REVIEW

The course of neuropsychiatric symptoms in community-dwelling patients with dementia: a systematic review

Petra Borsje,1,2 Roland B. Wetzel,1 Peter L. Lucassen,1 Anne Margriet Pot3,4,5 and Raymond T. Koopmans1,6

1Department of Primary and Community Care, Radboud University Medical Center, Nijmegen, the Netherlands
2Thebe, Region South-East, Gorle, Tilburg, the Netherlands
3Department of Clinical Psychology, VU University Amsterdam, the Netherlands
4Netherlands Institute of Mental Health and Addicition, Utrecht, the Netherlands
5School of Psychology, The University of Queensland, Australia
6Joachim en Anna, Center for specialized geriatric care, Nijmegen

ABSTRACT

Background: Neuropsychiatric symptoms (NPS) often occur in patients with dementia. Understanding the course of NPS in dementia is important for healthcare professionals for psycho-educational purposes and adequate and timely interventions to prevent or diminish NPS as much as possible.

Methods: We conducted a systematic literature search in several electronic databases. We combined search strings for the terms dementia, community-dwelling, cohort studies and NPS. Screening titles and abstracts, assessing the methodological quality and data-extraction were independently conducted by at least two authors.

Results: This literature search revealed 6605 unique records of which 23 studies were included in data synthesis. In total 7184 patients participated in the included studies with a mean number of 312. Sixty percent of the participants were female and the mean age of all participants was 74.8 years. Follow-up varied between 1 and 6 years; in 17 studies loss to follow-up was less than 20% per year. NPS are highly prevalent, incident and persistent although frequency parameters vary considerably across studies. Delusions/delusional misidentification, wandering/agitation, aberrant motor behavior/motor hyperactivity and apathy are the most common NPS. For hallucinations, delusions/delusional misidentification, paranoia, aggression, wandering/agitation, aberrant motor behavior/motor hyperactivity, disinhibition, apathy, and sleep disturbance increasing trends in point prevalence rates have been found.

Conclusions: NPS in community-dwelling patients are frequent and persistent. The increasing trends of several NPS in the course of dementia require a preventive approach of professional caretakers. For such an approach, a timely diagnosis and adequate professional support to prevent or diminish these problems is necessary.

Key words: dementia, neuropsychiatric symptoms, community-dwelling, systematic review

Introduction

Dementia is a chronic and progressive disorder with great impact on people with dementia and their family members (WHO, 2012). NPS, also termed behavioral and psychological symptoms of dementia (BPSD), frequently complicate the course of dementia. Examples of NPS are psychosis (delusions and hallucinations), depressive mood, anxiety, irritability/lability, apathy, euphoria, disinhibition, agitation/aggression, aberrant motor activity, sleep disturbance, and eating disorder. Studies from various countries reported NPS prevalence rates in community-dwelling people that ranged from 61% to 96% (Chow et al., 2002; Lyketsos et al., 2002; Ikeda et al., 2004; Craig et al., 2005). NPS result in lower quality of life for both the people with dementia and their caregivers and affect the quality of the relationship with the caregivers (de Vugt et al., 2003; Shin et al., 2005).
Nursing home admission is predicted by NPS, severity of cognitive impairment, Alzheimer’s dementia, high rates of functional dependence, and depressive symptoms (Gaugler et al., 2009). Major depression is a predictor of early institutionalization in the first year following the dementia diagnosis (Dorenlot et al., 2005). The future severity of NPS is predicted by the baseline severity of NPS, stage of dementia and use of support services (Asada et al., 2000).

For people with dementia and their informal caregivers, as well as general practitioners and the other professionals involved in long-term care, it is important to understand the course of NPS. If we are able to recognize patients at risk of persistent NPS, we can develop individual approaches in the different stages of dementia for both patients with dementia and their professional and informal caregivers. Knowledge on NPS in dementia is important for psycho-educational purposes and timely interventions to prevent or diminish NPS as much as possible. The aim of this systematic review is to study the prevalence and course of NPS in community-dwelling patients with dementia.

