Systematic review of the impact of arts for health activities on health, wellbeing and quality of life of older people living in care homes

Amy Curtis, Lucy Gibson, Mary O’Brien and Brenda Roe
Edge Hill University, UK

Abstract
Care home populations frequently feature older people who often experience poor physical health and cognitive difficulties, along with vulnerability to psychological and social stressors. To date there has been no systematic review which focuses on the impact of arts for health activities to the care home population. Evidence was sourced from several databases and 71 studies were deemed eligible for inclusion in this review. These studies underwent data extraction and quality appraisal and the findings associated with health, wellbeing and quality of life are presented within this paper.

Keywords
older adults, arts for health, care home, systematic review, wellbeing, quality of life

Introduction
As the population of older adults increases, it is likely that the number who require long-term care will also rise (WHO, 2014). Care homes are facilities offering varying levels of support to those who require nursing care or assistance carrying out activities of daily living, with differing levels offered depending upon need and care home type. In many countries of the world, care home facilities (such as nursing homes, residential homes, supported or assisted living or aged care homes) are operated by local councils, voluntary agencies
or are privately owned. Within the UK they must all be registered with the Care Quality Commission (Age UK, 2013; NICE, 2013).

Care home populations are diverse and residents often have chronic illnesses with many possessing more than one long-term health condition (Clift & Camic, 2015; Ham, Imison, Goodwin, Dixon, & South, 2011; O’Neill, 2009). A decrease in physical health can be a predictor for individuals moving into long-term care facilities (Age UK, 2015). In addition to physical health conditions, around 70% of individuals who reside within a care home may have dementia and those without a formal diagnosis will often have some degree of memory problems or cognitive difficulties (Alzheimer’s Society, 2014). Due to the physical and social changes associated with living within a care home environment, residents are also susceptible to depression (Age UK, 2016; Anderson et al., 2011; NICE, 2015) and loneliness (Cowl and Gaugler, 2014; Social Care Institute for Excellence, 2012).

There is recognition that care homes often adequately respond to the physical needs of residents; however, there is also suggestion that long-term care should incorporate a more holistic approach to meet psychological and sociological needs (New Economics Foundation, 2008; NICE, 2013; WHO, 2014, 2015). It has been demonstrated that not meeting such needs can increase incidences of anxiety, low mood and lead to behavioural disturbances in older people residing in care homes (deGuzman et al., 2011; Hancock, Woods, Challis, & Orrell, 2006). Quality of life is a quality indicator for care homes (Department of Health, 2014) along with the recommendation that older people in care homes are able to partake in ‘meaningful activities’ (NICE, 2015).

Research suggests residents can maintain quality of life, health and wellbeing by participation in arts-based activities. Benefits to older adults have been reported in terms of behaviour, psychological functioning, cognition, social relationships and overall health (Cowl and Gaugler, 2014; O’Connor, Ames, Gardner, & King, 2009a; O’Connor, Ames, Gardner, & King, 2009b, Kolanowski, Litaker, Buettnner, Moeller, & Costa, 2011; Vink, Bruinsma, and Sholten, 2011). Along with these benefits, the ability to continue learning throughout the lifespan and foster creativity have also been highlighted (Age UK, 2015; Clift and Camic, 2015; Cowl and Gaugler, 2014; Vella-Burrows, 2015; Young, Camic, and Tischler, 2015).

**Methods**

**Research aim**

Despite recognition of emerging benefits to older people, to date, there has been no systematic review focused exclusively on evaluating the evidence of arts for health activities for older people who reside in care homes. The aim of this systematic review was to locate, analyse, appraise and synthesise existing empirical studies which have evaluated the impact of arts for health activities with older people who reside in care homes. The outcome measures focused upon were health, wellbeing and quality of life.

**Registration of the review**

Prior to conducting the review, a protocol was developed and registered with the University’s Faculty Research Ethics Committee and the International Prospective Register of Systematic Reviews (PROSPERO, Curtis et al., 2015, CRD42015026264). PRISMA guidelines were followed for the methods and reporting of the review (Liberati
et al., 2009, Moher, Liberati, Ytetzlaff, & Altman, 2009) as well as the PICOS framework (Robinson, Saldanha, & Mckay, 2011).

**Inclusion criteria**

Published papers of empirical studies of arts for health activities with older people living in care homes were eligible for inclusion in this review. Arts for health activities could take place within or external to the care home (i.e. studies where residents visited arts establishments or where activities were carried out within a separate venue). The term ‘care home’ referred to nursing homes, dual registered care homes, aged care facilities and assisted living facilities. Study populations had a mean age of 60 years or above. Arts for health activities included those classified as performing arts, creative arts, visual arts and/or recreational activities, as defined by Arts Council England (2015). The arts activity could include active or passive engagement and be delivered individually or as a group. Study designs were qualitative, quantitative or mixed methods. For quantitative studies, there had to be a clear comparison between usual care or a comparison activity. However, uncontrolled pre–post/test studies were also eligible for inclusion where clear pre- and post-measures were reported.

**Exclusion criteria**

Studies were excluded where those taking part in the activity lived within their own homes, unless there was a proportion of residents who resided in a care home whose data could be extracted and included separately. Studies that reported on the effects of ‘reminiscence’ were also excluded unless they involved aspects of creativity such as ‘life story work’ or ‘musical based reminiscence’. Non-empirical studies were also excluded, along with those not published in English.

