APPENDIX 12: 2005 CLINICAL EVIDENCE – STUDY CHARACTERISTICS TABLES FROM PREVIOUS GUIDELINE (CG25)

Abbreviations

BPRS	Brief Psychiatric Scale Ratings
CGI	Clinical Global Impression
DSM(-III-R, -IV	Diagnostic Statistical Manual (third edition revised, fourth edition)
ECT	electroconvulsive therapy
EPS	extrapyramidal symptoms
IM	intramuscular
MOAS	Modified Overt Aggression Scale
n	number of participants
NOSIE	Nurses' Observation Scale for Inpatient Evaluation
OAS	Overt Aggression Scale
PANSS	Positive and Negative Syndrome Scale
RAPP	Routine Assessment of Patient Progress
SO	special observation
SOAS	Staff Observation Aggression Scale
VAS	visual analogue scale

1.1 ENVIRONMENT AND ALARM SYSTEMS

1.1.1 Mistral 2002

Source	Study design	Aims of study	Outcome measures	Results
Mistral and colleagues (2002) Country: UK Evidence level: 2-	Qualitative design: grounded theory and thematic content analysis and psychometric tests. Settings: psychiatric high care ward with a seclusion facility. Population: 36 staff.	To evaluate changes in attitude following upgrading of the physical environment, regular ward meetings, personal alarms, training in risk assessment, control and restraint techniques, and introduction of clear rules and sanctions.	Semi-structured interviews with 36 nursing and medical staff. Attitude measure (to measure attitude of staff to service users). Ward atmosphere scale. Records of admissions, staff illness and use of seclusion.	Pre- and postintervention phase: 7 months apart. Key themes from interviews: 5 positive areas of change – communication, team cohesiveness, relations with management, clarity and structure and quality of service user care. Five areas of continuing concern: resources and staffing levels, admittance criteria, divisions between nursing staff and medical staff, stress and safety. Attitude measure: significant change in 2 out of 7 subscales from preintervention to postintervention phase. Skill and knowledge adequacy: <0.05. Self-esteem in this work: <0.001. Ward atmosphere scale: significant change in 2/10 subscales from preintervention to postintervention phase. Involvement (activity levels of service users): <0.002. Practical orientation (preparation for release from hospital): <0.05. Records: admissions – no significant differences in a 2-year period. Seclusions: reduction from a mean of 3 times in 1996 to once a week in 1998.

				Staff illness: a reduction of 40% in staff sick leave over 2 years.

Reviewer's comments:

- Five staff refused to participate, however it is not clear whether the total staff compliment is 36 or 41. Reference is made to theoretical tradition of grounded theory without clarity on use of theory.
- The interview data is presented as frequency counts of coded data with limited contextually supportive evidence. The small size of interviews resulted in insignificant results on the subscales of the psychometric tests.
- Respondent validation was not undertaken.

1.1.2 Nijman 1999

Source	Study design	Aims of study	Outcome measures	Results
Nijman and colleagues (1999) Country: Belgium Evidence level: 2-	Correlation study (with weak control). Setting: 2 closed observation wards. Population: 354 (212 male) were admitted to the wards during the study period.	To examine association between ward crowding and increase in aggressive incidents. To examine if enlargement of ward space result in a decline in aggressive incidents.	All verbal and physical aggressive acts by service users admitted to the 2 observation wards. Acts were recorded using the revised Staff Observation Aggression Scale (SOAS). The study period was between 1 February and 15 December 1996. In the middle of the study (9 July), a courtyard was opened in 1 ward (ward 1). The inner courtyard was connected to 2 entrances to the ward, which increased the service users' opportunity to walk around freely. Ward 2 did not have a courtyard. The frequency of aggression on ward 1 was compared with that on ward 2, before and after the spatial enlargement of ward 1.	A total of 226 aggressive incidents were recorded during the study period. Aggressive episodes varied from 0 to 15, the average being 4.9 incidents per week. 18 (8%) of the incidents led to mild or moderate injury to the victims. A Pearson product-moment correlation was calculated between the weekly occupancy rates of the wards and the frequency of aggression, as measured by the number of incidents per service user. A modest correlation was found between weekly occupation rates and the total number of incidents per service user ($r=0.21$, $p<0.05$).

Reviewer's comments:

- No information is provided for the number of incidents of service user aggression in relation to ward setting. The reporting in the results section in extremely poor. Although the authors suggest that this study is a comparison of the 2 wards, they fail to provide any useful information that would support this suggestion.
- The statistical analysis (Pearson product-moment correlation) is not appropriate for the analysis of this data. The suggestion by the authors that "a modest correlation was found between weekly occupation rates and the total number of incidents per service user ($r=0.21$, $p<0.05$)" is not supported by the design or the summary statistic.
- The assumption made by the authors is that crowding is the only factor related to aggressive incidents, a position not supported by the literature.
- The results of this study should be treated with caution.

1.2 OBSERVATION

1.2.1 Bowles 2001

Source	Study design	Aims of study	Outcome measures	Results
Bowles and Doods (2001) Country: UK Evidence level: 2-	Before and after study design without controls. Setting/population: 21-bed acute in-patient ward for males below 65 years.	To assess the therapeutic value of dismantling formal observation and replacing it with 1-to-1 interaction and activities.	Levels of: • suicide • absconding • staff sickness • self-harm • use of staff time • costs.	After 6 months: • Formal observation rare. After 18 months: • 1-to-1 observation never used; 5-10 minute checks rare. • Nurses provided programme of weekly activities for service users. • Service users more involved in their care and ward decisions. • Deliberate self-harm reduced by almost two-thirds. • Violence and aggression reduced by almost one-third. • Staff sickness reduced by two-thirds. • Absconding reduced by almost half. • 95% of service users receive daily structured time with nurses. • No increase in suicides. Over 12 months: • £45,000 saved on staffing budget.

Reviewer' comments:

- Authors conclude that formal observation is an 'outmoded ritual of mental health nursing'.
- Authors maintain that the 'gift' of a nurse's time is the most effective intervention.
- Authors argue that nurses should decide how to 'gift' their time.
- The authors acknowledge that the study is too small for the results to be generalisable and is not adequate basis for policy or practice change.
- This is not an appropriate study design for assessing therapeutic value or effectiveness.
- This was an evaluation of a change in practice, rather than a research project.

