ARTICLE

Interventions for preventing and reducing the use of physical restraints in long-term geriatric care – a Cochrane review

Ralph Möhler, Tanja Richter, Sascha Köpke and Gabriele Meyer

Aims and objectives. To evaluate the effectiveness of interventions to prevent and reduce the use of physical restraints in older people requiring long-term nursing care.

Background. Physical restraints are commonly used in geriatric long-term care. However, they are associated with adverse outcomes. Therefore, freedom from physical restraints should be the aim of high-quality nursing care.

Design. Systematic review of randomised controlled trials.

Methods. This review followed the methods of the Cochrane Handbook of systematic reviews of interventions. The systematic search (September 2009) covered the Cochrane Dementia and Cognitive Improvement Group’s Specialized Register, MEDLINE, EMBASE, CINAHL, PsycINFO and LILACS.

Results. Six cluster-randomised controlled trials met the inclusion criteria. All studies investigated educational approaches targeting nursing staff. In addition, two studies offered consultation, two guidance and one support and free access to technical aids. Five studies examined nursing home residents and one study residents in group dwelling units. No studies in community settings were included. Overall, the methodological quality of studies was low. Their results were inconsistent. One study with good methodological quality in the nursing home setting documented an increase in physical restraints use in both groups, while the other four studies with lower quality found reduced use of physical restraints in the intervention group. The single study in group dwelling units found no change in physical restraints use in the intervention group but a significant increase in the control group.

Conclusions. There is insufficient evidence supporting the effectiveness of educational interventions targeting nursing staff for preventing or reducing the use of physical restraints in geriatric long-term care.

Relevance to clinical practice. Our findings indicate that educational programmes targeting nursing staff might not be effective in reducing the use of physical restraints in geriatric long-term care. It remains unclear which components should be included in educational programmes aiming to reduce physical restraints.

Key words: long-term care, older people, physical restraints, systematic review

Accepted for publication: 26 February 2012

Introduction

Physical restraints (PR) are commonly used in geriatric long-term care in different countries, as shown in several international studies (Hamers et al. 2004, Feng et al. 2009, Meyer et al. 2009a). Epidemiological studies revealed wide centre variation between and within countries (Feng et al. 2009, Meyer et al. 2009a). For community-dwelling older...
people, there is little evidence about the amount and types of PR (De Veer et al. 2009).

For the purpose of this review, PR are defined as ‘any device, material or equipment attached to or near a person’ s body and which cannot be controlled or easily removed by the person and which deliberately prevents or is deliberately intended to prevent a person’s free body movement to a position of choice and/or a person’s normal access to their body’ (Evans et al. 2002, Retsas 1998). This includes bilateral bedrails, limb or trunk belts, and fixed tables on a chair or chairs that prevent persons from getting up (Evans et al. 2002).

Nurses claim to use PR predominantly for patient safety, primarily for fall prevention (Evans et al. 2002, Hamers & Huizing 2005). Other reasons mentioned include control of challenging behaviour and safe use of medical devices (Hamers & Huizing 2005). However, there is strong evidence that PR use is not an adequate measure for reducing falls or fall-related injuries or for controlling challenging behaviour (Evans et al. 2002, Pellfolk et al. 2010). The use of PR is also related to several adverse effects, for example, direct injuries and mortality related to PR use through falls or fatal entrapment, decreased mobility or reduced psychological well-being (Castle & Engberg 2009, Engberg et al. 2008, Evans et al. 2002, Healey et al. 2008). Associations between PR use and cognitive impairment and disruptive behaviour have been shown (Huizing et al. 2007, Meyer et al. 2009a, Sullivan-Marx et al. 1999). It is unclear whether institutional characteristics, such as staff mix, significantly influence decisions on PR use (Huizing et al. 2007, Meyer et al. 2009a). Also, the ‘philosophy’ or ‘culture’ of care (i.e. attitudes and beliefs of nursing staff) is suspected to have a strong influence on the use of PR (Hamers & Huizing 2005, Meyer et al. 2009a).

