

Benoit AG, Campbell BI, Tanner JR, Staley JD, Wallbridge HR, Biehl DR, Bradley BD, Louridas G, Guzman RP, and Fromm RA (2005) Risk Factors and Prevalence of Perioperative Cognitive Dysfunction in Abdominal Aneurysm Patients, *Journal of Vascular Surgery,* 42(5):884-90.

Bickel H, Gradinger R, Kochs E, and Forstl H (2008) High Risk of Cognitive and Functional Decline After Postoperative Delirium: A Three-Year Prospective Study, *Dementia and Geriatric Cognitive Disorders,* 26(1):26-31.

Bogardus ST, Desai MM, Williams CS, Leo-Summers L, Acampora D, and Inouye SK (2003) The Effects of a Targeted Multicomponent Delirium Intervention on Postdischarge Outcomes for Hospitalized Older Adults, *American Journal of Medicine,* 114(5):383-90.

Bourdel-Marchasson I, Vincent S, Germain C, Salles N, Jenn J, Rasoamanarivo E, Emeriau JP, Rainfray M, and Richard-Harston S (2004) Delirium Symptoms and Low Dietary Intake in Older Inpatients Are Independent Predictors of Institutionalization: a 1-Year Prospective Population-Based Study, *Journals of Gerontology Series A-Biological Sciences and Medical Sciences,* 59(4):350-4.

Böhner H, Hummel TC, Habel U, Miller C, Reinbott S, Yang Q, Gabriel A, Friedrichs R, Müller EE, Ohmann C, Sandmann W, and Schneider F (2003) Predicting Delirium After Vascular Surgery: a Model Based on Pre- and Intraoperative Data, *Annals of Surgery,* 238(1):149-56.

Bracco D, Noiseux N, Dubois M-J, Prieto I, Basile F, Olivier J-F, and Hemmerling T (2007) Epidural Anesthesia Improves Outcome and Resource Use in Cardiac Surgery: A Single-Center Study of a 1293-Patient Cohort, *Heart Surgery Forum,* 10(6):301-10.

Brauer C, Morrison RS, Silberzweig SB, and Siu AL (2000) The Cause of Delirium in Patients With Hip Fracture, *Archives of Internal Medicine,* 160(12):1856-60.

Breitbart W, Marotta R, Platt MM, Weisman H, Derevenco M, Grau C, Corbera K, Raymond S, Lund S, and Jacobson P (1996) A Double-Blind Trial of Haloperidol, Chlorpromazine, and Lorazepam in the Treatment of Delirium in Hospitalized AIDS Patients, *American Journal of Psychiatry,* 153(2):231-7.

Breitbart W, Tremblay A, and Gibson C (2002) An Open Trial of Olanzapine for the Treatment of Delirium in Hospitalized Cancer Patients, *Psychosomatics,* 43(3):175-82.

Bryant RA, Harvey AG, Dang ST, Sackville T, and Basten C (1998) Treatment of Acute Stress Disorder: a Comparison of Cognitive-Behavioral Therapy and Supportive Counseling, *Journal of Consulting and Clinical Psychology,* 66(5):862-6.

Bucerius J, Gummert JF, Borger MA, Walther T, Doll N, Falk V, Schmitt DV, and Mohr FW (2004) Predictors of Delirium After Cardiac Surgery Delirium: Effect of Beating-Heart (Off-Pump) Surgery, *Journal of Thoracic and Cardiovascular Surgery,* 127(1):57-64.

Caeiro L, Ferro JM, Claro MI, Coelho J, Albuquerque R, and Figueira ML (2004) Delirium in Acute Stroke: a Preliminary Study of the Role of Anticholinergic Medications, *European Journal of Neurology,* 11(10):699-704.

Caeiro L, Ferro JM, Albuquerque R, and Figueira ML (2004) Delirium in the First Days of Acute Stroke, *Journal of Neurology,* 251(2):171-8.

Centorrino F, Albert MJ, Drago-Ferrante G, Koukopoulos AE, Berry JM, and Baldessarini RJ (2003) Delirium During Clozapine Treatment: Incidence and Associated Risk Factors, *Pharmacopsychiatry,* 36(4):156-60.

Christe C, Janssens JP, Armenian B, Herrmann F, and Vogt N (2000) Midazolam Sedation for Upper Gastrointestinal Endoscopy in Older Persons: a Randomized, Double-Blind, Placebo-Controlled Study, *Journal of the American Geriatrics Society,* 48(11):1398-403.

Clark M and Watts S (1994) The Incidence of Pressure Sores Within a National Health Service Trust Hospital During 1991, *Journal of Advanced Nursing,* 20(1):33-6.

Cole MG, Primeau FJ, Bailey RF, Bonnycastle MJ, Masciarelli F, Engelsmann F, Pepin MJ, and Ducic D (1994) Systematic Intervention for Elderly Inpatients With Delirium: a Randomized Trial, *CMAJ: Canadian Medical Association Journal,* 151(7):965-70.

Cole MG, McCusker J, Bellavance F, Primeau FJ, Bailey RF, Bonnycastle MJ, and Laplante J (2002) Systematic Detection and Multidisciplinary Care of Delirium in Older Medical Inpatients: a Randomized Trial, *CMAJ: Canadian Medical Association Journal,* 167(7):753-9.

Cole MG, Dendukuri N, McCusker J, and Han L (2003) An Empirical Study of Different Diagnostic Criteria for Delirium Among Elderly Medical Inpatients, *Journal of Neuropsychiatry and Clinical Neurosciences,* 15(2):200-7.

Contin AM (2005) Postoperative Delirium After Elective Orthopedic Surgery, *International Journal of Geriatric Psychiatry,* 20(6):595-7.

Dolan MM (2000) Delirium on Hospital Admission in Aged Hip Fracture Patients: Prediction of Mortality and 2-Year Functional Outcomes, *Journals of Gerontology Series A-Biological Sciences and Medical Sciences,* 55(9):M527-M534.

Douglas IJ and Smeeth L (2008) Exposure to Antipsychotics and Risk of Stroke: Self Controlled Case Series Study, *BMJ,* 337:a1227.

Drame M, Jovenin N, Novella J-L, Lang P-O, Somme D, Laniece I, Voisin T, Blanc P, Couturier P, Gauvain J-B, Blanchard F, and Jolly D (2008) Predicting Early Mortality Among Elderly Patients Hospitalised in Medical Wards Via Emergency Department: The Safes Cohort Study, *Journal of Nutrition, Health and Aging,* Serdi Publishing Company. 12(8):599-604.

Dubois M-J (2001) Delirium in an Intensive Care Unit: A Study of Risk Factors, *Intensive Care Medicine,* 27(8):1297-304.

Duppils GS and Wikblad K (2007) Patients' Experiences of Being Delirious, *Journal of Clinical Nursing,* 16(5):810-8.

Edelstein DM, Aharonoff GB, Karp A, Capla EL, Zuckerman JD, and Koval KJ (2004) Effect of Postoperative Delirium on Outcome After Hip Fracture, *Clinical Orthopaedics and Related Research,*(422):195-200.

Edlund A, Lundstrom M, Lundstrom G, Hedqvist B, and Gustafson Y (1999) Clinical Profile of Delirium in Patients Treated for Femoral Neck Fractures, *Dementia and Geriatric Cognitive Disorders,* 10(5):325-9.

Edlund A, Lundstrom M, Brannstrom B, Bucht G, and Gustafson Y (2001) Delirium Before and After Operation for Femoral Neck Fracture, *Journal of the American Geriatrics Society,* 49(10):1335-40.

Edlund A, Lundstrom M, Karlsson S, Brannstrom B, Bucht G, and Gustafson Y (2006) Delirium in Older Patients Admitted to General Internal Medicine, *Journal of Geriatric Psychiatry and Neurology,* 19(2):83-90.