1.2.2 Jones 2000

Source	Study design	Aims of study	Outcome measures	Results
Jones and colleagues (2000) Country: UK Evidence level: 2-	Survey. Setting: 3 site mental health care trust (108 acute beds). Population: 54 service users who were psychiatric inpatients and experienced close or constant observation (2 highest levels out of 4 possible).	To identify service user preferences and feelings about close and constant observation.	Repertory grid technique to measure service user's feeling and preferences about close and constant observation. Service users interviewed either while being observed or within 5 days of a period of observation ending.	Out of 54 service users, 25 agreed to be interviewed, but only 18 completed the interviews. Data was analysed using Flexigrid, SPSS, and t tests for paired and independent samples. Service users commented that they felt safest when they were being observed either by a nurse they knew or by a nurse who talked to them. The inverse was also true. Both were magnified for service users with risk of self-harm (p=0.002). Services users preferred to be observed by nurses who they knew (p <0.0002) or who talked to them (p <0.0002). Suicidal service users disliked being observed by nurses they didn't know (p=0.0001) and by nurses who didn't talk to them (p=0.0001).

Reviewer's comments:

- Authors comment that the role of the observer is the most important factor in shaping service user perceptions of observation.
- Small sample size; results are not generalisable.
- Limitations of study are discussed – suitability of Flexigrid for all service users.
- Only 2 service users of final sample were observed for the risk of harm to others.

1.2.3 Nielson 2001

Source	Study design	Aims of study	Outcome measures	Results
Neilson and Brennan (2001) Country: UK Evidence level: 2-	Cross-sectional audit within a retrospective study. Setting: high dependency psychiatric unit. Population: 34 staff (trained and untrained), 144 adult service users' special observation (SO) records (includes elderly).	To elicit staff attitudes towards and knowledge of SO policy. To ascertain any differences across wards in terms of implementation and staff attitudes towards and knowledge of hospital observation policy.	To ascertain implementation of SO and staff knowledge of hospital policy: • knowledge test questionnaire was given to selected staff (both trained and untrained) • semistructured interview were conducted with some staff across all 4 wards • scoring schedule for audit of 144 SORSs randomly chosen by random number table (schedule piloted and amended).	(Four levels of SO in order of severity – red, amber, blue, green) blue most commonly imposed for 3 wards (56.25%). Ward C used more of a mixture. Audit of SORSs: • All wards scored low on review date and authorising signature. • Assessed risk stated on only 26.4% of SORSs. • Assigning staff on block to SO led to missed time periods (64.5%). Staff interviews: • Nursing staff felt less involved in decision-making than they would have liked (94.2% – too medically dominated) – felt SO often used 'just in case' (82.4% – blue level used too frequently). • Impossible for staffing levels to meet current demands of SO (73.6%). • Communication and documentation had improved since introduction of SO policy (35.29%). • Poor medical review of SO (32.36%). • Red level could provoke disturbed service users (29.41%). • Gender needed greater consideration when allocating staff to SO (23.6%). Knowledge test: • All staff had good knowledge of hospital policy on SO.

Reviewer's comments:

- Authors acknowledge that lack of randomisation of staff limits generalisability.
- Amended tools not piloted or validated.
- Does not differentiate between SO used for to prevent self-harm and SO used to prevent harm to others.
- Authors conclude that the audit provides evidence that the SO is not being adhered to in practice, as intended.

1.2.4 Shugar 1990

Source	Study design	Aims of study	Outcome measures	Results
Shugar and Rehaluk (1990) Country: Canada Evidence level: 2-	Retrospective cohort with controls. (Control group was made up of service users entering unit immediately subsequent to each subject's admission). Setting: psychiatric teaching unit. Population: 102 adult, civil and forensic (also geriatric) with 102 control subjects.	To ascertain reason for CO and to assess the effectiveness of CO.	Incidence of CO.	102 incidences of CO identified. CO used for violence management: • Over-stimulation – 25 • Violence to property – 6 • Potential violence to others – 5 • Actual violence to others – 4. Service users requiring long-term observation distinguished from those requiring short-term observation by greater risk of self-harm ($p < 0.04$), history of violence to property ($p < 0.05$), multiple reasons for being placed on CO ($p < 0.04$). More likely to receive ECT ($p < 0.03$) or restraints ($p < 0.05$). Six demographic and clinical factors differentiating subjects requiring CO from those not requiring it – history of self-harm, involuntary status on admission, belonged to 2 lowest social classes, past history of violence to property, female, past history of violence. Authors offer tentative conclusion of positive effectiveness of CO, but note that study design makes these difficult to validate, because of confounders.
Reviewer's comments: • Authors admit that design constraints make effectiveness difficult to assess and therefore offers only tentative conclusions. • Authors recommend that CO is only used as a short-term measure, but offer no research evidence to back this up. • While this article contains some useful information, the study design is weak and the conclusions must be treated with caution.				

1.2.5 Yong 1992

Source	Study design	Aims of study	Outcome measures	Results
Yonge and Stewin (1992) Country: Canada Evidence level: 2-	Qualitative. Setting: unspecified psychiatric context. Population: 8. psychiatric nurses.	To examine nurses' responses to undertaking close observation (CO).	Interviews (taped, transcribed and analysed using 'ethnograph' – programme for textual analysis).	The following themes emerged: • service user and nurse both on CO • CO alters the passage, meaning and use of time • CO as a dynamic rather than static relationship • CO enhances nurse's sense of powerlessness • nurses prepare for CO in advance • strategies for difficult situations • issues around watching service user eat • no nurse went into bathroom with service user • nurses have personal preferences for certain CO service users.
Reviewer's comments: • All themes are treated as equally important – does not indicate frequency. • Highlighted various common sense issues related to the stressful nature of CO. These results need to be treated with caution, due to small sample size.				

1.3 RISK AND PREDICTION

1.3.1 Cheung 1996

Source	Study design	Aims of study	Outcome measures	Results
Cheung (1996) Country: Australia Evidence level: 2+	8-week prospective cohort study Setting/population: Large psychiatric hospital; 220 service users	To assess the prevalence and nature of aggressive behaviour and the risk factors associated with aggressive behaviour.	Aggressive behaviour (measured by the SOAS). Demographics. Ward environment.	Multiple logical regression was used to calculate the effects of various service user characteristics on aggressive behaviour. Only male gender ($p < 0.01$) and duration of admission ($p < 0.05$) correlated with aggression status. When considering types of aggression, only male gender correlated with physical aggression ($p < 0.02$) and only duration of admission correlated with verbal aggression ($p < 0.05$). The most severe incidents tended to occur in the afternoon ($p < 0.001$). No other ward factors were significant.
Reviewer's comments: • Authors note that more variables could have been considered and note that the lack of correlation between diagnosis and aggression could have resulted from the majority of service users having schizophrenia, therefore, not allowing diagnostic variables to be fully tested. These findings are not generalisable and need to be validated in a number of settings.				