A restraint-free nursing care environment has been recommended as the standard of care (Flaherty 2004). Accordingly, efforts have been made to reduce PR. Reduction approaches were first introduced in the United States in the 1980s (Castle & Mor 1998) with a number of studies conducted in hospitals and nursing homes. A systematic review (Evans & Fitzgerald 2002) analysed 13 studies, including only one randomised controlled trial (RCT). Since then, a number of randomised controlled studies have been conducted (Huizing et al. 2009a, Testad et al. 2005, 2010). All interventions were designed as complex interventions consisting of different components, including educational sessions aimed at changing nurses’ attitudes to PR use and information about and implementation of alternatives to the use of PR. Until our recently published Cochrane Review (Mohler et al. 2011), there was no systematic overview available evaluating the efficacy and safety of interventions aiming to reduce or prevent PR use.

Aims and methods

This systematic review aims to evaluate the effectiveness of interventions for preventing and reducing the use of PR in older people who require long-term nursing care (either in community nursing care or in residential care facilities) and describes the included complex interventions. The methods of the Cochrane Collaboration (Higgins & Greene 2011) were used. A review protocol has been published (Meyer et al. 2009b).

Inclusion and exclusion criteria

- **Types of studies**: Individual or cluster-randomised controlled trials without restrictions on publication date or language.
- **Types of participants**: Older people of either gender requiring long-term nursing care.
- **Types of interventions**: Interventions or groups of interventions evaluating a PR reduction or prevention programme.
- **Types of outcomes**: Primary outcomes were the number or proportion of residents with at least one PR, prevention of PR or the reduction of PR (i.e. withdrawing previously used PR). Secondary outcomes were types of PR, duration of PR use, prescription of psychotropic drugs, residents’ and caregivers’ quality of life, adverse effects of the interventions employed, duration of the interventions’ effect, and injuries and deaths during the study period.

Literature search

The literature search was performed in September 2009 and covered the following databases: The Cochrane Dementia and Cognitive Improvement Group’s Specialized Register, MEDLINE, EMBASE, CINAHL, PsycINFO, LILACS and a number of trial registers. In addition, reference lists of included publications were checked for additional trials. Authors of the included studies were contacted to identify unpublished or ongoing studies. Hand-search was performed in abstract of the books of the following scientific congresses: IAGG World Congress of Gerontology and Geriatrics (2001, 2005, 2009), The Gerontological Society of America’s Annual Scientific Meeting (2000–2009), Congress of the European Union Geriatric Medicine Society (2000–2008) and European Congress of Gerontology (2003, 2007). The following search terms were used: physical restraint*, bedrail*, bedchair*,
containment measure, elderly, old people, geriatric, aged, nursing home, care home, geriatric care, residential facil. The complete search strategy for Medline is published in the original Cochrane Review (Möhler et al. 2011).

Data collection and extraction
Two authors (RM, TR) evaluated eligibility and methodological quality of studies retrieved from the literature search. In cases of disagreement, a third author (GM) additionally reviewed the studies, and agreement was reached by consensus. RM and TR independently extracted and compared the data.

Quality assessment
Quality assessment followed the Cochrane Handbook for Systematic Reviews of Interventions (Higgins & Greene 2011). Two authors (RM, TR) independently assessed the methodological quality of studies to identify any potential sources of systematic bias. Criteria for appraisal of studies were internal validity and low risk of bias through selection bias, performance bias, attrition bias and detection bias. For the included cluster-randomised trials, additional design-related criteria were included in the quality assessment (loss of cluster bias, appropriateness of statistical methods for cluster randomisation). In case of unclear or missing information, the corresponding author of the trial was contacted.

Data synthesis
A data check revealed pronounced clinical heterogeneity of the primary studies with regard to PR definitions and interventions’ composition, for example, duration and frequency of educational sessions. Two studies showed pronounced baseline differences in the prevalence of PR between groups. In one study, the results were published as the mean number of PR per week. Given these inconsistencies, a meta-analysis on published data was not feasible.

Results
The literature search revealed 160 citations, 27 publications were screened in full text, and four publications fulfilled the eligibility criteria (Evans et al. 1997, Testad et al. 2005, Huizing et al. 2009a,b). Studies were excluded because they were not randomised controlled trials (n = 15) or did not meet the inclusion criteria regarding primary outcomes (n = 1), participants (n = 2) or intervention (n = 5). Two of the included publications reported on the same study (Huizing et al. 2009a,b); thus, the final report was used as the primary source of data Huizing et al. (2009a). Two further studies were identified through other sources (Pellfolk et al. 2010, Testad et al. 2010). One unpublished study (Koczy et al. 2005) and two ongoing studies were identified (Haut et al. 2009, Gulpers et al. 2010). The final Cochrane review analysed five studies. Two studies were published after the review’s publication. The study by Gulpers et al. (2011) does not fulfil the inclusion criteria (non-randomised controlled trial). The study by Koczy et al. (2011) has been included. Thus, our databases comprise six studies (Fig. 1).