**Search strategy**

No date restrictions were applied to the search; key databases were searched from inception to December 2016, with continual updates obtained monthly. Databases included PsychInfo, AHMED, MEDLINE, EMBASE and CINAHL. The Cochrane Library, Joanna Briggs Foundation and Web of Science Arts and Humanities and Social Sciences Citation Index were also searched. Additional hand-searching of reference lists from systematic reviews and relevant papers was also conducted. It was specified that the returned articles were available in English (where the database search allowed for this specification). Search terms included arts activities: ‘dance therapy’, ‘music therapy’, ‘art therapy’, ‘singing’, ‘reading’, ‘poetry’, ‘reminiscence therapy’, ‘memory box’, ‘sew’, ‘knit’, ‘creative art(s)’, ‘paint(ing)’, ‘performing art’, ‘arts and health’, ‘arts and humanities’, ‘life story’, ‘life narrative’, ‘life review’, ‘theatre’, ‘visual art’, and ‘arts and wellbeing’ combined with terms related to care homes: ‘nursing home’, ‘residential home’ and ‘care home’ (Literature search strategy available from the lead author and online supplementary table S1 – Literature Search Strategy). Search terms were piloted prior to use.

**Search outcome**

A total of 1091 potentially relevant studies were identified (see Figure 1). Following removal of duplicates, 790 papers were screened for eligibility at title and abstract level with 624 removed that did not meet inclusion criteria. The resulting 166 papers
underwent full-text screening independently by two reviewers using a bespoke inclusion/exclusion form and 74 met the inclusion criteria. The 74 papers represented a total of 71 studies: 59 (83%) quantitative, eight qualitative (11%) and four (6%) mixed methods (see Figure 1).

**Figure 1.** PRISMA flow chart of included studies.

Data from the included studies were then extracted using a bespoke data extraction form based on the PICOS framework (Robinson et al., 2011) to extract all relevant data by the lead reviewer/author. All data extraction forms and pdfs of included studies were distributed amongst the review team and data extraction independently checked for errors or disagreements and consensus reached.
Quality appraisal of included studies was also undertaken by the lead reviewer and independently assessed and verified by another reviewer with agreement reached. The Cochrane tool GRADE (Higgins and Green, 2011) was used for all randomised controlled trials (RCTs). Appraisal of quasi-RCT and uncontrolled pre/post-test studies was also undertaken using these criteria. A score of low risk (+), high risk (−) or unclear risk (?) of bias was determined for each of the factors with a rationale explaining each decision, in keeping with the standards for appraising quality in a systematic review (Guyatt et al., 2011; Higgins and Green, 2011; Higgins et al., 2011) (see Box 1). Risk of bias graphs was created for quantitative studies using Review Manager 5.3 (RevMan).

A tool developed by Walsh and Downe (2006), based on Lincoln and Guba (1985), was used to appraise and score qualitative studies (Cooke, Mills, & Lavender, 2010a; Downe, Finlayson, Walsh, & Lavender, 2009) (see Box 2).

### Box 1
Criteria based on GRADE (Higgins and Green, 2011) to assess the risk of bias in quantitative studies.
- Evidence adequate randomisation took place
- Evidence of allocation concealment
- Evidence of blinding of outcome assessment – given the nature of the arts activities it would not be possible to blind those taking part and therefore this was judged to have occurred if the person scoring the outcome assessment was blind to group allocation.
- Complete/incomplete outcome data
- Selective reporting
- Any other apparent risk of bias

### Box 2
Quality appraisal criteria for judgement and scoring of included qualitative studies. See Cooke et al. (2010a) and Downe et al. (2009).
- Whether the scope, purpose and aims were presented and clear
- If the study design was appropriate and well explained
- Sampling size adequacy and characteristics
- Whether there were data and evidence to support the author’s interpretation of the results
- Evidence of reflexivity from the author
- Whether ethical considerations had been adhered to
- Evidence of relevance and transferability.

A – No, or few flaws. The study credibility, transferability, dependability and confirmability are high.
B – Some flaws, unlikely to affect the credibility, transferability, dependability and/or confirmability of the study.
C – Some flaws that may affect the credibility, transferability, dependability and/or confirmability of the study.
D – Significant flaws that are very likely to affect the credibility, transferability, dependability and/or confirmability of the study.
All studies were allocated a level of evidence score: I results based on a RCT, IIa results based on well-designed controlled study without randomisation, IIb evidence from a well-designed quasi-experimental design or III evidence from well-designed descriptive studies (NICE guidelines, 2012).

**Data synthesis**

Given the number and heterogeneity of included studies, their differing study designs, methods and outcome measures, data are presented in a narrative descriptive summary. This provides an overview and high level comparison of the different studies, their characteristics, key findings in relation to health, wellbeing and quality of life and their quality appraisal.