1.3.2 Ehmann 2001

Source	Study design	Aims of study	Outcome measures	Results
Ehmann (2001) Country: US Evidence level: 2-	2-year prospective cohort study (no control). Setting: 20-bed locked in tertiary care facility. Population: 78 treatment resistant or difficult diagnosis service users (17-65) [64 for prediction].	To describe rates and characteristics of aggression. To assess accuracy of incident reports. To discern relationships between types of aggression. To delineate clinical, historical and demographic characteristics of violent versus non-violent service users that have predictive validity.	Demographic information: • Diagnosis. • Number of incidents (Modified Overt Aggression Scale [MOAS] scores compared to hospital incident forms). • Psychopathology [rated with the Positive And Negative Syndrome Scale (PANSS), Routine Assessment of Patient Progress (RAPP), the Global Assessment of Functioning (DSM-IV axis V), Clinical Global Impression (CGI), degree of treatment resistance, DSM-III-R diagnoses, and the premorbid adjustment scale]. • Aggression (injury or threat to people, property, self). • Assault (injury to person). • Violence (defined as MOAS 3 o r4).	Statistical analysis was used (p=0.05=significance). 64% service users were assaultive. 26% assaulted others more than once. Incident reports underestimated violence by 45%, self-harm by 65% and property damage by 73%. Violence spread over admission, not only in first few weeks in long stay service users. Assault correlated with self-harm (p <0.0001) and aggression to objects (p <0.0001). Aggression to objects correlated to self-harm (p <0.0001) and verbal aggression (p <0.0001). Serious assaults failed to correlated with other forms of aggression. In first 4 weeks, mean MOAS scores for assault correlated with self-harm (p=0.002, object aggression (p <0.001) and verbal aggression (p <0.001). Violent (MOAS 3 or 4) versus non-violent groups: Best predictors were alcohol abuse in past year, female and diagnosed with non-paranoid schizophrenia. Using PANSS sensitivity=67%, specificity=91%, positive predictive value=71% (base rate=24%) 47% improvement over chance. If RAPP safety score substituted for PANSS total score sensitivity=81%, specificity=96%, positive predictive value=87%, improvement over chance=62%.

				Logical regression formula substituting RAPP total for PANSS total gave negative predictive value of 95% and a positive predictive value in random subset 1 of 78% and 62% in random subset 2. Best univariate predictors were poor premorbid adjustment, early age at illness onset, greater psychopathology and poor functioning at admission.
Reviewer's comments: • As only 6% of assaults occurred during night shift in year 1, no ratings were taken during the night shift in year 2. • Authors argue that results indicated that the relationship between assault and verbal aggression declines over time. After first month, only related to property damage and self-harm. Authors note that correlates of violence are dependent on definition. • Authors note that the inclusion of a clinical judgement item (RAPP safety item) greatly enhanced predictive validity.				

1.3.3 Kay 1988

Source	Study design	Aims of study	Outcome measures	Results
Kay (1988) Country: US Evidence level: 2-	Of the 3 studies reported, 2 were cross-sectional and 1 was a 3-month prospective cohort (only the prospective cohort is discussed here). Setting: 600-bed urban psychiatric hospital Population: 37 psychiatric service users on a chronic care unit (mostly with schizophrenia).	To test the predictive validity of the aggression risk profile in predicting psychiatric in- patient violence.	39 items contained within the tool covering 4 main areas: demographics current psychiatric diagnosis history of aggression clinical profile. Incidents of aggression were measured using MOAS.	Significant predictors of violence were found, 7 of these were specific to verbal or physical violence but not to both. Aggression generally was predicted by: younger age, more acutely ill, more threatening of violence by history and previously rated more agitated and labile in affect. Verbal aggression was predicted by: motor excitement, difficulty with gratification, depressed feelings. Physical violence was predicted by: anger, hostility, history of attacks on others, history of greater total aggression. Noted that history of aggression, although a good predictor on its own, did not enter into the regression formula for the strongest predictive combination because subsumed by other variables in the tool. After stepwise multiple aggression all types of aggression were significantly predicted: verbal ($p < 0.025$), physical ($p < 0.01$) and total aggression ($p < 0.05$).
Reviewer's comments: Authors note that while the best predictors were established by a combination of demographic and clinical variables, greater specificity was achieved by clinical variables. Authors note that the results may not be generalisable to different service user populations or in different settings. Authors note that the work needs validating.				

1.3.4 Kho 1998

Source	Study design	Aims of study	Outcome measures	Results
Kho (1998) Country: UK Evidence level: 2+	5-month prospective cohort study. Setting: 5 wards (4 acute admission, 1 locked) in 2 hospitals. Population: 360 acute psychiatric in- patients (wards had same catchment areas or similar populations).	To confirm reliability of MOAS (modified overt aggression scale) for use in everyday clinical practice. To examine whether commonly cited factors (demographic and clinical) associated with aggression are applicable to acute psychiatric admission units in general.	• Stage of admission. • Gender. • Ethnic group. • Type of. Ward. • Primary diagnosis. • Age.	Levels of aggression varied significantly over stage of admission. Women were more likely than men to be aggressive against objects. Asian women were more likely to exhibit aggression than other groups after the first 2 weeks of admission. Aggression was likely on the locked ward, although ward E had high levels of aggression. A diagnosis on mania or substance misuse was most likely to lead to verbal aggression. Individuals aged <30 years were more likely to be aggressive in the first 2 weeks of admission – significant only for verbal aggression and aggression against objects. MOAS rating was weighted towards serious aggressive incidents. MOAS Inter-rater reliability was moderate (weighted kappa 0.58) Authors suggest that this could be improved by providing training, selecting only the most highly qualified nurses to act as raters and limiting the number of raters.

Reviewer's comments:

- This is a well-designed study, which suggests that the MOAS rating scale can be applied to a clinical environment.
- Confounders controlled for using statistical analysis.
- The authors note that the study design does not allow causes and effects to be discriminated so that factors truly predictive of aggression cannot be identified.
- Authors note that other factors that might have confounded the results – such as ward environment, management of service users and interactions with staff – are not addressed.
- Others stress that results did not show that young black Afro-Caribbean males were highly aggressive.