Characteristics of included studies
Five of six studies (Evans et al. 1997, Testad et al. 2005, 2010, Huizing et al. 2009a, Koczy et al. 2011) were conducted in nursing homes and one study (Pellfolk et al. 2010) in group dwelling units for persons with dementia. No studies were identified that investigated older people in the community. In all studies, cluster randomisation was used. A cluster was defined as a nursing home on its own (Evans et al. 1997, Testad et al. 2005, 2010, Koczy et al. 2011) or a nursing home ward (Huizing et al. 2009a) or a group dwelling unit (Pellfolk et al. 2010). Clusters were randomised to either an intervention group (IG) or a control group (CG). In three studies (Evans et al. 1997, Testad et al. 2005, 2010), study groups consisted of only one or two clusters. Follow-up ranged from 3 to 12 months (Table 1).

Risk of bias of included studies
Overall, the methodological quality of the studies was low to moderate (Fig. 2). Sequence generation was adequate in five studies (Evans et al. 1997, Testad et al. 2005, Huizing et al. 2009a, Pellfolk et al. 2010, Koczy et al. 2011) and unclear in one (Testad et al. 2010). Allocation concealment was adequate in only one study (Huizing et al. 2009a). A recruitment bias could have occurred in three studies, which identified participants after random allocation of clusters (Testad et al. 2005, 2010, Pellfolk et al. 2010). Baseline imbalances between groups were found in four studies. Evans et al. (1997) found statistically significant differences concerning the prevalence of PR and the dependency level of participants. Testad et al. (2010) reported statistically significant differences in the prevalence of PR, in psychotropic medication use and in mean agitation scores. In the study by Koczy et al. (2011), the groups significantly differed at baseline in terms of gender distribution and level of care dependency. The baseline PR rate of all residents living in the participating nursing homes (n = 5561) differed noticeably between groups (IG 7.2% vs. CG 5.0%). It is not reported whether this
difference is statistically significant. Pellfolk et al. (2010) documented significant baseline differences in the age of participants and in wandering behaviour, but it remains unclear whether these differences constitute a clinically relevant imbalance between groups. Outcome assessors were blinded to group allocation in four studies (Evans et al. 1997, Testad et al. 2005, 2010, Huizing et al. 2009a). No clusters were lost to follow-up in five studies (Evans et al. 1997, Testad et al. 2005, 2010, Pellfolk et al. 2010, Koczy et al. 2011), and one cluster of 15 clusters dropped out in the study by Huizing et al. (2009a).

Five studies did not consider the cluster design effect in any of their statistical analyses (Evans et al. 1997, Testad et al. 2005, 2010, Huizing et al. 2009a, Koczy et al. 2011). Only one study (Pellfolk et al. 2010) adjusted the likelihood of being restrained to the cluster design. Intra-cluster correlation coefficients were not reported in any of the studies.

For the study by Koczy et al. (2011), an outcome reporting bias could be verified. In the study protocol (Koczy et al. 2005), the main outcomes were defined as the ‘number of persons with PR’ and ‘the number of hours with PR’, while the published results refer to ‘complete cessation of PR use at the days 91–93 after start of the intervention’ in residents with PR at baseline as main outcome. Also differing from the study protocol, the published results only considered belts and fixed tables as PR, whereas the protocol also included other measures, including braked wheelchairs near a table and ‘protective sheets’ (Meyer & Köpke 2011).
Table 1 Characteristics of included studies