**Methods**

We performed a systematic review of prospective cohort studies according to the guidelines of the Cochrane Collaboration and the PRISMA-Statement (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) (Moher et al., 2009; Higgins and Green, 2011). Conform the PROGnosis RESearch Strategy (PROGRESS) we aimed for studies that can be classified as fundamental prognosis research (Hemingway et al., 2013). The following steps were described in a predefined research protocol: (1) inclusion criteria; (2) exclusion criteria; (3) search methods for identification of studies; (4) data extraction; (5) assessment of methodological quality; (6) data synthesis.

**Inclusion criteria**

*Types of studies.* This review included prospective cohort studies.

*Types of participants.* Patients with dementia.

*Setting.* Primary care or community dwelling patients.

*Study size.* At least 25 or more patients at baseline.

*Follow-up.* Three months or more.

*Types of outcome measures.* Neuropsychiatric symptoms.

**Exclusion criteria**

*Types of studies.* Case-studies, case-control studies, clinical trials, cross-sectional studies and trend studies (repeated cross-sectional studies).

*Types of participants.* Caregivers, patients with mild cognitive impairment (MCI).

*Setting.* Assisted living facilities, chronic care institutions, home of the aged, housing for the elderly, intermediate care facilities, nursing home and residential care.

*Search method for identification of studies*

On November 27th, 2012 we conducted an electronic search in the databases PubMed, EMBASE, CINAHL, PsycINFO and the Cochrane Library for all studies that were published until that date. We modified a previously used search strategy for residents with dementia in long-term care institutions for this review on patients with dementia in primary care (Zuidema et al., 2007; Wetzels et al., 2010). We combined search strings for dementia, primary care, cohort studies and NPS with the Boolean operator AND. An overview of the terms used in the computerized search strategy as performed in PubMed is presented in supplementary (see Table S1 available as supplementary material attached to the electronic version of this paper at www.journals.cambridge.org/jid_IPG). The search strategy was adapted for the other four databases to fit database-specific features. The reference lists of selected papers and previous reviews on the course of NPS in primary care were searched for papers not identified by the initial search. The results of the 5 databases were aggregated and duplicates were deleted. No books or dissertations were included in this review. There were no limitations regarding the language of the publication.

*Selection method*

Two authors (PB and RBW) independently screened titles and abstracts to identify eligible papers. When there was insufficient information to evaluate the inclusion and exclusion criteria, we retrieved the full text paper. We excluded all studies that clearly did not meet all inclusion criteria or that met at least one of the exclusion criteria. Inter-rater agreement about inclusion and exclusion based on titles and abstracts was measured and reported as Cohen’s $\kappa$ (Landis and Koch, 1977). Subsequently, two authors (PB and RBW) independently reviewed the full publications of the remaining papers. We discussed disagreements in consensus meetings. All discussions led to consensus about inclusion. When needed we corresponded with co-authors of
this paper for further information to clarify study eligibility.

**Data extraction and assessment of methodological quality**

At least two reviewers (PB, RBW and/or PLL) independently extracted the information from the selected publications by using standardized and pre-tested data-extraction forms. The extracted information involved data on study population, diagnostic criteria, inclusion and exclusion criteria, setting, type of prognostic factors, duration of follow-up, outcomes, and data on methodological quality. In a case of disagreement, we reached consensus after discussion with all three reviewers.