1.3.5 Oulis 1996

Source	Study design	Aims of study	Outcome measures	Results
Oulis 1996 Country: Greece Evidence level: 2+	Cross-sectional. Setting: 2 inner-city psychiatric clinics. Population: 136 acute and chronic psychiatric in-patients.	To determine the prevalence and types of violence and the correlates.	• Verbal aggression. • Aggression against property. • Self-harm. • Physical aggression. Measured using the aggression risk profile and the MOAS.	Clinical and demographic variables were not significant in distinguishing non-aggressive and aggressive service users. Verbal aggression was significantly associated with agitation, low tolerance of frustration, difficulty in delaying gratification and anger (adjusted R squared=0.392). Aggression against property was significantly associated with bizarre behaviour or rituals (negatively), delusions, disorganised thinking and anger (adjusted R squared=0.271). Self-harm was significantly correlated with anger (adjusted R squared=0.133). Physical aggression was significantly correlated with agitation, disorganised thinking, anger and anti-social behaviour (adjusted R squared=0.288). Total anger was significantly correlated with bizarre behaviour or rituals (negatively), disorganised thinking and anger (adjusted R squared=0.355).

Reviewer's comments:

- All forms of aggressive behaviour were considered, therefore, all service users who scored 1 or above were included in the aggressive group.
- Authors note that their study confirms that of Kay and colleagues (1988).
- Authors assert that the results indicate the best predictors of aggression. However, these need to be confirmed by a prospective study.

1.3.6 Palmstierna 1989

Source	Study design	Aims of study	Outcome measures	Results
Palmstierna 1989 Country: Sweden Evidence level: 2+	Prospective cohort study. Setting: acute psychiatric. Population: 105 admitted and involuntary psychiatric service users.	To determine the factors that best predict violence in the short term, at 8 days and at 28 days.	SOAS. Main outcomes considered: • age • sex • diagnosis • history of violence • previous conviction for violent crime.	At 8 days, the only significant predictor was known previous damage to property or physical injury to person ($p < 0.05$). At 28 days, the only significant predictor was known abuse of drugs other than alcohol ($p < 0.05$). Because determination coefficients are very low (3.9 and 5.4% respectively), authors state that results indicate that risk factors are of limited value in predicting violence inside acute institutions. Also note that at 28 days females tend to more aggressive, but the result is not significant.

Reviewer's comments:

- Authors argue that certain risk factors for aggressive behaviour in outpatient settings are of limited value in the short-term prediction of violence amongst acute involuntary service users.
- Authors comment that different time perspectives demand different prediction procedures.
- Analysis by multivariate approach could explain why several factors did not reach significance, where they did in other papers.
- Factors chosen were from a list published in 1983, probably different in 2003.
- Follow-up period rather long – 8 and 28 days – different from other papers.

1.3.7 Yesavage 1984

Source	Study design	Aims of study	Outcome measures	Results
Yesavage (1984) Country: US Evidence level: 2-	3-year prospective cohort study. Setting: PICU in veterans' medical centre. Population: 70 adult male service users with schizophrenia (DSM-III criteria).	To assess correlates of violence for service users with schizophrenia during first 8 days of admission.	Low neuroleptic serum levels. Degree of psychotic symptoms. Act leading to admission. Military combat experience. Childhood discipline.	Best correlates for in-patient assaults were: • Low neuroleptic serum levels, violence prior to admission and schizophrenia rating on Brief Psychiatric Scale Ratings BPRS ($p < 0.01$).
Reviewer's comments: • Author argues that the implication of these findings is that in-patients with low serum levels of their neuroleptic may become violent because of under-control of their core schizophrenic symptoms. He postulates that this usually appears in service users with command hallucinations who act on them unexpectedly.				

1.4 RAPID TRANQUILLISATION / PHARMACOLOGICAL STUDIES

1.4.1 Battaglia 1997

Study	Population	Methods	Main intervention(s) and comparisons	Follow-up period	Outcomes (primary, secondary and adverse events) effect size, p-value
Battaglia (1997) Country: US Source of funding: supported from a grant by Wyeth-Ayerst Research. Evidence level: 1-	Setting: emergency departments in 5 universities or general hospitals. Participants: 98 psychotic, agitated and aggressive patients. Inclusion criteria: exhibition of psychosis and behavioural dyscontrol, scoring at least 5 on a scale of 1-7 or 3 or more of 11 psychosis/anxiety items from BPRS. Exclusion criteria: alcohol intoxication, allergic hypersensitivity, central nervous system depression, delirium, neuroleptic malignant syndrome, airway	Allocation: randomised. Blindness: double blind. Duration: 24-hours (98 service users over an 18-month period). Setting: 5 sites (emergency department). Baseline comparability: yes.	Group 1: Lorazepam 4 mg IM. Group 2: Haloperidol 5 mg IM. Group 3: Lorazepam 4 mg and haloperidol 5 mg IM. Sample size for each group: Group 1 – 31 Group 2 – 35 Group 3 – 32.	Hourly for 24 hours	Agitated Behaviour Scale. 11 items of modified BPRS CGI. All drugs gave a significant reduction in Agitated Behaviour Scale and modified BPRS over time. More rapid onset of action for group 3 (compared to group 2 p=0.64) as contrasted with groups 1 and group 3 (p=0.0014). Greater reduction in MBPRS at 2 and 3 hours for group 3. No difference at any time points for CGI. Means adjusted by analysis of covariance statistical text for baseline levels. Time spent asleep: • Hourly assessment of whether participant was awake or could be aroused by verbal stimuli was made using an alertness scale (for a minimum of 12 hours after last injection). Significantly more time was spent asleep in groups 1 and 3 than in groups 2 at 3, 4, 5, 6, 7, 9 and 11 hours. Number of doses required for tranquillisation. Adverse reactions. No difference found between the number of incidences. More extrapyramidal symptoms (EPS) in group 2 (20%), than group 1 or 3.

	obstruction, severe hypo- or hyper - tension, glaucoma, benzodiazepine or neuroleptic within last 24-hours.				

Notes on quality assessment and comments:

- No objective measure of behaviour on entry into study.
- Many comparisons performed with no adjustment to p value.
- Considered sleep a therapeutic end-point.
- If sleep was considered as a therapeutic end-point for rapid tranquillisation, then combined treatment or lorazepam alone was superior to haloperidol alone.