<table>
<thead>
<tr>
<th>Reference</th>
<th>Design</th>
<th>Sample</th>
<th>Setting</th>
<th>Inclusion/exclusion criteria</th>
<th>Follow-up, months</th>
<th>Intervention and Control</th>
<th>Primary Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evans et al. (1997)</td>
<td>Cluster-RCT</td>
<td>n = 463</td>
<td>3 nursing homes, USA</td>
<td>All residents of included nursing homes, age 60+ years, non-comatose and conversant in English</td>
<td>12</td>
<td>Educational intervention plus consultation (IG1); Educational intervention only (IG2); Usual care (control)</td>
<td>Physical restraints status</td>
</tr>
<tr>
<td>Huizing et al. (2009a)</td>
<td>Cluster-RCT</td>
<td>n = 241</td>
<td>14 nursing home wards from 7 nursing homes, The Netherlands</td>
<td>All residents of each participating nursing home ward, no Korsakoff's syndrome, no psychiatric diseases</td>
<td>10</td>
<td>Educational intervention plus consultation; Usual care (control)</td>
<td>Physical restraint status</td>
</tr>
<tr>
<td>Koczy et al. (2011)</td>
<td>Cluster-RCT</td>
<td>n = 333</td>
<td>45 nursing homes, Germany</td>
<td>Nursing homes with at least five residents with PR use; all residents with PR use in a 3-day period before the start of the intervention</td>
<td>3</td>
<td>Educational intervention plus support and provision of technical aids; Usual care (control)</td>
<td>Complete cessation of physical restraints*</td>
</tr>
<tr>
<td>Pellfolk et al. (2010)</td>
<td>Cluster-RCT</td>
<td>n = 280</td>
<td>40 group dwelling units for people with dementia, Sweden</td>
<td>All residents of the included group dwelling units</td>
<td>6</td>
<td>Educational intervention; Usual care (control)</td>
<td>Physical restraint status</td>
</tr>
<tr>
<td>Testad et al. (2005)</td>
<td>Cluster-RCT</td>
<td>n = 141</td>
<td>4 nursing homes, Norway</td>
<td>All residents with dementia diagnosis (CDR)*</td>
<td>7</td>
<td>Educational intervention plus guidance; Usual care (control)</td>
<td>Physical restraints status</td>
</tr>
<tr>
<td>Testad et al. (2010)</td>
<td>Cluster-RCT</td>
<td>n = 145 baseline; n = 90 follow-up</td>
<td>4 nursing homes, Norway</td>
<td>All residents with dementia diagnosis, FAST* score ≥4</td>
<td>12</td>
<td>Educational intervention plus guidance; Usual care (control)</td>
<td>Physical restraints status (structural restraints)</td>
</tr>
</tbody>
</table>

*On three consecutive days at the end of the intervention period.

*FAST, Functional Assessment Staging; CDR, Clinical Dementia Rating Scale.
Definition of physical restraints

Definitions of PR in the studies were heterogeneous. Evans et al. (1997), Koczy et al. (2011) and Pellfolk et al. (2010) counted as PR all measures that inhibit a person’s free physical movement, for example, belts or chairs with tables, but bedrails were excluded. Huizing et al. (2009a), Testad et al. (2005, 2010) defined PR as any limitation of an individual’s freedom of movement, including belts or chairs with fixed tables, but also restrictive clothes (e.g. sleeping suits) and electronic measures, which could restrict a person’s movement (e.g. sensor mats or motion alarm systems). In addition to these structural restraints, Testad et al. (2010) defined a second form of restraint as interactional restraints, which were linked to treatment or care giving activities, for example, force or pressure in medical examination, treatment, or in activity of daily living procedures. These interactional restraints are not included in this review as their underlying concept is more related to elder abuse than to PR.

Description of interventions

In all studies (Evans et al. 1997, Testad et al. 2005, 2010, Huizing et al. 2009a, Pellfolk et al. 2010, Koczy et al. 2011), the interventions comprised an educational programme. In addition, consultation (Evans et al. 1997, Huizing et al. 2009a) or guidance (Testad et al. 2005, 2010) for nursing staff was offered. Koczy et al. (2011) offered a multifactorial intervention comprising nurses’ training as ‘change agents’, provision of technical aids such as hip protectors and sensor mats, and support for nurses by telephone or personal visits. Evans et al. (1997) provided two IG: in IG1, the educational programme was offered and additional consultation for nursing staff; in IG2, only the educational programme was offered. Figure 3 displays information on the intervention components investigated in the studies.

Underlying concepts of educational programmes

The educational programmes were based on existing educational programmes for hospitals (Huizing et al. 2009a), research by and experiences of clinical experts (Pellfolk et al. 2010), as well as on clinical experiences (Testad et al. 2010).