For assessing the methodological quality of the included studies we used a standardized checklist of predefined criteria (see Table 1), which has been used in previous prognostic reviews (Olde Hartman et al., 2009) and is based on theoretical considerations and methodological aspects described earlier (Hudak et al., 1998; Altman, 2001; Hayden et al., 2006). Two authors (PB and RBW) tested the quality assessment checklist in a pilot assessment. Each criterion was scored positive (+, design or conduct adequate), negative (−, design or conduct inadequate), or unclear (?, insufficient information). The total quality score is expressed as the sum of all criteria that are scored positive. The maximum quality score is 21 and we calculated the quality of a study as the percentage of the maximum score. We discussed disagreements in the scoring of quality items in consensus meetings and categorized the quality criteria into four major forms of bias: selection bias, completeness of follow-up, information bias, and confounding. For judging selection bias, quality criteria description of inception cohort, study population, relevant inclusion and exclusion criteria, definition of dementia and NPS, response rate ≥75% and information about non-responders versus responders were used. The quality criteria loss to follow-up <20% per year and information about completers versus those lost to follow-up/dropouts were used to judge completeness of follow-up. Furthermore, the quality criteria standardized assessment of symptoms and functional outcome, as well as potential prognostic factors were used to judge information bias; the quality criteria description of possible treatment in cohort and appropriate univariate crude estimates and multivariate analysis techniques were used to judge confounding. Finally, the quality criteria number of participants in study population ≥100 at baseline, follow-up of at least 12 months, prospective data collection, clinically relevant outcome measures, frequencies of most important outcome measures and prognostic factors presented and influence of prognostic factors presented were used to judge descriptive items. We defined studies with a quality score of 60% or higher as studies with high quality (Kuijpers et al., 2004; Olde Hartman et al., 2009).

**Data synthesis**

The following main study characteristics were extracted from the papers: setting/country, number enrolled in cohort, criteria for diagnosis (dementia and NPS), duration of follow-up (years and range), loss to follow-up (number and%), gender and age (years ± S.D. and range) at baseline, living situation of participants at baseline, number of assessments and presentation of results for number of patients per assessment (PPA) or completers.

Information on the course of NPS is presented in three subgroups according to the factor analysis of Aalten (Aalten et al., 2003). The three subgroups we present are a psychotic subgroup including hallucinations, delusions, delusional misidentification and paranoia, a hyperactivity subgroup including agitation, aggression, euphoria, disinhibition, irritability and aberrant motor behavior and finally an affective subgroup including depression, anxiety, apathy, night-time behavior disturbances and eating abnormalities. Data are presented as point and cumulative prevalence, (cumulative) incidence, persistence and resolution per assessment. Point prevalence is defined as the proportion of patients with specific NPS at each assessment. The cumulative prevalence is defined as the proportion of patients developing a specific NPS on at least one assessment over the follow-up period including baseline assessment. Incidence is defined as the proportion of patients who develop a specific NPS at one assessment but did not show the symptom on the preceding assessment. The cumulative incidence is defined as the proportion of patients who are symptom free at baseline, but develop a specific NPS at next assessments. A symptom is persistent if it is present on at least two subsequent assessments, regardless of time of first manifestation of the symptom. Resolution is defined as the proportion of patients who showed a specific NPS at one assessment but not at the next assessment and is displayed for each successive assessment (Aalten et al., 2003). Not statistically tested increasing or decreasing changes are presented as trends. In case of significant changes p-values are presented.
Table 1. Explanation of the criteria for assessing the methodological quality.

A. Description of inception cohort.
Positive if it is described in what setting the participants were recruited (i.e. general population, patients attending the general practitioner, inpatient or outpatient setting).

B. Description of study population.
Positive if it is described which participants from the inception cohort are recruited and if age and sex are described.

C. Description of relevant inclusion and exclusion criteria.
Positive if it is described how participants were identified with dementia.

+ = Dementia diagnosed by standardized diagnostic interview and/or assessment scales
− = Dementia not diagnosed by standardized diagnostic interview (including DSM) or assessment scale
? = Not clear

D. Definition of dementia and neuropsychiatric symptoms.
Positive if the definition of dementia and neuropsychiatric symptoms is described.

E. Number of participants in study population ≥ 100.
Positive if the number of participants with dementia in the study population was at least 100 at baseline.