1.4.2 Bieniek 1998

Study	Population	Methods	Main intervention(s) and comparisons	Follow-up period	Outcomes (primary, secondary and adverse events) effect size, p-value
Bieniek and colleagues (1998) Country: US Source of funding: not stated in Broadstock. Evidence level: 1-	20 acutely agitated newly admitted service users – at least 4 on Overt Aggression Scale (OAS). Exclusions not mentioned.	Allocation: randomised. Blindness: double-blind. Duration: 24 hours. Setting: psychiatric emergency services. Baseline comparability: yes.	Group 1: Lorazepam 2 mg IM. Group 2: Haloperidol 5 mg IM plus lorazepam 2 mg IM. Sample size for each group: Group 1 not stated in Broadstock. Group 2 not stated in Broadstock.	30, 60, 120, 180 minutes after first injection.	Both groups significant reduction at 60 min OAS, (75%) visual analogue scale VAS (50%), CGI (45%). No differences were noted with ANOVAS, but non-parametric tests indicated that a greater percentage improved post 60 minutes in combined group OAS, (100%) VAS (78%) whilst in group 1 OAS, (55%) VAS (27%). No difference on CGI. Sedation by VAS – no differences in time. No serious adverse events occurred.

Notes on quality assessment and comments:

- Small sample size.
- Short follow-up
- Many comparisons performed with no adjustment to p value.
- 2 service users received second injection in group 1 but not excluded, which disadvantages group 2.

1.4.3 Dorevitch 1999

Study	Population	Methods	Main intervention(s) and comparisons	Follow-up period	Outcomes (primary, secondary and adverse events) effect size, p-value
Dorevitch (1999) Country: Israel Source of funding: not stated. Evidence level: 1-	Presence of active psychosis, disruptive or aggressive behaviour, pronounced psychomotor agitation, or violent outburst and hospitalisation in an acute ward. Exclusions: not mentioned.	Allocation: randomised. Blindness: double-blind. Duration: 120 minutes. Setting: acute ward. Baseline comparability: only age and gender stated.	Group 1: Haloperidol 5 mg IM. Group 2: Flunitrazepam 1 mg IM. Sample size for each group: Group 1 – 13 Group 2 – 15. Numbers needed to treat: 8 (8.125).	15, 30, 45, 60, 90, 120 minutes after first injection.	Overt aggression scale (OAS)=50% reduction at 90 minutes postadministration – both groups significant (group 1=95%, group 2=80%) p <0.001. Effect of haloperidol lasted at least 120 minutes postadministration. Effect of flumitrazepam had worn off at 60 minutes. No significant difference in anti-aggressive response at 90 minutes. Group 2 reached maximum aggressive effect quicker <30 minutes). Overall response rate (defined as a reduction of a least 50% in overt aggression scale score at 90 minutes for both drugs – p <0.001). Adverse reactions: No EPS in either group. 3 in each group had marked sedation.

Notes on quality assessment and comments:

- Small sample size.
- Short follow-up.
- No objective measure of behaviour on entry into study.
- Concluded that flumitrazepam is convenient, rapid and safe.

1.4.4 Foster 1997

Study	Population	Methods	Main intervention(s) and comparisons	Follow-up period	Outcomes (primary, secondary and adverse events) effect size, p-value
Foster and colleagues (1997) Country: US Source of funding: part supported by a grant from the National Alliance for Research on Schizophrenia and Depression. Evidence level: 1-	37 service users with psychotic symptoms.	Allocation: not stated. Blindness: double-blind. Duration: every 30 minutes for 4 hours (until participant sedated or no longer a danger to themselves or others). Setting: emergency department. Baseline comparability: yes.	Group 1: Haloperidol 5 mg IM or oral concentrate. Group 2: Lorazepam 2 mg IM or oral concentrate. BPRS GCI Sample size for each group: Group 1 – 20 Group 2 – 17.	4 hours.	Aggression reduction (better GCI scores at 1, 2 and 3 hours in group 2). Both groups has significant decrease in BPRS scores ($p < 0.001$) and GCI scores ($p < 0.001$). No significant difference between oral and IM routes. Adverse reactions (none recorded) Sedation/sleep (2 service users group 1, 3 service users group 2). Physiological measures (blood pressure and so on).

Notes on quality assessment and comments:

- Clinical characteristics not well balanced in 2 groups (groups differences for diagnosis significant ($p < 0.05$), more bipolar service users received lorazepam and more psychotic service users received haloperidol by chance.
- Intoxicants weren't tested for.
- Doesn't state if allocation is sufficiently concealed.
- Small study.
- Very short time period.
- Authors conclude that Lorazepam may be safer, but this needs to be treated as tentative, at best.

1.4.5 Fruensgaard 1977

Study	Population	Methods	Main intervention(s) and comparisons	Follow-up period	Outcomes (primary, secondary and adverse events) effect size, p-value
Fruensgaard and colleagues (1977) Country: Denmark Source of funding: statistical evaluation by Fl. Abildgaard and drugs supplied by Lederle Laboratories, a division of American Cyanamide Corporations. Evidence level: 1-	Service users with acute psychosis characterised by agitation, excitement, aggressiveness, hostility, delusions and hallucinations. Excluded: pregnancy, manic-depressive illness, ECT in preceding 8 weeks, organic brain syndrome with marked dementia, convulsive disorders, alcoholism or drug dependence, serious impairment of renal, hepatic, cardiovascular or metabolic functions, and present or former increase intro-ocular pressure, no neuroleptic therapy within 12 hours preceding admission.	Allocation: randomised. Blindness: double. Duration: 3 days. Setting: multi-site. Baseline comparability: yes.	Group 1: Loxapine 50 mg, IM (maximum 150 mg injections in 24 hours). Group 2: Haloperidol 5 mg IM (maximum 15 mg injections in 24 hours). BPRS GCI Daily at 6-12 hours after last dose. Blood pressure and pulse rates measured at baseline and during treatment (specific intervals). Laboratory data included complete blood count, serum creatinine, urinalysis, electrocardiogram	No follow-up beyond 3 days reported in this study (up to 3 days of IM treatment, followed by oral treatment up to 4 weeks).	Aggression: No significant differences in effect of 2 drugs on BPRS or CGI. Sedation: More pronounced in loxapine group p <0.025 (2 hours hrs after first injection p <0.05). After loxapine, there was a higher sleeping period regardless of injection time, diagnosis or hospital (p <0.01). Adverse reactions (evaluated at least daily or as necessary): 7/15 in group 2 and 1/15 in group 1 experienced EPS. (Acute dystonia was recorded in 2 of these cases in group 2). Anticholingeric 5/15 group 1 and 3/15 group 2. Drowsiness/fatigue (where seen as problem by service user) 4/15 group 1 and 3/15 group 2. Dizziness 6/15 group 1 and 1/15 group 2. Palpitations 1/15 group 1. Injection site pain lasting for less than 1 hour 3/15 group 1 and 2/15 group 2 (a moderate reaction of the tissue could be noted). Decreased pulse rate and systolic and diastolic blood pressure during treatment – tendency in both groups. No subjective symptoms were noted. Systolic blood pressure didn't fall below 100 mmHg for any service user. Other drugs taken: Biperiden 1 ml.