Ehlers A, Clark DM, Hackmann A, McManus F, Fennell M, Herbert C, and Mayou R (2003) A Randomized Controlled Trial of Cognitive Therapy, a Self-Help Booklet, and Repeated Assessments As Early Interventions for Posttraumatic Stress Disorder, *Archives of General Psychiatry,* 60(10):1024-32.

Ekman M, Berg J, Wimo A, Jonsson L, and McBurney C (2007) Health Utilities in Mild Cognitive Impairment and Dementia: a Population Study in Sweden, *International Journal of Geriatric Psychiatry,* 22(7):649-55.

Elie M, Rousseau F, Cole M, Primeau F, McCusker J, and Bellavance F (2000) Prevalence and Detection of Delirium in Elderly Emergency Department Patients, *CMAJ: Canadian Medical Association Journal,* 163(8):977-81.

Ely EW, Margolin R, Francis J, May L, Truman B, Dittus R, Speroff T, Gautam S, Bernard GR, and Inouye SK (2001b) Evaluation of Delirium in Critically Ill Patients: Validation of the Confusion Assessment Method for the Intensive Care Unit (CAM-ICU), *Critical Care Medicine,* 29(7):1370-9.

Ely EW, Inouye SK, Bernard GR, Gordon S, Francis J, May L, Truman B, Speroff T, Gautam S, Margolin R, Hart RP, and Dittus R (2001) Delirium in Mechanically Ventilated Patients: Validity and Reliability of the Confusion Assessment Method for the Intensive Care Unit (CAM-ICU), *JAMA: Journal of the American Medical Association,* 286(21):2703-10.

Ely EW, Shintani A, Truman B, Speroff T, Gordon SM, Harrell FE, Jr., Inouye SK, Bernard GR, and Dittus RS (2004) Delirium As a Predictor of Mortality in Mechanically Ventilated Patients in the Intensive Care Unit, *JAMA: Journal of the American Medical Association,* 291(14):1753-62.

Ely EW, Girard TD, Shintani AK, Jackson JC, Gordon SM, Thomason JW, Pun BT, Canonico AE, Light RW, Pandharipande P, and Laskowitz DT (2007) Apolipoprotein E4 Polymorphism As a Genetic Predisposition to Delirium in Critically Ill Patients, *Critical Care Medicine,* 35(1):112-7.

Fabbri RM, Moreira MA, Garrido R, and Almeida OP (2001) Validity and Reliability of the Portuguese Version of the Confusion Assessment Method (CAM) for the Detection of Delirium in the Elderly, *Arquivos De Neuro-Psiquiatria,* 59(2-A):175-9.

Faezah SK, Zhang D, and Yin LF (2008) The Prevalence and Risk Factors of Delirium Amongst the Elderly in Acute Hospital, *Singapore Nursing Journal,* 35(1):11-4.

Foy A, O'Connell D, Henry D, Kelly J, Cocking S, and Halliday J (1995) Benzodiazepine Use As a Cause of Cognitive Impairment in Elderly Hospital Inpatients, *Journals of Gerontology Series A-Biological Sciences and Medical Sciences,* 50(2):M99-106.

Francis J, Martin D, and Kapoor WN (1990) A Prospective Study of Delirium in Hospitalized Elderly, *JAMA: Journal of the American Medical Association,* 263(8):1097-101.

Francis J and Kapoor WN (1992) Prognosis After Hospital Discharge of Older Medical Patients With Delirium, *Journal of the American Geriatrics Society,* 40(6):601-6.

Franco K (2001) The Cost of Delirium in the Surgical Patient, *Psychosomatics,* 42(1):68-73.

Furlaneto ME and Garcez-Leme LE (2006) Delirium in Elderly Individuals With Hip Fracture: Causes, Incidence, Prevalence, and Risk Factors, *Clinics,* 61(1):35-40.

Gagnon P, Charbonneau C, Allard P, Soulard C, Dumont S, and Fillion L (2002) Delirium in Advanced Cancer: a Psychoeducational Intervention for Family Caregivers, *Journal of Palliative Care,* 18(4):253-61.

Galanakis P, Bickel H, Gradinger R, Von Gumppenberg S, and Forstl H (2001) Acute Confusional State in the Elderly Following Hip Surgery: Incidence, Risk Factors and Complications, *International Journal of Geriatric Psychiatry,* 16(4):349-55.

Gamberini M, Bolliger D, Lurati Buse GA, Burkhart CS, Grapow M, Gagneux A, Filipovic M, Seeberger MD, Pargger H, Siegemund M, Carrel T, Seiler WO, Berres M, Strebel SP, Monsch AU, and Steiner LA (2009) Rivastigmine for the Prevention of Postoperative Delirium in Elderly Patients Undergoing Elective Cardiac Surgery--a Randomized Controlled Trial, *Critical Care Medicine,* 37(5):1762-8.

Gill SS, Rochon PA, Herrmann N, Lee PE, Sykora K, Gunraj N, Normand SL, Gurwitz JH, Marras C, Wodchis WP, and Mamdani M (2005) Atypical Antipsychotic Drugs and Risk of Ischaemic Stroke: Population Based Retrospective Cohort Study, *BMJ,* 330(7489):445.

Givens JL, Sanft TB, and Marcantonio ER (2008) Functional Recovery After Hip Fracture: the Combined Effects of Depressive Symptoms, Cognitive Impairment, and Delirium, *Journal of the American Geriatrics Society,* 56(6):1075-9.

Goldenberg G (2006) Predicting Post-Operative Delirium in Elderly Patients Undergoing Surgery for Hip Fracture, *Psychogeriatrics,* 6(2):43-8.

Gonzalez M, de Pablo J, Fuente E, Valdes M, Peri JM, Nomdedeu M, and Matrai S (2004) Instrument for Detection of Delirium in General Hospitals: Adaptation of the Confusion Assessment Method, *Psychosomatics,* 45(5):426-31.

Granberg A, Bergbom E, I, and Lundberg D (1998) Patients' Experience of Being Critically Ill or Severely Injured and Cared for in an Intensive Care Unit in Relation to the ICU Syndrome. Part I, *Intensive and Critical Care Nursing,* 14(6):294-307.

Greene NH, Attix DK, Weldon BC, Smith PJ, McDonagh DL, and Monk TG (2009) Measures of Executive Function and Depression Identify Patients at Risk for Postoperative Delirium, *Anesthesiology,* 110(4):788-95.

Grieve R, Hutton J, Bhalla A, Rastenyte D, Ryglewicz D, Sarti C, Lamassa M, Giroud M, Dundas R, and Wolfe CD (2001) A Comparison of the Costs and Survival of Hospital-Admitted Stroke Patients Across Europe, *Stroke,* 32(7):1684-91.

Gustafson Y, Brannstrom B, Berggren D, Ragnarsson JI, Sigaard J, Bucht G, Reiz S, Norberg A, and Winblad B (1991) A Geriatric-Anesthesiologic Program to Reduce Acute Confusional States in Elderly Patients Treated for Femoral Neck Fractures, *Journal of the American Geriatrics Society,* 39(7):655-62.

Hamann J, Bickel H, Schwaibold H, Hartung R, and Forstl H (2005) Postoperative Acute Confusional State in Typical Urologic Population: Incidence, Risk Factors, and Strategies for Prevention, *Urology,* 65(3):449-53.

Han JH, Morandi A, Ely EW, Callison C, Zhou C, Storrow AB, Dittus RS, Habermann R, and Schnelle J (2009) Delirium in the Nursing Home Patients Seen in the Emergency Department: Brief Reports, *Journal of the American Geriatrics Society,* 57(5):889-94.