Content and delivery of educational programmes

The time frame in which the educational programmes were administered ranged from 1 day to 6 months. The total amount of education ranged from six to ten hours, with different numbers of educational sessions ranging from one to ten sessions. The duration of individual educational sessions ranged from 30- to 45-minute sessions to full-day seminars (for details see Fig. 3). Testad et al. (2010) offered a seminar lasting 2 days, the exact number of hours was not mentioned. Furthermore, Testad et al. (2010) provided insufficient information on the content of the educational programmes. Koczy et al. (2011) offered 6 hours of education to one nurse as ‘change agent’ per nursing home. These change agents were asked to disseminate the education’s content to nursing staff at their nursing homes. Educational programmes covered the following topics:

- Information on dementia, aggression and challenging behaviour (Testad et al. 2005, Pellfolk et al. 2010),
- Strategies for analysing and handling aggression or challenging behaviour (Evans et al. 1997, Huizing et al. 2009a, Pellfolk et al. 2010, Koczy et al. 2011),

© 2012 Blackwell Publishing Ltd
Journal of Clinical Nursing, 21, 3070–3081
Information on PR, for example, legal issues, adverse events, experiences of being restrained, correct use (Evans et al. 1997, Huizing et al. 2009a, Pellfolk et al. 2010, Koczy et al. 2011),

• Decision-making processes and alternatives to use of PR (Testad et al. 2005, Huizing et al. 2009a, Pellfolk et al. 2010, Koczy et al. 2011),


Consultation
Two interventions (Evans et al. 1997, Huizing et al. 2009a) comprised consultation provided by a nurse specialist at registered nurse (RN) level (Huizing et al. 2009a) or by a master’s-prepared gerontological nurse specialist (Evans et al. 1997). Evans et al. (1997) offered 6-months’ consultation for the corresponding IG, Huizing et al. (2009a) 8 months. The consultations included:

• multidisciplinary meetings, evaluating the use of physical restraints on individual residents, discussing difficulties in achieving PR-free care and stimulating the use of PR alternatives or less restrictive measures (Huizing et al. 2009a).

• discussions about residents with challenging behaviour or a history of multiple falls (Evans et al. 1997).

Guidance
Two studies (Testad et al. 2005, 2010) offered a monthly one-hour guidance session over 6 months following the single educational session. Guidance in the study by Testad et al. (2005) addressed the development of care plans for individual residents, taking into account the content of the educational session. Guidance in the study by Testad et al. (2010) addressed the implementation and reinforcement of new skills.

Support
Koczy et al. (2011) offered telephone support (advice by investigators) for change agents during the intervention period and in-house visits by a member of the research team on request.

Control group
In all studies, no intervention was offered in the CG (usual care). Characteristics of usual care were not reported in any of the studies.

Feasibility/Pilot test of interventions
None of the studies provided any information on a pilot test or a feasibility test of the intervention.

Implementation of interventions
Four studies (Evans et al. 1997, Testad et al. 2005, 2010, Pellfolk et al. 2010) offered some information on the implementation strategy. None of the studies evaluated the implementation process.

Nurse attendance at educational sessions
Huizing et al. (2009a) reported that 90% of staff attended at least four of five educational sessions. Evans et al. (1997) reported that 81% of nursing staff in IG1 and 78% in IG2 attended at least one of ten educational sessions, whereas 42% of IG1 and 39% of IG2 attended five or more sessions. In the study by Testad et al. (2010), all the nursing staff attended every educational and guidance session. In the study by Koczy et al. (2011), all change agents attended the
educational session. Pellfolk et al. (2010) and Testad et al. (2005) gave no information on the proportion of nurses attending educational session.

Nursing staff turnover
Attrition rates of nursing staff were reported in two studies. Testad et al. (2010) reported that 56 staff members (53.8%) in the IG and 53 (57.0%) in the CG were still employed at the end of the follow-up period. Reasons for attrition included retirement, pregnancy, long-term sick leave, moving or changing jobs. In the study by Testad et al. (2005), nursing staff attrition was only presented as the number of nurses who left the study, without reporting on the corresponding proportions.

Effects of interventions
Five of six studies presented the proportion of residents with PR as primary outcome (Evans et al. 1997, Testad et al. 2005, 2010, Huizing et al. 2009a, Pellfolk et al. 2010), and one study gave the number of residents without physical restraints (Koczy et al. 2011). None of the studies reported residents with newly applied PR.