			and, in some service users, liver parameters, both before, during (specific intervals) and if necessary after trial. Sample size for each group: Group 1 – 15 Group 2 – 15.		

Notes on quality assessment and comments:

- The numbers in each group are equal, which suggests that this trial is not properly randomised. Method of randomisation is not specified.
- The study has a small sample size, which makes comparisons between the 2 drugs difficult.
- The authors stress that further trials that compare loxapine and haloperidol are necessary.

1.4.6 Garza-Trevino 1989

Study	Population	Methods	Main intervention(s) and comparisons	Follow-up period	Outcomes (primary, secondary and adverse events) effect size, p-value
Garza-Trevino and colleagues (1989) Country: US Source of funding: not stated. Evidence level: 1-	68 service users (study 1); 53 service users (study 2) judged to require immediate treatment for acute agitation - scoring between 50 and 100 on a VAS. Exclusion criteria: no service user had received a dose of centrally acting depressant at least 2 hours before baseline.	Allocation: randomised. Blindness: open. Duration: not mentioned. Setting: general psychiatric hospital. Baseline comparability: yes.	Study 1: Group 1 Lorazepam 4 mg IM. Group 2: Haloperidol 5 mg IM. Group 3: Both of the above. Sample size for each group: Group 1 (not stated by Broadstock) Group 2 (not stated by Broadstock) Group 3 (not stated by Broadstock). Study 2: Group 1: Haloperidol 5mg IM and phenobarbital	30, 60, >60 minutes (usually within 3.5 minutes after first administration).	Study 1: Combination treatment was more likely to lead to tranquillisation than either of the single drugs within 30 minutes 18/24=75% versus 16/44=36% Chi-squared. Finding replicated in ANOVAS. Adverse reactions: Not reported. Study 2: 3 participants in group 1 and 1 in group 2 failed to reach tranquillisation after third dose. Adverse reactions: Not reported.

			sodium (IM) 130 mg Group 2: Thiothixene 5 mg (IM) and lorazepam 4 mg IM. Sample size for each group: Group 1 (not stated) Group 2 (not stated).		
Notes on quality assessment and comments: In study 1, more women were in the haloperidol only group than the combined group. Very short follow-up period for both studies. Side-effects not described for both studies. Neither study was double-blind.					

1.4.7 Paprocki 1977

Study	Population	Methods	Main intervention(s) and comparisons	Follow-up period	Outcomes (primary, secondary and adverse events) effect size, p-value
Paprocki & Versiani (1977) Country: Brazil Source of funding: supported by a grant from Lederies Laboratories, a division of American Cyanamid Company. Evidence level: 1-	35 female service users with psychotic symptoms characterised by agitation, excitement, aggressiveness, hostility, delusions and hallucinations. Excluded: known hypersensitivity dibenzazepine compounds; ECT, insulin coma, or subcoma therapy within previous 8 weeks, organic brain syndrome with marked dementia or inability to communicate during interview, history of convulsive disorders, alcoholism or drug dependence as a significant feature of clinical history, serious impairment of renal or hepatic function, increased intra-ocular pressure or history of narrow angle glaucoma or urinary retention, cardiovascular or metabolic disorder,	Allocation: randomised. Blindness: double. Duration: every 30 minutes for 4 hours (until participant sedated or no longer a danger to themselves or others). Setting: fourth ward of state hospital. Baseline comparability: yes.	Group 1 Haloperidol 5 mg IM or oral concentrate for 4 days (in 1 ml ampules) at 6-12 hour intervals (or until symptoms diminished) then oral equivalent for 3 days and then 2.5 mg doses for 4 weeks (adjusted to suit service user response). Group 2 Loxapine 50 mg IM or oral concentrate for 4 days (in 1 ml ampules) at 6-12 hour intervals (or until symptoms diminished) then oral equivalent for 3 days and then 25 mg doses for 4 weeks (adjusted to suit service user response). The initial IM dose was either 0.5 or 1 ml (no more than 3 ml in 24 hours). Oral phase maximum dose was	4 weeks. BPRS Nurses' Observation Scale for Inpatient Evaluation (NOSIE) CGI At 24-hour intervals and then weekly during oral phase.	Loss to follow-up: 25 service users had sufficient response to enter oral phase (group 1: 14, group 2: 11) 22 reached end of 4 weeks. All dropouts were for inadequate response (except 1 in haloperidol for toxicity). Aggression reduction: No significant differences between the 2 drugs were detected on any of the rating scales. Both drugs showed significant improvement on most items and total scores, except in motor retardation on BPRS and NOSIE which worsen from day 2–end of trial (haloperidol) and from day 2–5 (loxapine). Adverse reactions: 1 toxicity withdrawal in group 1, rigidity and drowsiness were noted in each group. Anti-Parkinson medication (trihexyphenidyl – 4 mg/day) group 1–6 IM phase, group 2–2 IM phase. Sedation/sleep (loxapine groups significantly less somatic effect p=0.05 at day 4). Sedative effected peaked at 6 hours for loxapine and 8 hours for haloperidol. Sleep was not considered an undesirable outcome. On day 1, only 6 loxapine and 11 haloperidol service users were awake prior to their pm injection. Physiological measures significant alterations in several parameters relative to vital signs – mean

	pregnancy suspected or confirmed (urine test).		either 150mg loxapine or 15mg haloperidol. Laboratory tests of haematology, blood chemistry and urinalysis at baseline, during parenteral phase and at end of oral phase. Sample size for each group Group 1 – 18 – 14 in oral phase Group 2 – 17 – 11 in oral phase.		lying pulse (5.0 beats per minute), means lying and systolic blood pressure reduced (5.9 and 6.9 mm Hg), no significant changes in diastolic blood pressure. No significant difference between 2 groups.
Notes on quality assessment and comments: Clinical characteristics were well balanced in 2 groups Small study. Authors note the need to take possible hypertension into account when using IM neuroleptics. Authors conclude that loxapine is superior to haloperidol in controlling agitation/excitement and aggressiveness as assessed under the conditions of this trial. However, this difference was only noted over a period of 5 days, and was not significant in the first 24 hours, and is therefore not relevant to rapid tranquillisation.					