Han L, McCusker J, Cole M, Abrahamowicz M, Primeau F, and Elie M (2001) Use of Medications With Anticholinergic Effect Predicts Clinical Severity of Delirium Symptoms in Older Medical Inpatients, *Archives of Internal Medicine,* 161(8):1099-105.

Harari D, Hopper A, Dhesi J, Babic-Illman G, Lockwood L, and Martin F (2007) Proactive Care of Older People Undergoing Surgery ('POPS'): Designing, Embedding, Evaluating and Funding a Comprehensive Geriatric Assessment Service for Older Elective Surgical Patients, *Age and Ageing,* 36(2):190-6.

Harding R, Martin C, and Holmes J (2008) Dazed and Confused: Making Sense of Delirium After Hip Fracture, *International Journal of Geriatric Psychiatry,* 23(9):984-6.

Hendriks MR, Evers SM, Bleijlevens MH, van Haastregt JC, Crebolder HF, and van Eijk JT (2008) Cost-Effectiveness of a Multidisciplinary Fall Prevention Program in Community-Dwelling Elderly People: a Randomized Controlled Trial, *International Journal of Technology Assessment in Health Care,* 24(2):193-202.

Henon H, Lebert F, Durieu I, Godefroy O, Lucas C, Pasquier F, and Leys D (1999) Confusional State in Stroke: Relation to Preexisting Dementia, Patient Characteristics, and Outcome, *Stroke,* 30(4):773-9.

Herrick IA, Ganapathy S, Komar W, Kirkby J, Moote CA, Dobkowski W, and Eliasziw M (1996) Postoperative Cognitive Impairment in the Elderly. Choice of Patient-Controlled Analgesia Opioid, *Anaesthesia,* 51(4):356-60.

Herrmann N, Mamdani M, and Lanctot KL (2004) Atypical Antipsychotics and Risk of Cerebrovascular Accidents, *American Journal of Psychiatry,* 161(6):1113-5.

Hestermann U, Backenstrass M, Gekle I, Hack M, Mundt C, Oster P, and Thomas C (2009) Validation of a German Version of the Confusion Assessment Method for Delirium Detection in a Sample of Acute Geriatric Patients With a High Prevalence of Dementia, *Psychopathology,* 42(4):270-6.

Hofste WJ, Linssen CA, Boezeman EH, Hengeveld JS, Leusink JA, and de Boer A (1997) Delirium and Cognitive Disorders After Cardiac Operations: Relationship to Pre- and Intraoperative Quantitative Electroencephalogram, *International Journal of Clinical Monitoring and Computing,* 14(1):29-36.

Holden J, Jayathissa S, and Young G (2008) Delirium Among Elderly General Medical Patients in a New Zealand Hospital, *Internal Medicine Journal,* 38(8):629-34.

Holmes J and House A (2000) Psychiatric Illness Predicts Poor Outcome After Surgery for Hip Fracture: A Prospective Cohort Study, *Psychological Medicine,* 30(4):921-9.

Holroyd S and Rabins PV (1994) A Retrospective Chart Review of Lithium Side Effects in a Geriatric Outpatient Population, *American Journal of Geriatric Psychiatry,* 2(4):346-51.

Hu H (2006) Olanzapine and Haloperidol for Senile Delirium: A Randomized Controlled Observation, *Chinese Journal of Clinical Rehabilitation,* 10(42):188-90.

Iglesias CP, Manca A, and Torgerson DJ (2008) The Health-Related Quality of Life and Cost Implications of Falls in Elderly Women, *Osteoporos International,*

Inouye SK, Viscoli CM, Horwitz RI, Hurst LD, and Tinetti ME (1993) A Predictive Model for Delirium in Hospitalized Elderly Medical Patients Based on Admission Characteristics, *Annals of Internal Medicine,* 119(6):474-81.

Inouye SK, Rushing JT, Foreman MD, Palmer RM, and Pompei P (1998) Does Delirium Contribute to Poor Hospital Outcomes? A Three-Site Epidemiologic Study, *Journal of General Internal Medicine,* 13(4):234-42.

Inouye SK, Bogardus ST, Jr., Charpentier PA, Leo-Summers L, Acampora D, Holford TR, and Cooney LM, Jr. (1999) A Multicomponent Intervention to Prevent Delirium in Hospitalized Older Patients, *New England Journal of Medicine,* 340(9):669-76.

Inouye SK, Bogardus ST, Jr., Baker DI, Leo-Summers L, and Cooney LM, Jr. (2000) The Hospital Elder Life Program: a Model of Care to Prevent Cognitive and Functional Decline in Older Hospitalized Patients, *Journal of the American Geriatrics Society,* 48(12):1697-706.

Inouye SK, Zhang Y, Jones RN, Kiely DK, Yang F, and Marcantonio ER (2007) Risk Factors for Delirium at Discharge: Development and Validation of a Predictive Model, *Archives of Internal Medicine,* 167(13):1406-13.

Inouye SK (1998) Delirium in Hospitalized Older Patients: Recognition and Risk Factors, *Journal of Geriatric Psychiatry and Neurology,* 11(3):118-25.

Inouye SK (1999) Predisposing and Precipitating Factors for Delirium in Hospitalized Older Patients, *Dementia and Geriatric Cognitive Disorders,* 10(5):393-400.

Joint Formulary Committee (2009) *British National Formulary*, 57th Edition. London: British Medical Association and Royal Pharmaceutical Society of Great Britain. Available from: http://www.bnf.org.uk

Jones RN, Yang FM, Zhang Y, Kiely DK, Marcantonio ER, and Inouye SK (2006) Does Educational Attainment Contribute to Risk for Delirium? A Potential Role for Cognitive Reserve, *Journals of Gerontology Series A-Biological Sciences and Medical Sciences,* 61(12):1307-11.

Kagansky N, Rimon E, Naor S, Dvornikov E, Cojocaru L, and Levy S (2004) Low Incidence of Delirium in Very Old Patients After Surgery for Hip Fractures, *American Journal of Geriatric Psychiatry,* 12(3):306-14.

Kakuma R, Galbaud du FG, Arsenault L, Perrault A, Platt RW, Monette J, Moride Y, and Wolfson C (2003) Delirium in Older Emergency Department Patients Discharged Home: Effect on Survival, *Journal of the American Geriatrics Society,* 51(4):443-50.

Kalisvaart KJ, de Jonghe JFM, Bogaards MJ, Vreeswijk R, Egberts TCG, Burger BJ, Eikelenboom P, and van Gool WA (2005) Haloperidol Prophylaxis for Elderly Hip-Surgery Patients at Risk for Delirium: A Randomized Placebo-Controlled Study, *Journal of the American Geriatrics Society,* 53(10):1658-66.

Kaneko T (1999) Prophylactic Consecutive Administration of Haloperidol Can Reduce the Occurrence of Postoperative Delirium in Gastrointestinal Surgery, *Yonago Acta Medica,* 42(3):179-84.

Kanis JA, Brazier JE, Stevenson M, Calvert NW, and Lloyd JM (2002) Treatment of Established Osteoporosis: a Systematic Review and Cost-Utility Analysis, *Health Technology Assessment,* 6(29):1-146.

Kawaguchi Y, Kanamori M, Ishihara H, Abe Y, Nobukiyo M, Sigeta T, Hori T, and Kimura T (2006) Postoperative Delirium in Spine Surgery, *Spine Journal,* 6(2):164-9.

Kazmierski J, Kowman M, Banach M, Pawelczyk T, Okonski P, Iwaszkiewicz A, Zaslonka J, Sobow T, and Kloszewska I (2006) Preoperative Predictors of Delirium After Cardiac Surgery: a Preliminary Study, *General Hospital Psychiatry,* 28(6):536-8.