**Findings**

**Characteristics of included studies**

A table of included studies is available online, supplementary table 2 – table of included studies. Of the 71 included studies publication dates ranged from 1983 to 2015 with the majority published from 2000 onward. There was an increase in publications per decade, with four from the 1980s (4%), 14 from the 1990s (20%), 29 (41%) between 2000 and 2010 and 28 (39%) from the five-year period, 2010–2015. Study dates were only reported by 13 (18%) studies (Chang, Lu, Lin, & Chen, 2013; Cooke et al., 2010a; Doric-Henry, 1997; Guetin et al., 2009; Houser, 2014; Lin, 2010; Liu, Niu, Feng, & Liu, 2014; Myskja & Nord, 2008; Palo-Bengtsson, Winblad, & Ekman, 1998; Philips, 2010; Ragneskog, Kihlgren, Karlsson, & Norberg, 1996b; Ridder, Stige, Qvale, & Gold, 2013; Roe, 2016) and these ranged from ‘spring’ 1991 to December 2012.

Sixty-three (89%) studies were intervention studies: 59 were quantitative (83%) and four mixed methods (6%). Eight studies were qualitative (11%), all of which were descriptive. A total of 27 quantitative studies used uncontrolled pre/post-test designs (38%), 23 (32%) reported on RCTs, 12 quasi-experimental designs (17%) and one being a retrospective cohort study.

Each of the studies was broadly classified via arts type and consisted of the following: 48 (68%) involved activities classified as ‘music’, eight (11%) as multi-sensory activities, six (8%) as ‘spoken and written word’, six (8%) as performance and three as applied arts and crafts.

**Location and setting**

A third of studies took place within the USA (24, 33%). Second most cited countries were the UK (seven, 10%), Australia (seven, 10%) and Taiwan (seven, 10%). Six were carried out in Sweden (8%), five within Canada (7%) and two in Israel. The remaining studies were carried out in Belgium, Iceland, Norway, France, Holland, Spain, China, Japan, South Korea, Iran, Czech Republic and the Philippines, with one study originating from each country. Only one study featured data from more than one country, with participants included from 14 care homes in total, 10 from Norway and four from Denmark (Ridder et al., 2013).

The term ‘nursing home’ was used by 29 of the studies (41%), 11 (15%) referred to ‘care homes’, seven (10%) to ‘long-term care facilities’ and five (7%) reported residents lived
within ‘residential homes’. Lesser featured terms were ‘dementia specialty facility’ (n = 3), ‘residential facility’ (n = 2), ‘senior living residence’ (n = 2), ‘residential units’ (n = 1), ‘care facility’ (n = 1), ‘care retirement community’ (n = 1), ‘continuing care facility’ (n = 1), ‘memory care residence’ (n = 1) and ‘supported living facility’ (n = 1). Five studies did not specify a term for where participants resided.

Most studies (49, 69%) reported the arts activity took place within the care home itself. There was just one study which carried out an activity external to the home, where residents visited a gallery and museum (Roe, 2016). The remaining 21 (30%) studies did not explicitly state where the arts activity took place, although, it is implied from the description of the activities that they were carried out within the home environment.

**Aims of the included studies**

Most studies aimed to assess the impact of the activity on more than one outcome measure. The most popular being the impact of the activity on health behaviour symptoms reported by just under half of the studies (n = 32, 45%). Many of these studies focused on determining whether there were reductions in agitated behaviours (n = 17, 24%). In addition, four studies aimed to assess the impact on verbally disruptive behaviours (VDB), five on ‘behavioural symptoms associated with dementia’, two exploring ‘irritability and restlessness’ and one focusing specifically on ‘exit-seeking behaviours’ exhibited by those who resided upon a ward within a care home. Five studies explored positive behaviours, aiming to determine levels of engagement with the arts activity.

The second most reported aim was assessing the impact of the arts activity on measures of psychological wellbeing, cited by a total of 24 studies (34%). Specific aims included the effects on depression, anxiety and general wellbeing measures.

Assessments of cognition were reported by 10 studies (14%), nine (13%) focused on the influence of the arts activity on the care process and eight (11%) explored quality of life. Lesser reported outcomes included feasibility of carrying out arts for health activities (n = 6, 8%), physical health (n = 4), satisfaction with living environment (i.e. the care home) (n = 2), socialisation (n = 2) and neuropsychiatric symptoms associated with dementia (n = 2).

**Characteristics of included populations**

The total number of older adults included was 2086 (reported for a total of 70 out of the 71 studies, 99%). There was large variation in sample size, from 3 to 117. Just over 70% of the studies (n = 51) reported the mean age of those involved with the arts activity (or data which allowed for the calculation of a mean age) and across these the mean age of residents was 84 years, with a range between 69 and 87.5 years. Forty-four of these 51 studies (86% of those which reported a mean age) recorded their sample as aged 80 years or above and a further six studies reported a mean age above 70 years (Casby & Holm, 1994; Garland, 2006; Ho et al., 2011; Hong & Choi, 2011; Kincaid & Peacock, 2003; Sung, Chang, Lee, & Lee, 2007). This left only one study which reported a mean age of less than 70 years; however, this was very close at 69.4 years (Mohammadi, 2011).

The gender of those taking part was reported by 62 studies (87%), with women making up three quarters of the total sample across the studies (75%, n = 2068), and men representing 25% of the total (n = 508). Only three studies featured more men than women (Mohammadi, 2011; Sung et al., 2006; Sung, Chang, & Lee, 2010), two reported equal
numbers of men and women (Hilliard, 2004; Ragneskog, Braane, Karlsson, & Kihlgren, 1996a) and two reported on data obtained from women only (Carruth, 1997; deGuzman, 2009).