1.4.8 Reschke 1974

Study	Population	Methods	Main intervention(s) and comparisons	Follow-up period	Outcomes (primary, secondary and adverse events) effect size, p-value
Reschke (1974) Country: US Source of funding: Not specified Evidence level: 1-	48 female and 2 male psychiatric emergencies. Excluded: pregnant women, acute or chronic brain syndrome, acute alcoholic intoxication, epilepsy, psychoneurosis, drug addiction, epilepsy, psychoneurosis, personality disorder.	Allocation: randomised. Blindness: double-blind. Duration: 24 hours. Setting: ward. Baseline comparability: groups 4 and 5 each contained 1 male service user.	Group 1: Haloperidol 5 mg IM. Group 2: Haloperidol 2 mg IM. Group 3: Haloperidol 1 mg IM. Group 4: Chlorpromazine 25 mg IM. Group 5: Placebo. blood pressure, pulse, respiration at baseline after each injection at each target symptom. Laboratory data blood, liver and urine profiles, chest X-ray and ECG at baseline and end of study. Sample size for each group: Group 1 – 10 Group 2 – 11 Group 3 – 8	24 hours or 6 hours after last dose – whichever was later. 5-point target symptoms rating scale (0=absent, 4=very severe) at baseline at every 30 minutes for 2 hours after first injection. BPRS at baseline and immediately after first injection. Global therapeutic effect (at IM and oral stages).	Aggression Symptoms adequately controlled in significantly more service users in groups 1 and 2 ($p < 0.05$). In group 1, 2.8 injections were required for adequate control and in group 2, 3.7 injections were required for adequate control. Loss to follow-up: 1 in group 1 due to transient hypotensive episode. In group 5, 6 transferred to oral medication. Somnolence (not evaluated at 2-hour evaluation point): 1 in group 2, 5 in group 4. Adverse reactions: • Transient hypertension – haloperidol 3, chlorpromazine 1, placebo 0. • Drowsiness – awake – haloperidol 12, chlorpromazine 1, placebo 0. • Drowsiness – asleep – haloperidol 1, chlorpromazine 6, placebo 0. • Dry mouth – haloperidol 4, chlorpromazine 1, placebo 0. • Mild EPS – haloperidol 6, chlorpromazine 1, placebo 0. Other drugs taken Trihexyphenidyl HC1 2 mg for EPS.

			Group 4 – 10 Group 5 – 11.		
Notes on quality assessment and comments: Sample size was small. Subsequent treatment with oral haloperidol versus oral chlorpromazine favoured haloperidol, but results are not reported in sufficient detail. Chlorpromazine is not recommended for rapid tranquillisation as it is hazardous in the doses required for this procedure.					

1.4.9 Tuason 1986

Study	Population	Methods	Main intervention(s) and comparisons	Follow-up period	Outcomes (primary, secondary and adverse events) effect size, p-value
Reschke (1974) Country: US Source of funding: Not specified Evidence level: 1-	48 female and 2 male psychiatric emergencies. Excluded: pregnant women, acute or? chronic brain syndrome, acute alcoholic intoxication, epilepsy, psychoneurosis, drug addiction, epilepsy, psychoneurosis, personality disorder.	Allocation: randomised. Blindness: double-blind. Duration: 24 hours. Setting: ward. Baseline comparability: groups 4 and5 each contained 1 male service user.	Group 1: Haloperidol 5 mg IM. Group 2: Haloperidol 2 mg IM. Group 3: Haloperidol 1MG IM. Group 4: Chlorpromazine 25 mg IM. Group 5: Placebo. Blood pressure, pulse, respiration at baseline after each injection at each target symptom. Laboratory data blood, liver and urine profiles, chest X-ray and	10 days (IM for 24-72 hours and then orally up to 10 days).	Response rate. Hostility, uncooperativeness. Sedation – considered as therapeutic end-point and noted in the first hour for most participants. Within 12 hours, 24/25 group 1 and 22/27 group 2 were asleep. Therapeutic response did not differ significantly between the 2 treatment groups (p >0.05). Adverse reactions: Dystonia (14), akathisia (14); 4 removed from study due to adverse reactions (2 in groups 1 (increased blood pressure, tachycardia), 2 in group 2 (severe akathisia and severe dystonia). No significant difference between the 2 groups in the number and severity of adverse events.

			ECG at baseline and end of study. Sample size for each group: Group 1 – 10 Group 2 – 11 Group 3 – 8 Group 4 – 10 Group 5 – 11.		
Notes on quality assessment and comments: Analysis of dropouts mentioned. Drug administration not blinded, but evaluation of effects blinded. Medical history of service users not known/reported.					

2 ECONOMICS EVIDENCE – COMPLETED METHODOLOGY CHECKLISTS

2.1 MODIFICATIONS TO THE ENVIRONMENT

2.1.1 Nanda 2011

Study identification: Nanda U, Eisen S, Zadeh RS, Owen D. Effect of visual art on patient anxiety and agitation in a mental health facility and implications for the business case. Journal of Psychiatric and Mental Health Nursing. 2011;18:386-93.			
Guideline topic: Violence and aggression			
Section 1: Applicability (relevance to specific guideline review question(s) and the NICE reference case)		**Yes/ Partly/ No/Unclear /NA**	**Comments**
1.1	Is the study population appropriate for the guideline?	Yes	
1.2	Are the interventions and services appropriate for the guideline?	Yes	
1.3	Is the healthcare system in which the study was conducted sufficiently similar to the current UK NHS context?	Partly	US
1.4	Are costs measured from the NHS and personal social services (PSS) perspective?	No	
1.5	Are non-direct health effects on individuals excluded?	Yes	
1.6	Are both costs and health effects discounted at an annual rate of 3.5%?	NA	1 year
1.7	Is the value of health effects expressed in terms of quality-adjusted life years (QALYs)?	No	
1.8	Are changes in health-related quality of life (HRQoL) reported directly from patients and/or carers?	NA	
1.9	Is the valuation of changes in HRQoL (utilities) obtained from a representative sample of the general public?	NA	
1.10	**Overall judgement:** Partially applicable		
Other comments: None			