Kim KS, Pae CU, Chae JH, Bahk WM, and Jun T (2001) An Open Pilot Trial of Olanzapine for Delirium in the Korean Population, *Psychiatry and Clinical Neurosciences,* 55(5):515-9.

Kim KY, McCartney JR, Kaye W, Boland RJ, and Niaura R (1996) The Effect of Cimetidine and Ranitidine on Cognitive Function in Postoperative Cardiac Surgical Patients, *International Journal of Psychiatry in Medicine,* 26(3):295-307.

Kind P, Dolan P, Gudex C, and Williams A (1998) Variations in Population Health Status: Results From a United Kingdom National Questionnaire Survey, *BMJ,* 316(7133):736-41.

Koebrugge B, Koek HL, Van Wensen RJA, Dautzenberg PLJ, and Bosscha K (2009) Delirium After Abdominal Surgery at a Surgical Ward With a High Standard of Delirium Care: Incidence, Risk Factors and Outcomes, *Digestive Surgery,* 26(1):63-8.

Korevaar JC, van Munster BC, and de Rooij SE (2005) Risk Factors for Delirium in Acutely Admitted Elderly Patients: a Prospective Cohort Study, *BMC Geriatrics,* 5:6.

Koster S, Oosterveld FG, Hensens AG, Wijma A, and van der PJ (2008) Delirium After Cardiac Surgery and Predictive Validity of a Risk Checklist, *Annals of Thoracic Surgery,* 86(6):1883-7.

Laitinen H (1996) Patients' Experience of Confusion in the Intensive Care Unit Following Cardiac Surgery, *Intensive and Critical Care Nursing,* 12(2):79-83.

Landefeld CS, Palmer RM, Kresevic DM, Fortinsky RH, and Kowal J (1995) A Randomized Trial of Care in a Hospital Medical Unit Especially Designed to Improve the Functional Outcomes of Acutely Ill Older Patients, *New England Journal of Medicine,* 332(20):1338-44.

Laurila JV, Pitkala KH, Strandberg TE, and Tilvis RS (2002) Confusion Assessment Method in the Diagnostics of Delirium Among Aged Hospital Patients: Would It Serve Better in Screening Than As a Diagnostic Instrument?, *International Journal of Geriatric Psychiatry,* 17(12):1112-9.

Laurila JV, Pitkala KH, Strandberg TE, and Tilvis RS (2003) The Impact of Different Diagnostic Criteria on Prevalence Rates for Delirium, *Dementia and Geriatric Cognitive Disorders,* 16(3):156-62.

Laurila JV, Pitkala KH, Strandberg TE, and Tilvis RS (2004) Impact of Different Diagnostic Criteria on Prognosis of Delirium: a Prospective Study, *Dementia and Geriatric Cognitive Disorders,* 18(3-4):240-4.

Laurila JV, Pitkala KH, Strandberg TE, and Tilvis RS (2004) Delirium Among Patients With and Without Dementia: Does the Diagnosis According to the DSM-IV Differ From the Previous Classifications?, *International Journal of Geriatric Psychiatry,* 19(3):271-7.

Lee K-U (2005) Amisulpride Versus Quetiapine for the Treatment of Delirium: A Randomized, Open Prospective Study, *International Clinical Psychopharmacology,* 20(6):311-4.

Leslie DL, Zhang Y, Holford TR, Bogardus ST, Leo-Summers LS, and Inouye SK (2005) Premature Death Associated With Delirium at 1-Year Follow-Up, *Archives of Internal Medicine,* 165(14):1657-62.

Leung JM, Sands LP, Vaurio LE, and Wang Y (2006) Nitrous Oxide Does Not Change the Incidence of Postoperative Delirium or Cognitive Decline in Elderly Surgical Patients, *British Journal of Anaesthesia,* 96(6):754-60.

Leung JM (2007) Apolipoprotein E E4 Allele Increases the Risk of Early Postoperative Delirium in Older Patients Undergoing Noncardiac Surgery, *Anesthesiology,* 107(3):406-11.

Levkoff SE, Safran C, Cleary PD, Gallop J, and Phillips RS (1988) Identification of Factors Associated With the Diagnosis of Delirium in Elderly Hospitalized Patients, *Journal of the American Geriatrics Society,* 36(12):1099-104.

Levkoff SE, Evans DA, Liptzin B, Cleary PD, Lipsitz LA, Wetle TT, Reilly CH, Pilgrim DM, Schor J, and Rowe J (1992) Delirium. The Occurrence and Persistence of Symptoms Among Elderly Hospitalized Patients, *Archives of Internal Medicine,* 152(2):334-40.

Lewis LM, Miller DK, Morley JE, Nork MJ, and Lasater LC (1995) Unrecognized Delirium in ED Geriatric Patients, *American Journal of Emergency Medicine,* 13(2):142-5.

Lin S-M (2004) The Impact of Delirium on the Survival of Mechanically Ventilated Patients, *Critical Care Medicine,* 32(11):2254-9.

Lin S-M, Huang C-D, Liu C-Y, Lin H-C, Wang C-H, Huang P-Y, Fang Y-F, Shieh M-H, and Kuo H-P (2008) Risk Factors for the Development of Early-Onset Delirium and the Subsequent Clinical Outcome in Mechanically Ventilated Patients, *Journal of Critical Care,* 23(3):372-9.

Lindgren P, Glader EL, and Jonsson B (2008) Utility Loss and Indirect Costs After Stroke in Sweden, *European Journal of Cardiovascular Prevention and Rehabilitation,* 15(2):230-3.

Liptzin B, Laki A, Garb JL, Fingeroth R, and Krushell R (2005) Donepezil in the Prevention and Treatment of Post-Surgical Delirium, *American Journal of Geriatric Psychiatry,* 13(12):1100-6.

Lundstrom M, Edlund A, Karlsson S, Brannstrom B, Bucht G, and Gustafson Y (2005) A Multifactorial Intervention Program Reduces the Duration of Delirium, Length of Hospitalization, and Mortality in Delirious Patients, *Journal of the American Geriatrics Society,* 53(4):622-8.

Magarey JM and McCutcheon HH (2005) 'Fishing With the Dead'--Recall of Memories From the ICU, *Intensive and Critical Care Nursing,* 21(6):344-54.

Marcantonio E, Ta T, Duthie E, and Resnick NM (2002) Delirium Severity and Psychomotor Types: Their Relationship With Outcomes After Hip Fracture Repair, *Journal of the American Geriatrics Society,* 50(5):850-7.

Marcantonio ER, Juarez G, Goldman L, Mangione CM, Ludwig LE, Lind L, Katz N, Cook EF, Orav EJ, and Lee TH (1994) The Relationship of Postoperative Delirium With Psychoactive Medications, *JAMA: Journal of the American Medical Association,* 272(19):1518-22.

Marcantonio ER, Goldman L, Mangione CM, Ludwig LE, Muraca B, Haslauer CM, Donaldson MC, Whittemore AD, Sugarbaker DJ, and Poss R (1994) A Clinical Prediction Rule for Delirium After Elective Noncardiac Surgery, *JAMA: Journal of the American Medical Association,* 271(2):134-9.

Marcantonio ER, Flacker JM, Michaels M, and Resnick NM (2000) Delirium Is Independently Associated With Poor Functional Recovery After Hip Fracture, *Journal of the American Geriatrics Society,* 48(6):618-24.

Marcantonio ER, Flacker JM, Wright RJ, and Resnick NM (2001) Reducing Delirium After Hip Fracture: a Randomized Trial, *Journal of the American Geriatrics Society,* 49(5):516-22.