Just under half the studies specified within inclusion criteria that residents required a dementia diagnosis to participate (n = 35, 48%). Whilst not specifying a dementia diagnosis, the requirement for participants to display ‘cognitive difficulties’ was specified by Klages, Zecevic, Organce, and Hobson (2011). In addition, Simard and Volcier (2010) reported most participants to have ‘severe’ Cognitive Performance Scale scores. In addition, Martin et al. (2004) explicitly stated no residents were excluded from taking part based on cognitive abilities.

In contrast, nine studies featured inclusion criteria which may have prevented those with dementia (or other forms of cognitive impairment) from partaking (Bennett & Maas, 1988; Burrack, Jefferson, & Libow, 2004; Chang et al., 2013; Chen, Lin, & Jane, 2009; Hagen, 2004; Houston, Mckee, Carroll, & Marsh, 1998; Scott, Masser, & Pachana, 2014; Suzuki, 1998; Vankova et al., 2014). Both Hagen (2003) and Houston et al. (1998) excluded participants who had ‘severe cognitive impairment’. AQ: The reference “Hagen (2004)” is cited in the text but is not listed in the references list. Please either delete in-text citation or provide full reference details following journal style.] Bennett and Maas (1988) excluded people with ‘mental dysfunction or confusion’, and Chang et al. (2013) included only those with ‘good cognitive functioning’. Suzuki (1998) specified similar criteria outlining that residents required ‘adequate reality orientation’ to complete the outcome measures. Mini Mental State Examination (MMSE) cut-off scores were adopted as inclusion criteria by Burrack et al. (2004), Chen et al. (2009) and Vankova et al. (2014), with Scott et al. (2014) also assessing MMSE scores, although the cut-off was not reported.

**Outcome measures**

Of the studies which reported changes in behaviour as an outcome measure, 16 (23%) used the Cohen–Mansfield Agitation Inventory (CMAI). Five further studies used observational methods to assess behavioural changes, of which four devised their own schedules and one used an existing tool, the Behavioural Pathology in Alzheimer’s Disease Rating Scale (Svansdottir & Snaedal, 2006). Two studies assessing engagement used the Menorah Park Engagement Scale (Heyn, 2003; Skrajner & Camp, 2007), another used the Myers-Research Institute Engagement Scale (Orsulic-Jeras, Schneider, & Camp, 2000).

The most commonly assessed area of psychological wellbeing was depression as measured by the Geriatric Depression Scale (Cooke et al., 2010a; Guetin et al., 2009; Scott et al., 2014; Subramaniam, Woods, & Whitaker, 2014; Vankova et al., 2014) and Cornell Scale for Depression (Ashida, 2000; Clarkson, 2007; Orsulic-Jeras et al., 2000; Philips, 2010). Lesser utilised measures were the Montgomery Asberg Rating Scale (Myskja, 2008), Gottfries–Brane–Steen Scale (Ragneskog et al., 1996a), Self-Rating Depression Scale (Liu, 2014), Dementia Mood Assessment Scale (Koike et al., 2012) and Beck Depression Inventory (Doric-Henry, 1997). Levels of anxiety were measured using the Rating Anxiety in Depression tool (Cooke, Moyle, Shum, Harrison, & Murfield, 2010b; Sung et al., 2010), Hamilton Scale for Anxiety (Guetin et al., 2009) and the Geriatric Anxiety Inventory (Scott et al., 2014). Combined measures of both anxiety and depression were obtained via the Depression and Anxiety Stress Scale (Mohammadi, 2011) and the Hospital Anxiety Depression Scale (Houston et al., 1998). In addition, the use of psychotropic medications,
often used to moderate anxiety and agitation, was collected in five studies (Brontons, 2009; Clarkson, 2007; Houser, 2014; Ridder et al., 2013; Simard & Volicer, 2010). Assessment of general mood and affect was assessed almost exclusively via observational methods using the Observed Emotion Rating Scale (Hammar, Emami, Gotell, & Engstrom, 2011; Philips, 2010), the Philadelphia Geriatric Affect Rating Scale (Fritsch et al., 2009) and the Affect Rating Scale (Cox, Burns, & Savage, 2004). Un-validated measures of mood were reported in three studies (Heyn, 2003; Lord & Garner, 1993; Sole, Mercadal-Brontons, & Galati, 2014).

Quality of life measures were assessed via the Quality of Life in Alzheimer’s Disease Scale (Burrack et al., 2004), Dementia Quality of Life Scale (Cooke et al., 2010a), Alzheimer’s Disease Rated Quality of Life, the Life Satisfaction Index-A (Bennett & Maas, 1988) and an un-validated scale created by Vanderark, Newman, and Bell (1983). Resident’s satisfaction with their living environment was measured in two studies, each devising their own measure (Chang et al., 2013; Scott et al., 2014) but only the first study showed validation of the questionnaire within their study (Satisfaction with Living Environment Nursing Home).

Cognition was evaluated via scoring of the MMSE (Hong & Choi, 2011; Smith, 1986), Memory Retrieval Test (Suzuki, 1998), Cognitive Assessment Scale (Hagen, Armstrong-Esther, & Sandilands, 2003), Functional Assessment of Communication Skills (Philips, 2010) and Autobiographical Memory Interview (Subramaniam et al., 2014). One study used face–name recognition (Carruth, 1997) and another assessed ‘alert responses’ (Clair, 1996). Finally, Noice and Noice (2006) administered a range of measures which assessed word recall, listening span and the Means End Problem Solving Procedure.