Section 2: Study limitations (the level of methodological quality)		**Yes/ Partly/ No/Unclear/ NA**	**Comments**
2.1	Does the model structure adequately reflect the nature of the health condition under evaluation?	NA	
2.2	Is the time horizon sufficiently long to reflect all important differences in costs and outcomes?	Yes	
2.3	Are all important and relevant health outcomes included?	No	
2.4	Are the estimates of baseline health outcomes from the best available source?	No	Observational study
2.5	Are the estimates of relative treatment effects from the best available source?	No	Observational study
2.6	Are all important and relevant costs included?	Yes	
2.7	Are the estimates of resource use from the best available source?	No	Observational study
2.8	Are the unit costs of resources from the best available source?	No	Local sources
2.9	Is an appropriate incremental analysis presented or can it be calculated from the data?	Yes	Only cost minimisation
2.10	Are all important parameters whose values are uncertain subjected to appropriate sensitivity analysis?	No	
2.11	Is there no potential conflict of interest?	Yes	
2.12	**Overall assessment:** potentially serious limitations		
Other comments: None			

2.2 RAPID TRANQUILLISATION / PHARM

2.2.1 Freeman 2009

Study identification: Freeman DJ, DiPaula BA, Love RC. Intramuscular haloperidol versus intramuscular olanzapine for treatment of acute agitation: A cost-minimization study. Pharmacotherapy. 2009;29:930-6.			
Guideline topic: Violence and aggression			
Section 1: Applicability (relevance to specific guideline review question(s) and the NICE reference case)		**Yes/ Partly/ No/Unclear /NA**	**Comments**
1.1	Is the study population appropriate for the guideline?	Yes	
1.2	Are the interventions and services appropriate for the guideline?	Yes	
1.3	Is the healthcare system in which the study was conducted sufficiently similar to the current UK NHS context?	Partly	US
1.4	Are costs measured from the NHS and personal social services (PSS) perspective?	No	
1.5	Are non-direct health effects on individuals excluded?	Yes	
1.6	Are both costs and health effects discounted at an annual rate of 3.5%?	NA	Episode based approach
1.7	Is the value of health effects expressed in terms of quality-adjusted life years (QALYs)?	No	
1.8	Are changes in health-related quality of life (HRQoL) reported directly from patients and/or carers?	NA	
1.9	Is the valuation of changes in HRQoL (utilities) obtained from a representative sample of the general public?	NA	
1.10	**Overall judgement:** Partially applicable		
Other comments: None			

Section 2: Study limitations (the level of methodological quality)		**Yes/ Partly/ No/Unclear/ NA**	**Comments**
2.1	Does the model structure adequately reflect the nature of the health condition under evaluation?	NA	
2.2	Is the time horizon sufficiently long to reflect all important differences in costs and outcomes?	Yes	
2.3	Are all important and relevant health outcomes included?	No	
2.4	Are the estimates of baseline health outcomes from the best available source?	No	Retrospective medical record review
2.5	Are the estimates of relative treatment effects from the best available source?	No	Retrospective medical record review
2.6	Are all important and relevant costs included?	No	
2.7	Are the estimates of resource use from the best available source?	No	Retrospective medical record review
2.8	Are the unit costs of resources from the best available source?	No	Local sources
2.9	Is an appropriate incremental analysis presented or can it be calculated from the data?	No	Cost minimisation
2.10	Are all important parameters whose values are uncertain subjected to appropriate sensitivity analysis?	No	None
2.11	Is there no potential conflict of interest?	Yes	
2.12	**Overall assessment:** potentially serious limitations		
Other comments: None			

2.3 CHILDREN AND YOUNG PEOPLE

2.3.1 LeBel 2005

Study identification: LeBel J, Goldstein R. The economic cost of using restraint and the value added by restraint reduction or elimination. Psychiatric services. 2005;56:1109-1114.			
Guideline topic: Violence and aggression			
Section 1: Applicability (relevance to specific guideline review question(s) and the NICE reference case)		**Yes/ Partly/ No/Unclear /NA**	**Comments**
1.1	Is the study population appropriate for the guideline?	Yes	
1.2	Are the interventions and services appropriate for the guideline?	Partly	No comparator
1.3	Is the healthcare system in which the study was conducted sufficiently similar to the current UK NHS context?	Partly	US
1.4	Are costs measured from the NHS and personal social services (PSS) perspective?	No	
1.5	Are non-direct health effects on individuals excluded?	Yes	
1.6	Are both costs and health effects discounted at an annual rate of 3.5%?	NA	Episode based approach
1.7	Is the value of health effects expressed in terms of quality-adjusted life years (QALYs)?	No	
1.8	Are changes in health-related quality of life (HRQoL) reported directly from patients and/or carers?	NA	
1.9	Is the valuation of changes in HRQoL (utilities) obtained from a representative sample of the general public?	NA	
1.10	**Overall judgement:** Partially applicable		
Other comments: None			

Section 2: Study limitations (the level of methodological quality)		**Yes/ Partly/ No/Unclear/ NA**	**Comments**
2.1	Does the model structure adequately reflect the nature of the health condition under evaluation?	NA	
2.2	Is the time horizon sufficiently long to reflect all important differences in costs and outcomes?	No	Long term effects may exist
2.3	Are all important and relevant health outcomes included?	No	
2.4	Are the estimates of baseline health outcomes from the best available source?	NA	
2.5	Are the estimates of relative treatment effects from the best available source?	NA	
2.6	Are all important and relevant costs included?	No	
2.7	Are the estimates of resource use from the best available source?	No	Retrospective medical record review
2.8	Are the unit costs of resources from the best available source?	No	Local sources
2.9	Is an appropriate incremental analysis presented or can it be calculated from the data?	No	Cost minimisation
2.10	Are all important parameters whose values are uncertain subjected to appropriate sensitivity analysis?	NA	
2.11	Is there no potential conflict of interest?	No	1 author from consultancy
2.12	**Overall assessment**: potentially serious limitations		
Other comments: None			