Marcantonio ER, Kiely DK, Simon SE, John OE, Jones RN, Murphy KM, and Bergmann MA (2005) Outcomes of Older People Admitted to Postacute Facilities With Delirium, *Journal of the American Geriatrics Society,* 53(6):963-9.

Margiotta A, Bianchetti A, Ranieri P, and Trabucchi M (2006) Clinical Characteristics and Risk Factors of Delirium in Demented and Not Demented Elderly Medical Inpatients, *Journal of Nutrition, Health and Aging,* 10(6):535-9.

Martin NJ, Stones MJ, Young JE, and Bedard M (2000) Development of Delirium: a Prospective Cohort Study in a Community Hospital, *International Psychogeriatrics,* 12(1):117-27.

Mason J, Nicolson D, and Wilson D. (2002) Systematic Review Methods for National Guidelines (Unpublished Discussion Paper).

McAlpine JN, Hodgson EJ, Abramowitz S, Richman SM, Su Y, Kelly MG, Luther M, Baker L, Zelterman D, Rutherford TJ, and Schwartz PE (2008) The Incidence and Risk Factors Associated With Postoperative Delirium in Geriatric Patients Undergoing Surgery for Suspected Gynecologic Malignancies, *Gynecologic Oncology,* 109(2):296-302.

McAvay GJ, Van Ness PH, Bogardus ST, Jr., Zhang Y, Leslie DL, Leo-Summers LS, and Inouye SK (2006) Older Adults Discharged From the Hospital With Delirium: 1-Year Outcomes, *Journal of the American Geriatrics Society,* 54(8):1245-50.

McCaffrey R and Locsin R (2004) The Effect of Music Listening on Acute Confusion and Delirium in Elders Undergoing Elective Hip and Knee Surgery, *Journal of Clinical Nursing,* 13(6b):91-6.

McCaffrey R and Locsin R (2006) The Effect of Music on Pain and Acute Confusion in Older Adults Undergoing Hip and Knee Surgery, *Holistic Nursing Practice,* 20(5):218-26.

McCurren C and Cronin SN (2003) Delirium: Elders Tell Their Stories and Guide Nursing Practice, *MEDSURG Nursing,* 12(5):318-23.

McCusker J, Cole M, Abrahamowicz M, Han L, Podoba JE, and Ramman-Haddad L (2001) Environmental Risk Factors for Delirium in Hospitalized Older People, *Journal of the American Geriatrics Society,* 49(10):1327-34.

McCusker J, Cole MG, Dendukuri N, and Belzile E (2003) Does Delirium Increase Hospital Stay?, *Journal of the American Geriatrics Society,* 51(11):1539-46.

McNamee P, Bond J, Buck D, and Resource Implications Study of the Medical Research Council Cognitive Function and Ageing Study. (2001) Costs of Dementia in England and Wales in the 21st Century, *British Journal of Psychiatry,* 179:261-6.

McNicoll L, Pisani MA, Zhang Y, Ely EW, Siegel MD, and Inouye SK (2003) Delirium in the Intensive Care Unit: Occurrence and Clinical Course in Older Patients, *Journal of the American Geriatrics Society,* 51(5):591-8.

Mentes JC and Culp K (2003) Reducing Hydration-Linked Events in Nursing Home Residents, *Clinical Nursing Research,* 12(3):210-25.

Milbrandt EB, Deppen S, Harrison PL, Shintani AK, Speroff T, Stiles RA, Truman B, Bernard GR, Dittus RS, and Ely EW (2004) Costs Associated With Delirium in Mechanically Ventilated Patients, *Critical Care Medicine,* 32(4):955-62.

Milisen K, Foreman MD, Abraham IL, De Geest S, Godderis J, Vandermeulen E, Fischler B, Delooz HH, Spiessens B, and Broos PLO (2001) A Nurse-Led Interdisciplinary Intervention Program for Delirium in Elderly Hip-Fracture Patients, *Journal of the American Geriatrics Society,* 49(5):523-32.

Miyaji S, Yamamoto K, Hoshino S, Yamamoto H, Sakai Y, and Miyaoka H (2007) Comparison of the Risk of Adverse Events Between Risperidone and Haloperidol in Delirium Patients, *Psychiatry and Clinical Neurosciences,* 61(3):275-82.

Monette J, G, Fung SH, Massoud F, Moride Y, Arsenault L, and Afilalo M (2001) Evaluation of the Confusion Assessment Method (CAM) As a Screening Tool for Delirium in the Emergency Room, *General Hospital Psychiatry,* 23(1):20-5.

Moretti R, Torre P, Antonello RM, Cattaruzza T, and Cazzato G (2004) Cholinesterase Inhibition As a Possible Therapy for Delirium in Vascular Dementia: a Controlled, Open 24-Month Study of 246 Patients, *American Journal of Alzheimer's Disease & Other Dementias,* 19(6):333-9.

Morrison RS, Magaziner J, Gilbert M, Koval KJ, McLaughlin MA, Orosz G, Strauss E, and Siu AL (2003) Relationship Between Pain and Opioid Analgesics on the Development of Delirium Following Hip Fracture, *Journals of Gerontology Series A-Biological Sciences and Medical Sciences,* 58(1):76-81.

National Clinical Guideline Centre. Alcohol use disorders: diagnosis and clinical management of alcohol-related physical complications. NICE clinical guideline 100 (2010). Available from http://guidance.nice.org.uk/CG100

National Collaborating Centre for Mental Health. Alcohol Dependence and Harmful Alcohol Use. NICE Clinical Guideline (Expected in 2011).

National Collaborating Centre for Nursing and Supportive Care (2003) *Infection Control: Prevention of Healthcare-Associated Infections in Primary and Community Care. National Clinical Guideline Number 2*, Brentford: Richard Wells Research Centre, Thames Valley University. Available from: http://guidance.nice.org.uk/CG2

National Collaborating Centre for Nursing and Supportive Care (2005) *Violence - The Short-Term Management of Disturbed/Violent Behaviour in in-Patient Psychiatric Settings and Emergency Departments. National Clinical Guideline Number 25*, London: Royal College of Nursing. Available from: http://guidance.nice.org.uk/CG25

National Collaborating Centre for Chronic Conditions (2006) *Parkinson's Disease - National Clinical Guideline for Diagnosis and Management in Primary and Secondary Care. National Clinical Guideline Number 35*, London: Royal College Of Physicians. Available from: http://guidance.nice.org.uk/CG35

National Collaborating Centre for Acute Care (2006) *Nutrition Support in Adults: Oral Nutrition Support, Enteral Tube Feeding and Parenteral Nutrition. National Clinical Guideline Number 32*, London: National Collaborating Centre for Acute Care. Available from: http://guidance.nice.org.uk/CG32

National Collaborating Centre for Mental Health (2007) *Dementia. A NICE-SCIE Guideline on Supporting People With Dementia and Their Carers in Health and Social Care. National Clinical Guideline Number 42*, Leicester & London: The British Psychological Society & The Royal College Of Psychiatrists. Available from: http://guidance.nice.org.uk/CG42

National Collaborating Centre for Chronic Conditions (2008) *Stroke: National Clinical Guideline for Diagnosis and Initial Management of Acute Stroke and Transient Ischaemic Attack (TIA). National Clinical Guideline Number 68*, London: Royal College Of Physicians. Available from: http://guidance.nice.org.uk/CG68

National Collaborating Centre for Mental Health (2008) *Drug Misuse: Opioid Detoxification. National Clinical Guideline Number 52*, Leicester and London: The British Psychological Society and The Royal College of Psychiatrists. Available from: http://guidance.nice.org.uk/CG52