Primary outcomes

Educational programme only
Evans et al. (1997) showed an absolute decrease in PR use after 12 months from 23% to 19% in the second IG and a decrease from 49% to 43% in the CG. No information is available regarding statistical significance. Pellfolk et al. (2010) presented data of participants with data at both baseline and follow-up as well as of all participants (including participants admitted during the study follow-up). For participants with data at both baseline and follow-up (n = 288), the proportion of PR use after 6 months was almost unchanged in the IG (21.5% to 21.1%), while PR use increased significantly in the CG (from 20.1% to 38.1%; p < 0.001). The difference between study groups was statistically significant (p = 0.001). Residents in the IG were less likely to be physically restrained at follow-up (OR 0.21; 95% CI 0.08–0.57; p = 0.002). For all participants (baseline n = 353; follow-up n = 350), PR use decreased significantly in the IG (25.7–16.8; p = 0.03) and increased non-significantly in the CG (24.7–33.3). Also, the difference between the study groups was statistically significant (p = 0.001).

Educational programme plus consultation
In the study by Evans et al. (1997), the prevalence in the first IG decreased after 12 months from 32% to 14% and in the CG from 49% to 43%. No information is available regarding whether these changes were statistically significant. Huizing et al. (2009a) found a significant increase in PR use in both study groups after 10 months. PR use increased in the IG from 54% to 64% (p = 0.02) and in the CG from 49% to 60% (p = 0.007).

Educational programme plus guidance
Testad et al. (2005) documented a decrease in the mean number of PR per week and resident after 7 months from 3.3 to 1.5 in the IG compared to an increase from 3.1 to 3.7 in the CG, a statistically significant difference between the study groups (p = 0.016). Testad et al. (2010) documented a decrease in PR use from 60% at baseline to 18% after 12 months in the IG, while PR use remained nearly unchanged in the CG (13% at baseline as well as 13% at follow-up visit).

Educational programme plus support and provision of technical aids
In the study by Koczy et al. (2011), 35 of 208 residents (16.8%) in the IG were not restrained after 3-month follow-up compared to 11 of 125 residents (8.8%) in the CG (OR 2.16, 95% CI 1.05–4.46).

Secondary outcomes

Types of restraints
Only one study presented results on types of restraints (Huizing et al. 2009a). In this study, bilateral bedrails were the most commonly used measures (baseline 48% IG and 41% CG; follow-up 52% both groups). At baseline, no differences between groups were found referring to the types of PR. At follow-up, statistically significant differences in the use of sleep suits were found (4% IG vs. 17% CG; p ≤ 0.01).

Multiple restraints
Huizing et al. (2009a) documented 17% of the participants with two different types of PR, 10% with three and 2% with four different measurements within 24 hours (mean types of PR per resident 0.93 ± 1.10). No significant difference in the frequency of multiple restraints was found between the study groups at baseline and follow-up. In both groups, multiple restraints significantly increased (p = 0.04, respectively, p = 0.002).

Restraint intensity
Restrain intensity indicates the number of observations per resident with PR in a given time period. In the study by

© 2012 Blackwell Publishing Ltd
Journal of Clinical Nursing, 21, 3070–3081
Huizing et al. (2009a), the mean number of observations per resident with PR in a 24-hour period at baseline was 1.36 ± 1.62. During follow-up, PR intensity increased significantly in both study groups (p = 0.001 for both groups). Koczy et al. (2011) documented the duration of PR in residents with PR during the 72-hour observation period at follow-up compared to the 72 hours at baseline and reported less time with PR for participants in the IG compared to participants in the CG.

Psychotropic medications
A further publication (Siegl er et al. 1997) about Evans et al. (1997) reported a non-significant increase in residents with at least one neuroleptic medication in both IG from baseline to follow-up (IG1 from 18.2% to 19.0%, IG2 from 13.5% to 15.5%). In the CG, the proportion of residents with at least one neuroleptic medication decreased significantly from 18.6% to 11.3% (p = 0.014). Benzodiazepines decreased significantly in all study groups (IG1 from 22.3% to 18.2%, IG2 from 37.2% to 27.0% and CG from 32.8% to 26.6%). In the study by Testad et al. (2005), psychotropic medication decreased in the IG from 71% to 55% and in the CG from 61% to 52%. Pellfolk et al. (2010) documented no difference in benzodiazepines and neuroleptics between study groups. Koczy et al. (2011) showed a slight decrease of the number of psychoactive drugs in both groups. In the study by Testad et al. (2010), study groups showed significant baseline differences in the frequency of participants with at least one psychotropic medication (28% IG vs. 9% CG; p = 0.03). At follow-up, psychotropic medication remained nearly unchanged in both study groups.