Physical health was measured by balance and joint movement range (Hagen et al., 2003), oxygen saturation, temperature and pulse (Koike et al., 2012) and balance tests (Klages et al., 2011). Overall functioning was assessed via use of the Multidimensional Observational Scale of Elderly Subjects by Orsulic-Jeras et al. (2000).

Measures of behavioural changes and agitation were therefore always obtained via observations of residents, along with quality of care. General expression of ‘emotion’ and ‘affect’ was mostly obtained through observational methods, except for Koike et al. (2012). Residents always completed quality of life measures and depression outcome measures for all but two of the studies (Myskja, 2008; Orsulic-Jeras et al., 2000) both of whom used care staff to complete perceived depression. Similarly, anxiety was self-reported for all but one of the studies (Sung et al., 2006).

**Quantitative findings**

**Behaviour and symptoms associated with dementia**

Significant changes over time were reported by seven of the studies which utilised the CMAI (Garland, Beer, Eppingstall, & O’Connor, 2007; Goddaer & Abraham, 1994; Hicks-Moore & Robinson, 2007; Ho et al., 2011; Ledger & Baker, 2007; Lin et al., 2010; Vink et al., 2013). Despite this, four of the five studies that featured a comparison group taking part in other activities failed to show a significant difference between the two groups (Hicks-Moore & Robinson, 2007; Ledger & Baker, 2007; Remington, 2002; Vink et al., 2013). In contrast, two studies showed significant differences between the scores obtained between the comparison and control group when the comparison group took part in ‘normal daily activities’ (Lin et al., 2010; Sung et al., 2006), suggesting that arts activities were more effective than
normal daily activities. Further evidencing this, Garland et al. (2007) showed both the arts activity (music) and comparison activity (simulated family presence) to significantly reduce agitation, but not one more so than the other. Despite this, both were more effective than the third control condition which involved no activity.

In respect of other behavioural symptoms, both Hagen et al. (2003) and Svansdottir and Snaedal (2006) showed changes over time for both groups (those taking part in the arts activity and the control group). No significant differences were found between the two groups for the first study (Svansdottir & Snaedal, 2006) and the second did not provide a comparison (Hagen et al., 2003). Houser (2014) and Orsulic-Jeras et al. (2000) failed to show significant effects of the arts activity. In contrast, Brontons (1998) did show the control group to have significantly increased behavioural disturbances at post-test; however, these scores fluctuated throughout the study for both groups and did not change over time.

Studies specifically focused on VDBs produced mixed findings. Cohen-Mansfield and Werner (1997) showed decreases over time, but not more so for the music group compared with the control group. Casby and Holm (1994) also suggested decreases over time but with only three participants and did not report any statistical analyses. Increases in verbal disruptions were evidenced in three studies (Clarkson, Cassidy, & Eskes, 2007; Cooke et al., 2010a; Nair et al., 2011); however, none explicitly stated what types of verbalisations increased. Ziv, Granot, Hai, Dassa, and Halmov (2007) showed a significant increase in positive verbalisations whereas negative verbalisations decreased. Similar anecdotal evidence was provided by both Millard and Smith (1989) and Sole et al. (2014). Further to this, four studies provided evidence of increased engagement with the activity (Fritsch et al., 2009; Holmes, Knights, Dean, Hodkinson, & Hopkins, 2006; Orsulic-Jeras et al., 2000; Skrajner & Camp, 2007). Again, descriptive accounts were provided by Heyn (2003) and Millard and Smith (1989) but no statistical evidence was provided.

None of the studies which specifically focused on dementia symptomatology showed improvements (Billington, Carroll, Davis, Healey, & Kinderman, 2013; Houser, 2014; Philips, 2010). The only study to assess door testing (Kincaid & Peacock, 2003) showed a statistically significant fall in exit-seeking behaviours.

**Psychological wellbeing**

Nine studies analysing depression showed significantly reduced scores over time (Doric-Henry, 1997; Guetin et al., 2009; Houston et al., 1998; Koike et al., 2012; Liu et al., 2014; Mohammadi, Shahabi, & Panah, 2011; Myskja, 2008; Ragneskog et al., 1996a; Vankova et al., 2014). However, eight failed to show any significant impact of the activity on depression levels (Clarkson et al., 2007; Cooke et al., 2010a; Houser, 2014; Orsulic-Jeras et al., 2000; Philips, 2010; Scott et al., 2014; Simard & Volicer, 2010; Subramaniam et al., 2014).