National Collaborating Centre for Women's and Children's Health (2008) *Surgical Site Infection - Prevention and Treatment of Surgical Site Infection. National Clinical Guideline Number 74*, London: Royal College of Obstetricians and Gynaecologists. Available from: http://guidance.nice.org.uk/CG74

National Collaborating Centre for Primary Care (2009) *Medicines Adherence - Involving Patients in Decisions About Prescribed Medicines and Supporting Adherence. National Clinical Guideline Number 76*, London: National Collaborating Centre for Primary Care and Royal College of General Practitioners. Available from: http://guidance.nice.org.uk/CG76

National Collaborating Centre for Mental Health (2009) *Schizophrenia. Core Interventions in the Treatment and Management of Schizophrenia in Primary and Secondary Care (Update). National Clinical Guideline Number 82*, London: National Collaborating Centre for Mental Health. Available from: http://guidance.nice.org.uk/CG82

National Institute for Health and Clinical Excellence (2002) *Schizophrenia - the Clinical Effectiveness and Cost Effectiveness of Newer Atypical Antipsychotic Drugs for Schizophrenia. NICE Technology Appraisal Guidance 43*, London: National Institute for Health and Clinical Excellence. Available from: http://guidance.nice.org.uk/TA43

National Institute for Health and Clinical Excellence (2008) *Guide to the Methods of Technology Appraisal*, London: National Institute for Health and Clinical Excellence. Available from: http://www.nice.org.uk/media/B52/A7/TAMethodsGuideUpdatedJune2008.pdf

National Institute for Health and Clinical Excellence (2009) *The Guidelines Manual*, London: National Institute for Health and Clinical Excellence. Available from: http://www.nice.org.uk/media/5F2/44/The_guidelines_manual_2009_-_All_chapters.pdf

National Institute for Health and Clinical Excellence (2009) *Donepezil, Galantamine, Rivastigmine (Review) and Memantine for the Treatment of Alzheimer's Disease (Amended). NICE Technology Appraisal Guidance 111*, London: National Institute for Health and Clinical Excellence. Available from: http://guidance.nice.org.uk/TA111

Naughton BJ, Moran MB, Kadah H, Heman-Ackah Y, and Longano J (1995) Delirium and Other Cognitive Impairment in Older Adults in an Emergency Department, *Annals of Emergency Medicine,* 25(6):751-5.

Naughton BJ, Saltzman S, Ramadan F, Chadha N, Priore R, and Mylotte JM (2005) A Multifactorial Intervention to Reduce Prevalence of Delirium and Shorten Hospital Length of Stay, *Journal of the American Geriatrics Society,* 53(1):18-23.

Netten A and Curtis L (1996) *Unit Costs of Health and Social Care 1996*, University of Canterbury, Kent: Personal Social Services Research Unit. Available from: www.pssru.ac.uk

Netten A, Bebbington A, Darton R, Forder J, and Miles K (1998) 1996 Survey of Care Homes for Elderly People. Final Report, *PSSRU Discussion Paper 1423/2,* University of Canterbury, Kent:Personal Social Services Research Unit.

Netten A, Darton R, Bebbington A, Forder J, Brown P, and Mummery K (2001) Residential and Nursing Home Care of Elderly People With Cognitive Impairment: Prevalence, Mortality and Costs, *Aging and Mental Health,* 5:14-22.

Ni Chonchubhair A, Valacio R, Kelly J, and O'Keefe S (1995) Use of the Abbreviated Mental Test to Detect Postoperative Delirium in Elderly People, *British Journal of Anaesthesia,* 75(4):481-2.

NICE Short Clinical Guidelines Technical Team (2007) *Acutely Ill Patients in Hospital: Recognition of and Response to Acute Illness in Adults in Hospital. National Clinical Guideline Number 50*, London: National Institute for Health and Clinical Excellence . Available from: http://guidance.nice.org.uk/CG50

Nightingale S, Holmes J, Mason J, and House A (2001) Psychiatric Illness and Mortality After Hip Fracture, *Lancet,* 357(9264):1264-5.

Nitschke LF, Schlosser CT, Berg RL, Selthafner JV, Wengert TJ, and Avecilla CS (1996) Does Patient-Controlled Analgesia Achieve Better Control of Pain and Fewer Adverse Effects Than Intramuscular Analgesia? A Prospective Randomized Trial, *Archives of Surgery,* 131(4):417-23.

O'Keefe ST and Lavan JN (1999) Clinical Significance of Delirium Subtypes in Older People, *Age and Ageing,* 28(2):115-9.

O'Keeffe S and Lavan J (1997) The Prognostic Significance of Delirium in Older Hospital Patients, *Journal of the American Geriatrics Society,* 45(2):174-8.

O'Keeffe ST and Lavan JN (1996) Subcutaneous Fluids in Elderly Hospital Patients With Cognitive Impairment, *Gerontology,* 42(1):36-9.

O'Keeffe ST, Mulkerrin EC, Nayeem K, Varughese M, and Pillay I (2005) Use of Serial Mini-Mental State Examinations to Diagnose and Monitor Delirium in Elderly Hospital Patients, *Journal of the American Geriatrics Society,* 53(5):867-70.

Olin K, Eriksdotter-Jönhagen M, Jansson A, Herrington MK, Kristiansson M, and Permert J (2005) Postoperative Delirium in Elderly Patients After Major Abdominal Surgery, *British Journal of Surgery,* 92(12):1559-64.

Ouimet S, Kavanagh BP, Gottfried SB, and Skrobik Y (2007) Incidence, Risk Factors and Consequences of ICU Delirium, *Intensive Care Medicine,* 33(1):66-73.

Owens JF and Hutelmyer CM (1982) The Effect of Preoperative Intervention on Delirium in Cardiac Surgical Patients, *Nursing Research,* 31(1):60-2.

Pae C-U, Lee S-J, Lee C-U, Lee C, and Paik I-H (2004) A Pilot Trial of Quetiapine for the Treatment of Patients With Delirium, *Human Psychopharmacology,* 19(2):125-7.

Pandharipande P, Shintani A, Peterson J, Pun BT, Wilkinson GR, Dittus RS, Bernard GR, and Ely EW (2006) Lorazepam Is an Independent Risk Factor for Transitioning to Delirium in Intensive Care Unit Patients, *Anesthesiology,* 104(1):21-6.

Pandharipande P, Cotton BA, Shintani A, Thompson J, Pun BT, Morris JA, Jr., Dittus R, and Ely EW (2008) Prevalence and Risk Factors for Development of Delirium in Surgical and Trauma Intensive Care Unit Patients, *Journal of Trauma,* 65(1):34-41.

Papaioannou A, Fraidakis O, Michaloudis D, Balalis C, and Askitopoulou H (2005) The Impact of the Type of Anaesthesia on Cognitive Status and Delirium During the First Postoperative Days in Elderly Patients, *European Journal of Anaesthesiology,* 22(7):492-9.

Parellada E, Baeza I, de Pablo J, and Martinez G (2004) Risperidone in the Treatment of Patients With Delirium, *Journal of Clinical Psychiatry,* 65(3):348-53.

Patten SB, Williams JV, Haynes L, McCruden J, and Arboleda-Florez J (1997) The Incidence of Delirium in Psychiatric Inpatient Units, *Canadian Journal of Psychiatry - Revue Canadienne De Psychiatrie,* 42(8):858-63.

Personal Social Services Research Unit (PSSRU) (2007) *Dementia UK: The Full Report*, London: The Alzheimer's Society. Available from: http://www.alzheimers.org.uk/site/scripts/download.php?fileID=2

Peterson JF, Pun BT, Dittus RS, Thomason JW, Jackson JC, Shintani AK, and Ely EW (2006) Delirium and Its Motoric Subtypes: a Study of 614 Critically Ill Patients, *Journal of the American Geriatrics Society,* 54(3):479-84.