Falls and fall-related injuries
In the study by Evans et al. (1997), the CG showed a significantly lower baseline rate of residents with at least one fall event (CG 20.1% vs. IG1 33.1% and IG2 37.5%; p = 0.001). At follow-up, the CG showed a statistically significant higher number of residents with at least one fall event compared to both IG (CG 53.3% vs. IG1 37.8% and IG2 32.2%; p < 0.001). The total number of residents with fall-related injuries during follow-up was small (CG 2.2% vs. IG1 0% and IG2 5.3%; p = 0.02). In the study by Koczy et al. (2011), a non-significant difference of residents with at least one fall was reported at baseline (7% IG vs. 3.4% CG). After 3 months, the proportion of residents with at least one fall was 16.3% in the IG compared to 8.0% in the CG, also reported to be not statistically significant (OR 2.08; 95% CI 0.98–4.40). In both study groups, one resident with a fracture was documented. Pellfolk et al. (2010) assessed the proportion of residents with falls during a 1-month period before and after the intervention in both study groups and documented a non-significant decrease (IG from 11.4% to 10.1%, CG from 14.4% to 8.6%). No other studies reported on falls and fall-related injuries.

Adverse outcomes
Three studies reported challenging behaviour. For agitation, assessed by three subscales of the Cohen-Mansfield Agitation Inventory (CMAI), Koczy et al. (2011) found no significant differences between the study groups. In the study by Testad et al. (2010), the total CMAI score declined in the IG after 12 months, while it increased slightly in the CG. Testad et al. (2005) assessed agitation with the Brief Agitation Rating Scale (BARS). In this study, the BARS score in the IG increased, while it was nearly unchanged in the CG.

Discussion
Six cluster-randomised controlled trials were included in this review, showing inconsistent results.

Main outcome
Nursing homes
For nursing homes, one study with good methodological quality (Huizing et al. 2009a) showed no reduction but an increase in PR use at follow-up, while four studies with higher risk of bias (Evans et al. 1997, Testad et al. 2005, 2010, Koczy et al. 2011) found that their interventions reduced the use of PR. Although Koczy et al. (2011) reported a statistically significant difference in the proportion of residents free of restraints in favour of the IG, the results of this study must be interpreted with caution. The primary endpoint has been redefined, the statistical effect size estimate has been changed from the study protocol (Koczy et al. 2005), and a lower number of nursing homes was included than had been planned. The definition of PR also differs from the protocol. The cluster design was not properly considered. Use of adequate statistical methods would most likely have resulted in loss of statistical significance (Meyer & Köpke 2011).

Three studies included only one (Evans et al. 1997) or two (Testad et al. 2005, 2010) nursing homes per study group. In all studies, analyses did not consider an adjustment for cluster randomisation. Therefore, a unit of analysis bias cannot be excluded. Three studies showed baseline imbalance in PR use between study groups, in two studies these differences were statistically significant (Evans et al. 1997, Testad et al. 2010), while for one study this remains unclear (Koczy et al. 2011). Three studies (Testad et al. 2005, 2010, Huizing et al. 2009a) used a comparable definition of PR, whereas

© 2012 Blackwell Publishing Ltd
Journal of Clinical Nursing, 21, 3070–3081
two studies (Evans et al. 1997, Koczy et al. 2011) defined PR more narrow (i.e. excluding bedrails).

**Group dwelling units**

The study by Pellfolk et al. (2010) showed an almost unchanged proportion of PR in the IG compared to an increase in the CG. The intervention led to a statistically significant effect after including all residents admitted consecutively to the clusters during follow-up. Thus, the study suggests a preventive rather than a reductive effect. However, the results of the study were difficult to interpret and comparable studies investigating group dwelling units are lacking.