All but one of the studies (Cooke et al., 2010a) which assessed changes in anxiety showed a significant reduction in related scores over time (Guetin et al., 2009; Houston et al., 1998; Mohammadi et al., 2011; Sung et al., 2010; Svansdottir & Snaedal, 2006). Additionally, the four studies which featured a comparison group all showed significantly lower anxiety scores at follow-up for those who had taken part in the arts activity (Guetin et al., 2009; Houston et al., 1998; Sung et al., 2010; Svansdottir & Snaedal, 2006).
Significantly improved scores over time were also reported for general mood measures (Hammar, Emami, Engstrom, & Gotell, 2010; Lord & Garner, 1993; Sole et al., 2014; Suzuki, 1998) with evidence that the arts activity group improved significantly more than a comparison group (Lord & Garner, 1993). Noice and Noice (2006) failed to show improvements related to self-esteem or general psychological health. Increases in expression of positive emotions were also observed to significantly improve in Hammar et al. (2010), although this finding was not shown by Fritsch et al. (2009) or Suzuki (1998). However, Suzuki (1998) demonstrated a significant reduction in expressions of negative affect indicative that whilst positive emotions were not more apparent, there were reduced expressions of negative emotions. Fritsch et al. (2009) found that whilst those taking part in the arts activity displayed more levels of anger and fear, they also displayed less neutral affect.

Of the studies assessing medication changes, two reported that there was not enough data to analyse any trends (Brontons, 1998; Clarkson et al., 2007) and a further study found no significant changes (Houston, 2014). In contrast, two studies showed an impact of the arts activity, one via increased prescriptions of antipsychotic medications for the control group (Ridder et al., 2013), and the second through decreased anti-anxiety prescriptions throughout but no changes in antidepressants (Simard & Volicer, 2010).

Quality of life and general wellbeing
Evidence of benefit on general quality of life and wellbeing measures was limited, with improvements being reported by six studies (Bennett & Maas, 1988; Chang et al., 2013; Hagen et al., 2003; Hilliard, 2004; Philips, 2010; Subramaniam et al., 2014). Even these findings were tentative, Hilliard (2004) assessed length of life and determined that those taking part in music therapy lived longer. However, this failed to account for the fact that these individuals may have exhibited less advanced disease than those not taking part and therefore may have had greater life expectancy related to disease stage irrespective of the arts activity. Chang et al. (2013) also reported positive findings for only one measure of satisfaction, the living environment. This was associated with recalling pleasant memories which may indicate improvements in memory rather than quality of life. Sole et al. (2014) actually reported worsening of quality of life scores, which occurred for both the intervention and control group, although not to a significant degree. Both Hagen et al. (2003) and Subramaniam et al. (2014) showed improvements for those who took part in the arts activity at post-test; however, this was not sustained at follow-up, where both groups showed comparative improvements.

Cognition
With respect to cognition, both studies using the MMSE showed improvements over time for those taking part in the music activity (Hong & Choi, 2011; Smith, 1986). Scores were also significantly improved compared with the control group for the RCT (Hong & Choi, 2011). Both Hagen et al. (2003) and Subramaniam et al. (2014) showed improvements in cognitive assessment and memory scores, respectively, over time; however, this was not compared with other activities. Similarly, Clair (1996) showed alert responses to have increased for the music condition significantly more than the silent condition, but not significantly more so than for the reading condition. Unpleasant memories decreased for those taking part in music reminiscence but the percentage of pleasant memories did not
increase (Suzuki, 1998). Working memory did not significantly improve when taking part in drama in Noice and Noice (2006), although word recall and problem solving did. Social communication was not shown to significantly increase (Philips, 2010) and there were very limited findings for face–name recognition (Carruth, 1997), where ‘four participants showed an increase, three did not’ with no statistical comparisons, possibly due to the limited sample size.

**Quality of care and physical health**

Three quantitative studies reported the aim of assessing quality of care. Two of these explored similar outcomes (Ragneskog et al., 1996a, 1996b) focusing on nutritional intake during mealtimes. The first showed nurses fed the residents significantly more during the music conditions and residents spent more time with them during meals. The second showed residents consumed larger portions of food. Hammar et al. (2011) was the only study to focus on restiveness to care. It was found that significantly less pulling away, grabbing objects and behaviour adduction were observed during the Music Therapeutic Caregiving condition as opposed to the normal morning care routines.

Significant physical health improvements were only shown by Koike et al. (2012) who found pulse and temperature to have improved. Balance (Hagen et al., 2003; Klages et al., 2011) and overall physical ‘functioning’ (Orsulic-Jeras et al., 2000) did not improve due to the arts activity.

**Qualitative finding**

Qualitative findings overlapped with quantitative findings in terms of identified themes and benefits. All qualitative studies outlined improvements to wellbeing. Specified by several of these was the ability for the arts activity to improve a sense of self in residents (Billington et al., 2013; Chen et al., 2009; DeGuzman, 2011; La Cour, 2005; Martin et al., 2004). These five studies also reported on the participation of residents, despite physical barriers which may exist. Creativity and the opportunity to take part in meaningful activity were reported by Billington et al. (2013), Chen et al. (2009), Houser (2014) and La Cour, Josephsson and Luborsky (2005). Cognitive improvements were reported by three studies including benefits to listening, memory and attention (Billington et al., 2013) and concentration (Palo-Bengtsson et al., 1998).

Difficulties with the usual care process were reported by studies which conducted interviews with staff members (Gotell, 2007; Hammar et al., 2010; Houser, 2014). This was contrasted with providing care during and following the arts activity which was reported as being more unified and cooperative. The ability for staff members to use their creativity from the sessions for problem solving during work was also highlighted by both Billington et al. (2013) and Guzman-Garcia (2013) along with general improvements to the overall atmosphere of the care home (Houser, 2014). Improved communication between the residents and staff members was also evidenced (Gotell, 2007; Hammar et al., 2010; Roe, 2016).