Pisani MA, Araujo KL, Van Ness PH, Zhang Y, Ely EW, and Inouye SK (2006) A Research Algorithm to Improve Detection of Delirium in the Intensive Care Unit, *Critical Care,* 10(4):R121.

Pisani MA, Murphy TE, Van Ness PH, Araujo KL, and Inouye SK (2007) Characteristics Associated With Delirium in Older Patients in a Medical Intensive Care Unit, *Archives of Internal Medicine,* 167(15):1629-34.

Pisani MA, Murphy TE, Araujo KL, Slattum P, Van Ness PH, and Inouye SK (2009) Benzodiazepine and Opioid Use and the Duration of Intensive Care Unit Delirium in an Older Population, *Critical Care Medicine,* 37(1):177-83.

Pitkala KH (2005) Prognostic Significance of Delirium in Frail Older People, *Dementia and Geriatric Cognitive Disorders,* 19(2-3):158-63.

Pitkala KH, Laurila JV, Strandberg TE, and Tilvis RS (2006) Multicomponent Geriatric Intervention for Elderly Inpatients With Delirium: a Randomized, Controlled Trial, *Journals of Gerontology Series A-Biological Sciences and Medical Sciences,* 61(2):176-81.

Pitkala KH, Laurila JV, Strandberg TE, Kautiainen H, Sintonen H, and Tilvis RS (2008) Multicomponent Geriatric Intervention for Elderly Inpatients With Delirium: Effects on Costs and Health-Related Quality of Life, *Journals of Gerontology Series A-Biological Sciences and Medical Sciences,* 63A(1):56-61.

Pompei P, Foreman M, Rudberg MA, Inouye SK, Braund V, and Cassel CK (1994) Delirium in Hospitalized Older Persons: Outcomes and Predictors, *Journal of the American Geriatrics Society,* 42(8):809-15.

Pompei P, Foreman M, Cassel CK, Alessi C, and Cox D (1995) Detecting Delirium Among Hospitalized Older Patients, *Archives of Internal Medicine,* 155(3):301-7.

Porta M,Greenland S, and Last J (eds), *A dictionary of epidemiology,* 5th edition, New York, Oxford University Press, 2008.

Porthouse J, Cockayne S, King C, Saxon L, Steele E, Aspray T, Baverstock M, Birks Y, Dumville J, Francis R, Iglesias C, Puffer S, Sutcliffe A, Watt I, and Torgerson DJ (2005) Randomised Controlled Trial of Calcium and Supplementation With Cholecalciferol (Vitamin D3) for Prevention of Fractures in Primary Care, *BMJ,* 330(7498):1003.

Prakanrattana U and Prapaitrakool S (2007) Efficacy of Risperidone for Prevention of Postoperative Delirium in Cardiac Surgery, *Anaesthesia and Intensive Care,* 35(5):714-9.

Radtke FM, Franck M, Schneider M, Luetz A, Seeling M, Heinz A, Wernecke KD, and Spies CD (2008) Comparison of Three Scores to Screen for Delirium in the Recovery Room, *British Journal of Anaesthesia,* 101(3):338-43.

Rahkonen T, Eloniemi-Sulkava U, Paanila S, Halonen P, Sivenius J, and Sulkava R (2001) Systematic Intervention for Supporting Community Care of Elderly People After a Delirium Episode, *International Psychogeriatrics,* 13(1):37-49.

Ramirez-Bermudez J (2006) Frequency of Delirium in a Neurological Emergency Room, *Journal of Neuropsychiatry and Clinical Neurosciences,* 18(1):108-12.

Ranhoff AH, Rozzini R, Sabatini T, Cassinadri A, Boffelli S, and Trabucchi M (2006) Delirium in a Sub-Intensive Care Unit for the Elderly: Occurrence and Risk Factors, *Aging-Clinical and Experimental Research,* 18(5):440-5.

Redelmeier DA, Thiruchelvam D, and Daneman N (2008) Delirium After Elective Surgery Among Elderly Patients Taking Statins, *CMAJ: Canadian Medical Association Journal,* 179(7):645-52.

Rizzo JA, Bogardus ST, Jr., Leo-Summers L, Williams CS, Acampora D, and Inouye SK (2001) Multicomponent Targeted Intervention to Prevent Delirium in Hospitalized Older Patients: What Is the Economic Value?, *Medical Care,* 39(7):740-52.

Roberts B, Rickard CM, Rajbhandari D, Turner G, Clarke J, Hill D, Tauschke C, Chaboyer W, and Parsons R (2005) Multicentre Study of Delirium in ICU Patients Using a Simple Screening Tool, *Australian Critical Care,* 18(1):6-9.

Robinson SB and Rosher RB (2002) Can a Beverage Cart Help Improve Hydration?, *Geriatric Nursing,* 23(4):208-11.

Robinson TN, Raeburn CD, Angles EM, and Moss M (2008) Low Tryptophan Levels Are Associated With Postoperative Delirium in the Elderly, *American Journal of Surgery,* 196(5):670-4.

Robinson TN, Raeburn CD, Tran ZV, Angles EM, Brenner LA, and Moss M (2009) Postoperative Delirium in the Elderly: Risk Factors and Outcomes, *Annals of Surgery,* 249(1):173-8.

Rockwood K (1994) Increasing the Recognition of Delirium in Elderly Patients, *Journal of the American Geriatrics Society,* 42(3):252-6.

Rockwood K, Cosway S, Carver D, Jarrett P, Stadnyk K, and Fisk J (1999) The Risk of Dementia and Death After Delirium, *Age and Ageing,* 28(6):551-6.

Rolfson DB, McElhaney JE, Rockwood K, Finnegan BA, Entwistle LM, Wong JF, and Suarez-Almazor ME (1999) Incidence and Risk Factors for Delirium and Other Adverse Outcomes in Older Adults After Coronary Artery Bypass Graft Surgery, *Canadian Journal of Cardiology,* 15(7):771-6.

Rolfson DB, McElhaney JE, Jhangri GS, and Rockwood K (1999b) Validity of the Confusion Assessment Method in Detecting Postoperative Delirium in the Elderly, *International Psychogeriatrics,* 11(4):431-8.

Rudolph JL, Babikian VL, Birjiniuk V, Crittenden MD, Treanor PR, Pochay VE, Khuri SF, and Marcantonio ER (2005) Atherosclerosis Is Associated With Delirium After Coronary Artery Bypass Graft Surgery, *Journal of the American Geriatrics Society,* 53(3):462-6.

Rudolph JL, Jones RN, Grande LJ, Milberg WP, King EG, Lipsitz LA, Levkoff SE, and Marcantonio ER (2006) Impaired Executive Function Is Associated With Delirium After Coronary Artery Bypass Graft Surgery, *Journal of the American Geriatrics Society,* 54(6):937-41.

Rudolph JL, Jones RN, Rasmussen LS, Silverstein JH, Inouye SK, and Marcantonio ER (2007) Independent Vascular and Cognitive Risk Factors for Postoperative Delirium, *American Journal of Medicine,* 120(9):807-13.

Rudolph JL, Marcantonio ER, Culley DJ, Silverstein JH, Rasmussen LS, Crosby GJ, and Inouye SK (2008) Delirium Is Associated With Early Postoperative Cognitive Dysfunction, *Anaesthesia,* 63(9):941-7.

Salkeld G, Cameron ID, Cumming RG, Easter S, Seymour J, Kurrle SE, and Quine S (2000) Quality of Life Related to Fear of Falling and Hip Fracture in Older Women: a Time Trade Off Study, *BMJ,* 320(7231):341-6.