**Secondary outcomes**

Results from three studies (Evans et al. 1997, Pellfolk et al. 2010, Koczy et al. 2011) suggest that interventions do not lead to an increase in falls or fall-related injuries although only few fall-related fractures were observed in all three studies. As fall prevention and the fear of liability are common reason for the use of physical restraints (Evans & Fitzgerald 2002, Hamers et al. 2004), educational approaches should strongly emphasise the inappropriateness of PR use for fall prevention.

None of the four studies presenting results on psychotropic medication use (Evans et al. 1997, Pellfolk et al. 2010, Testad et al. 2010, Koczy et al. 2011) documented an increase in psychotropic medication in the IGs compared to CGs. Hence, it is unlikely that PR use has been replaced by psychotropic medication use.

The impact of physical restraint reduction programmes on challenging behaviour remains unclear because the studies reported inconsistent results.

The methodological quality of the studies was low to moderate. Only one study (Huizing et al. 2009a) showed a low risk of bias for selection of participants (i.e. adequate randomisation, allocation concealment and absence of baseline imbalances between study groups). As the cluster design was not adequately taken into account in most of the studies, centre effects might have influenced the results.

The included studies demonstrated significant clinical heterogeneity in terms of PR definition. Bedrails were not always counted as PR, although they have been shown to be the most commonly used PR measures in nursing homes (Meyer et al. 2009a). A consensus statement for conducting clinical trials aimed at reducing PR use, comparable to those published for fall injury prevention trials (Lamb et al. 2005) or hip protectors trials (Cameron et al. 2010), could help overcome the arbitrariness of research methods and PR definition. Such a consensus statement is currently in preparation (J.P.H. Hamers, 24 October 2011, personal announcement).

All included studies evaluated complex interventions (Craig et al. 2008). The quality of reporting with regard to the development and evaluation of the included interventions was insufficient. A more careful development of complex interventions on PR reduction is called for, including theory-based modelling of components and pilot testing of feasibility and acceptability (Craig et al. 2008). Furthermore, process evaluation is recommended to evaluate the degree of implementation of the interventions and barriers for implementation in clinical practice. Reporting of complex interventions should adhere to existing reporting statements, for example, CONSORT Statement for randomised trials of non-pharmacological interventions (Boutron et al. 2008) or cluster-randomised controlled trials (Campbell et al. 2004), as well as to the recommendations specific to complex interventions (Craig et al. 2008, Möhler et al. 2012). Only one study investigated group dwelling units for persons with dementia and no studies in the community setting could be identified. More high-quality research in residential care as well as in community care is needed.

**Conclusion**

The number of studies included in this review was small; studies were at low to high risk of bias and showed differences in terms of components of the interventions and definitions of PR. As the studies with higher risk of bias showed the interventions to be more effective than the single study with low risk of bias, the effects could be overestimated. In conclusion, there is insufficient evidence supporting the effectiveness of educational interventions targeting nursing staff for preventing or reducing the use of PR in geriatric long-term care. More research is needed, adhering to the recommendations for development and evaluation of complex interventions as well as using rigorous methodological approaches.

We have recently finished our own cluster-RCT, which will add further evidence of the effectiveness of educational interventions on reducing the use of PR. The intervention is based on an evidence-based guideline on PR reduction plus additional components for implementation (Haut et al. 2009). Results are promising, and a publication is expected in the near future. An update of our Cochrane Review is in preparation.

**Relevance to clinical practice**

The findings of this review indicate that educational programmes targeting nursing staff might not be effective
for reducing the use of PR in geriatric long-term care. However, this conclusion is based on a limited number of studies with various methodological shortcomings. At least one additional study evaluating an educational programme and an institutional policy might add valuable evidence on the effectiveness of educational approaches on PR reduction.

Acknowledgements

This project was funded by the German Ministry of Education and Research. We acknowledge the contributions of the consumer editors Johanna Waßmuß (Nurse Manager, Max-Brauer-Nursing Home, Hamburg, Germany) and Helga Schneider-Schelte (German Alzheimer Society, Berlin, Germany) to the development of the protocol. We also thank Pia von Lützau for her contributions to the development process of the review, in particular, for hand-searching the abstract books and data preparation. We would also like to acknowledge the Cochrane Dementia and Cognitive Improvement review group for their support.

Contributions

Study design: GM, SK, RM; data collection and analysis: RM, TR and manuscript preparation: RM, SK, GM, TR.

Conflict of interest

There are no conflicts of interest.

References