Increased socialisation was also reported in terms of interacting with other residents and between the care staff and residents (Billington et al., 2013; Chen et al., 2009; Guzman-Garcia, 2013; Martin et al., 2004; Roe, 2016). Sharing of memories prompted by
the arts activity was highlighted in all five of these studies to aide and promote these interactions.

The feasibility of carrying out the arts activity was described within Billington et al. (2013), Cox et al. (2004) and Roe (2016). Aspects of the programme that were discussed included the process of physically situating residents at the arts activity, staff requirements, physical barriers to participation and carrying out the arts activity itself. Two studies that featured changes in environments, Cox et al. (2004, Snoezelen and garden activity) and Roe (2016, museum and gallery) described the benefits of exploring novel surroundings that differed from the care home. However, both also contained accounts of times when this was distressing for certain residents.

Quality of included studies

Levels of evidence scores were assigned to each of the studies. Twenty-three (32%) studies provided level I evidence based on an RCT, 11 (15%) provided level IIa results based on well-designed controlled study without randomisation and most, 29 (41%) level IIb based on well-designed quasi-experimental studies. The remaining eight studies (11%) corresponded to an evidence score of level III, well-designed descriptive studies.

RCTs

The summary of RCT judgements of quality can be found in Figure 2. Twelve of the RCTs provided adequate accounts of randomisation to obtain a low risk score for this measure and nine of these also gave detail as to methods of allocation concealment. Blinding of outcome assessment was reported for a total of four of the studies, with nine reporting no blinding to have taken place. The remaining studies did not provide enough information to make a judgement either way. Nine studies received judgements of low risk for incomplete data, just two receiving high-risk judgements, and the remaining studies did not contain enough data to make a judgement. Fourteen appeared to be free of selective reporting, four showed selective reporting and the remaining studies did not provide enough information. Finally, eight studies possessed other biases which may affect the credibility of the studies.

Quasi-RCTs

Summary of quasi-RCT judgements is presented in Figure 3. Two of the quasi-RCTs demonstrated blinding of the individual scoring the outcome assessment, five stated this did not happen (therefore judged as high risk) and four did not contain enough information to make a judgement. Eight were unclear as to whether there was any missing outcome data, two were judged as high risk and just one was deemed to be low risk. In terms of selective reporting, there was only evidence of this occurring within one study, a further six were judged as low risk and it was not clear whether this had taken place in the remaining two studies. Two demonstrated risks of other biases, six were deemed low risk and two did not provide enough information.

Uncontrolled pre/post-test

Summary of uncontrolled pre/post-test judgements is shown in Figure 4. Evidence of blinding of outcome assessments was shown for just two studies; this explicitly did not
Figure 2. Risk of bias judgement for randomised controlled trials.
occur for 10 studies, and there was insufficient information to form a judgement about the others. Incomplete data were evidenced for seven studies, with 18 unclear and just one judged as having a low risk of bias. Selective reporting was also apparent in 14 studies, with just five receiving a judgement of low risk for this category. Finally, most of the studies

Figure 3. Risk of bias judgments for quasi-randomised controlled trials.
Figure 4. Risk of bias judgments for uncontrolled pre/post-test studies.
(15 in total) showed evidence of other biases, with only five low-risk judgements for this category.

Qualitative

Overall quality of the qualitative studies was judged to be good, with five studies (deGuzman, 2011; Gotell, 2007; Guzman-Garcia, 2013; La Cour et al., 2005; Roe, 2016) classified as Grade B for risk of bias and three judged at grade C (Chen et al., 2009; Martin et al., 2004; Palo-Bengtsson et al., 1998).

Discussion

The interest in arts activities for older people in care homes is illustrated by the increase in publications over time. Eighty per cent of the papers sourced within this review had been published since the year 2000, with 41% published between 2010 and 2015 alone. A higher proportion of studies focused upon music, accounting for almost 70% of the papers meaning that much of the evidence relates to this specific arts type. Future research would therefore benefit from explorations of other arts for health activities to achieve a greater evidence of potential benefits associated with older people living in care homes.

The focus of studies has also shifted over time. Apart from Hong and Choi (2011), studies with the aim of purely assessing cognitive changes were carried out in 1986 (Smith) and 1996 (Clair). Whilst subsequent studies have also explored the role of cognition, these also aimed to explore other areas such as quality of life or psychological wellbeing. Another (Philips, 2010) determined changes in social communication therefore focused more on socialisation. Similarly, the measures of quality of care within earlier studies (Ragneskog et al., 1996a, 1996b) provide evidence in regards to food consumption or nutritional intake, whereas latter studies explored the relationship between the carer and residents, along with restiveness to care (Gotell, 2007; Hammar et al., 2011; Houser, 2014).

Evidence was sourced from a range of study designs and there were disparities in quality across the studies. Unsurprisingly RCTs fared well in terms of quality, with other study designs showing a lower quality. This may partially be due to use of GRADE criteria for all study designs. A good standard of evidence was reflected in the scores obtained via qualitative evidence, particularly in those carried out within the last 10 years.