Sandberg O, Franklin KA, Bucht G, and Gustafson Y (2001) Sleep Apnea, Delirium, Depressed Mood, Cognition, and ADL Ability After Stroke, *Journal of the American Geriatrics Society,* 49(4):391-7.

Santana Santos F, Wahlund LO, Varli F, Tadeu Velasco I, and Eriksdotter Jönhagen M (2005) Incidence, Clinical Features and Subtypes of Delirium in Elderly Patients Treated for Hip Fractures, *Dementia and Geriatric Cognitive Disorders,* 20(4):231-7.

Santos FS, Velasco IT, and Fraquas R (2004) Risk Factors for Delirium in the Elderly After Coronary Artery Bypass Graft Surgery, *International Psychogeriatrics,* 16(2):175-93.

Sasajima Y, Sasajima T, Uchida H, Kawai S, Haga M, Akasaka N, Kusakabe M, Inaba M, Goh K, and Yamamoto H (2000) Postoperative Delirium in Patients With Chronic Lower Limb Ischaemia: What Are the Specific Markers?, *European Journal of Vascular and Endovascular Surgery,* 20(2):132-7.

Schor JD, Levkoff SE, Lipsitz LA, Reilly CH, Cleary PD, Rowe JW, and Evans DA (1992) Risk Factors for Delirium in Hospitalized Elderly, *JAMA: Journal of the American Medical Association,* 267(6):827-31.

Scott NB, Turfrey DJ, Ray DA, Nzewi O, Sutcliffe NP, Lal AB, Norrie J, Nagels WJ, and Ramayya GP (2001) A Prospective Randomized Study of the Potential Benefits of Thoracic Epidural Anesthesia and Analgesia in Patients Undergoing Coronary Artery Bypass Grafting, *Anesthesia and Analgesia,* 93(3):528-35.

Shea JD (1975) Pressure Sores: Classification and Management, *Clinical Orthopaedics and Related Research,*(112):89-100.

Sheng AZ (2006) Delirium Within Three Days of Stroke in a Cohort of Elderly Patients, *Journal of the American Geriatrics Society,* 54(8):1192-8.

Shulman KI (2005) Incidence of Delirium in Older Adults Newly Prescribed Lithium or Valproate: A Population-Based Cohort Study, *Journal of Clinical Psychiatry,* 66(4):424-7.

Skrobik YK, Bergeron N, Dumont M, and Gottfried SB (2004) Olanzapine Vs Haloperidol: Treating Delirium in a Critical Care Setting, *Intensive Care Medicine,* 30(3):444-9.

Stenwall E, Jönhagen ME, Sandberg J, and Fagerberg I (2008) The Older Patient's Experience of Encountering Professional Carers and Close Relatives During an Acute Confusional State: an Interview Study, *International Journal of Nursing Studies,* 45(11):1577-85.

Stenwall E, Sandberg J, Jönhagen ME, and Fagerberg I (2008) Relatives' Experiences of Encountering the Older Person With Acute Confusional State: Experiencing Unfamiliarity in a Familiar Person, *International Journal of Older People Nursing,* 3(4):243-51.

The NHS Information Centre Social Care Statistics (2009) *Personal Social Services Expenditure and Unit Costs England, 2007-08*, Leeds: The NHS Information Centre. Available from: http://www.ic.nhs.uk/statistics-and-data-collections/social-care/adult-social-care-information/personal-social-services-expenditure-and-unit-costs:-england-2007-08

Thomason JW, Shintani A, Peterson JF, Pun BT, Jackson JC, and Ely EW (2005) Intensive Care Unit Delirium Is an Independent Predictor of Longer Hospital Stay: a Prospective Analysis of 261 Non-Ventilated Patients, *Critical Care,* 9(4):R375-R381.

Uldall KK, Harris VL, and Lalonde B (2000) Outcomes Associated With Delirium in Acutely Hospitalized Acquired Immune Deficiency Syndrome Patients, *Comprehensive Psychiatry,* 41(2):88-91.

van der Mast RC, van den Broek WW, Fekkes D, Pepplinkhuizen L, and Habbema JD (1999) Incidence of and Preoperative Predictors for Delirium After Cardiac Surgery, *Journal of Psychosomatic Research,* 46(5):479-83.

van Munster BC, Korevaar JC, de Rooij SE, Levi M, and Zwinderman AH (2007) The Association Between Delirium and the Apolipoprotein E Epsilon4 Allele in the Elderly, *Psychiatric Genetics,* 17(5):261-6.

Van Rompaey B., Elseviers MM, Schuurmans MJ, Shortridge-Baggett LM, Truijen S, and Bossaert L (2009) Risk Factors for Delirium in Intensive Care Patients: A Prospective Cohort Study, *Critical Care,* 13(3)

Veliz-Reissmüller G, Agüero Torres H, van der Linden J, Lindblom D, and Eriksdotter Jönhagen M (2007) Pre-Operative Mild Cognitive Dysfunction Predicts Risk for Post-Operative Delirium After Elective Cardiac Surgery, *Aging-Clinical and Experimental Research,* 19(3):172-7.

Wanich C, Sullivan-Marx E, Gottlieb G, and Johnson J (1992) Functional Status Outcomes of a Nursing Intervention in Hospitalized Elderly, *Image - the Journal of Nursing Scholarship,* 24(3):201-8.

Ward S, Lloyd JM, Pandor A, Holmes M, Ara R, Ryan A, Yeo W, and Payne N (2007) A Systematic Review and Economic Evaluation of Statins for the Prevention of Coronary Events, *Health Technology Assessment,* 11(14):1-iv.

Weed HG, Lutman CV, Young DC, and Schuller DE (1995) Preoperative Identification of Patients at Risk for Delirium After Major Head and Neck Cancer Surgery, *Laryngoscope,* 105(10):1066-8.

Williams-Russo P, Urquhart BL, Sharrock NE, and Charlson ME (1992) Post-Operative Delirium: Predictors and Prognosis in Elderly Orthopedic Patients, *Journal of the American Geriatrics Society,* 40(8):759-67.

Wong DM, Niam T, Bruce JJ, and Bruce DG (2005) Quality Project to Prevent Delirium After Hip Fracture, *Australasian Journal on Ageing,* 24(3):174-7.

World Health Organization (1992) *The ICD-10 Classification of Mental and Behavioural Disorders Clinical Descriptions and Diagnostic Guidelines*, Geneva: World Health Organization. Available from: http://www.who.int/classifications/icd/en/bluebook.pdf

Yates C, Stanley N, Cerejeira JM, Jay R, and Mukaetova-Ladinska EB (2009) Screening Instruments for Delirium in Older People With an Acute Medical Illness, *Age and Ageing,* 38(2):235-7.

Yildizeli B (2005) Factors Associated With Postoperative Delirium After Thoracic Surgery, *Annals of Thoracic Surgery,* 79(3):1004-9.

Yoshimura Y, Kubo S, Shirata K, Hirohashi K, Tanaka H, Shuto T, Takemura S, and Kinoshita H (2004) Risk Factors for Postoperative Delirium After Liver Resection for Hepatocellular Carcinoma, *World Journal of Surgery,* 28(10):982-6.

Zakriya KJ, Christmas C, Wenz JF, Sr., Franckowiak S, Anderson R, and Sieber FE (2002) Preoperative Factors Associated With Postoperative Change in Confusion Assessment Method Score in Hip Fracture Patients, *Anesthesia and Analgesia,* 94(6):1628-32.

Zou Y, Cole MG, Primeau FJ, McCusker J, Bellavance F, and LaPante JL (1998) Detection and Diagnosis of Delirium in the Elderly: Psychiatrist Diagnosis, Confusion Assessment Method, or Consensus Diagnosis?, *International Psychogeriatrics,* 10(3):303-8.

Appendices A–K are in separate files