Overall there was mixed evidence of benefit for arts for health activities within care home populations. It may be suggested that negative findings may still be attributed to benefits. There was evidence that increased verbalisations were associated with increases in positive behaviours (Millard & Smith, 1989; Sole et al., 2014; Ziv et al., 2007), such as joining in with the arts activity which consequently reduced apathy in residents. Studies reporting on increased verbal disruptions did not discuss whether these were in respect of negative behaviours indicative of agitation or anxiety, or, whether this represents increased participation (Clarkson et al., 2007; Cooke et al., 2010a; Nair et al., 2011). Further evidence of the latter is provided by increased engagement with the arts activity (Fritsch et al., 2009; Holmes et al., 2006; Orsulic-Jeras et al., 2000; Skrajner & Camp, 2007) and increased expressions of positive affect (Hammar et al., 2010) along with a decrease in negative affect (Suzuki, 1998). Whilst Fritsch et al. (2009) showed increased incidences of fear and anxiety, there were also reductions in neutral affect, indicative that residents could express their emotions, even if these were negative, therefore improving levels of communication. Furthermore, qualitative
evidence showed a unique outcome regarding increased sense of self, which was not captured by any quantitative studies (Billington et al., 2013; Chen et al., 2009; DeGuzman, 2011; La Cour, 2010; Martin et al., 2004). A decrease in self-reported ratings of quality of life outlined in Sole et al. (2014) may also be attributed to this increased self-awareness, again suggesting residents may become more aware of their emotions. Interestingly Noice and Noice (2006) failed to quantitatively capture increases in self-awareness and was the only study within this review which stated that residents were told when they were making mistakes when acting and therefore may have felt discouraged. Further analysis on the aspects of arts for health activities would be useful in determining what content works best.

There is also evidence that arts for health activities may be particularly beneficial for those with existing clinical need, such as pre-existing depression and anxiety. Both Liu (2014) and Suzuki (1998) described populations with clinical depression who showed improvements when taking part in arts activities. In direct contrast, Orsulic-Jeras et al. (2000) comment on low depression scores at baseline may explain a lack of effect within this study, given that residents already functioned well psychologically. A sub-analysis carried out by Simard and Volicer (2010) showed that improvements to behaviour were only evidenced in patients with low social interaction levels at baseline, which further evidences the fact that low pre-existing functional levels may be needed to show improvements.

Despite positive findings over time for those taking part in arts activities, there was limited evidence for arts activities to foster improvements more than comparison activities. Many quantitative studies exploring the effects of the arts activity on behaviour and psychological wellbeing showed improvements over time, but not compared with another activity. However, there was evidence that taking part in such arts activities improved outcomes for those in care homes more so than when no activities were offered. Further illustration is provided by qualitative findings that would be difficult to capture quantitatively, the most relevant being the opportunity for creativity (Billington et al., 2013; Chen et al., 2009; Houser, 2014; La Cour et al., 2005), which is unlikely to improve taking part in other activities. Improved socialisation with both peers and staff members was also captured by qualitative evidence (Billington et al., 2013; Chen et al., 2009; Guzman-Garcia, 2013; Martin et al., 2004; Roe, 2016) meaning both types of evidence are needed to ascertain the true benefits associated with the arts.

**Limitations of the review**

This review was confined to studies published in English which therefore may have impacted upon the countries evidence was sourced from. This may therefore mean not all cultural practices or countries being represented. A meta-analysis was also not able to be performed given the differences and heterogeneity of the included studies and this should be undertaken where possible in the future when evidence from similar studies and outcomes becomes available.

**Conclusion**

International evidence emerging to date for arts for health activities for older people in care homes whether within or external to the home looks promising and can benefit residents’ health, wellbeing and quality of life. This review identifies evidence for including arts
activities within care homes and which activities may ascertain the greatest benefits. However, more research is required, particularly in relation to lesser evidenced arts activities such as crafts. Larger mixed methods complex intervention studies with qualitative elements embedded that adhere to international recognised standards for conduct and reporting such as CONSORT (Moher, 1998) and the MRC Framework for complex interventions/studies (Moore et al., 2015) are warranted. Where possible future meta-synthesis of qualitative evidence and meta-analysis of quantitative outcomes should be conducted.

**Declaration of conflicting interests**

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Funding**

The author(s) received no financial support for the research, authorship, and/or publication of this article.

**References**


**Amy Curtis** is currently undertaking her PhD at Edge Hill University, having previously completed a Master’s Degree by Research (MRes) with a focus on epidemiology. Prior to beginning her PhD Amy worked in many areas of mental health and her research interests include healthy ageing and ways to maintain positive mental health and wellbeing.

**Lucy Gibson** is a senior lecturer in Applied Health and Social Care at Edge Hill University. She has published and presented research on popular music and ageing, the extension of youth cultural practices in adulthood, using email interviews/online research methods, and the ageing body.

**Mary O’Brien** is a professor of Palliative and Supportive Care in the Faculty of Health and Social Care at Edge Hill University. Her research centres on the needs of people with life
limiting illness and their family carers with a particular focus on end of life issues and bereavement. Recent studies include hospice at home, end of life care and carers needs assessment.

**Brenda Roe** is a professor of Health Research, EPRC Edge Hill University and Honorary Fellow, PSSRU, University of Manchester. Her research interests include health and social care, organisation and service delivery and people’s experiences of living with chronic conditions in a variety of institutional, community and care home settings. Recent projects have included incontinence, falls, dementia, quality of life, end of life and partnerships with older people.