F. Response rate ≥ 75%.
Positive if response rate is at least 75%. Response rate: the number of patients in the study population at baseline, divided by the number of participants in the inception cohort.

G. Information about non-responders versus Responders.
Positive if demographic or clinical information (such as age and sex) was presented for responders and non-responders, or if there was no selective response, or no non-response.

H. Follow-up of at least 12 months.
Positive if the follow-up period was at least 12 months.

I. Loss to follow-up < 20% per year.
Positive if mean number of patients with dementia is less than 20% per year. Loss to follow-up: the number of patients in the study population at baseline minus the number of patients at the main NPS outcome measure for each year, divided by the number of patients in the study population at baseline.

J. Information about completers versus those lost to follow-up/dropouts.
Positive if demographic or clinical information (such as age and gender, disease characteristics, and other potential prognostic predictors) was presented for completers with dementia and those lost to follow-up at the main moment of outcome measurement, or if there was no selective loss to follow-up, or no loss to follow-up.

K. Prospective data collection.
Positive if main outcome measures on potential prognostic predictors were collected prospectively.

L. Description of possible treatment in cohort.
Positive if treatment subsequent to inclusion in cohort is fully described or standardized. Also positive if no treatment is given.

+ = treatment/multivariate correction for treatment in analysis, or no treatment given.
− = different treatment regimens, not clear how outcome is influenced by it.
? = not clear if any treatment is given.

M. Clinically relevant outcome measures.
Positive if at least one clinically relevant outcome measures is presented.

N. Standardized assessment of symptom outcome.
Positive if standardized questionnaires or objective outcome measurements of NPS were used for each follow-up measurement.

O. Standardized assessment of functional outcome.
Positive if standardized questionnaires or objective outcome measurements were used for each follow-up measurement.

P. Standardized assessment of potential prognostic factors.
Positive if standardized questionnaires or objective measurements of potential prognostic factors were used at baseline.

Q. Appropriate univariate crude estimates.
Positive if separate univariate (repeated measures) analysis of variance was calculated for each dependent measure.

R. Appropriate multivariate analysis techniques.
Positive if multivariate (repeated measures) analysis of variance was calculated for changes among the dependent measures occurring during the follow-up interval.

S. Frequencies of most important outcome measures presented.
Positive if frequency, percentage or mean, median (interquartile range), and standard deviation/confidence intervals are reported of the most important outcome measures.

T. Frequencies of most important prognostic factors presented.
Positive if:
- a. frequency of percentage is reported, or
- b. mean and standard deviation or standard error are reported, or
- c. median and interquartile range are reported, or
- d. if the influence of each separate factor is reported

U. Influence of prognostic factors presented.
Positive if the influence of each separate prognostic factor on the natural course of NPS is presented.
Results

Search results and study selection

The process of selecting publications for the review is illustrated in Figure 1. We retrieved a total of 9167 publications from searches of the various electronic databases (PubMed 2370, EMBASE 2428, CINAHL 2899, PsychINFO 1123 and the Cochrane Library 347) and 15 through hand-search of reference lists of other studies. After screening the titles and abstracts, 53 publications seemed eligible according to the inclusion and exclusion criteria. The inter-observer agreement (unweighted $\kappa$) for inclusion between the two reviewers (PB, RBW) for screening titles and abstracts was $\kappa = 0.60$ (95% CI: 0.43–0.76), which is considered “moderate” agreement (Landis and Koch, 1977). Proportion of agreement was 0.9965 (95% CI: 0.9947–0.9977). After assessing the full publications, 23 papers were definitively included in our review (Rosen and Zubenko, 1991; Forstl et al., 1993; McShane et al., 1995; Swearer et al., 1996; Ballard et al., 1997; Devanand et al., 1997; Juva et al., 1997; Keene and Hope, 1998; McShane et al., 1998; Asada et al., 1999; Haupt et al., 2000; Paulsen et al., 2000; Wilson et al., 2000; Ballard et al., 2001; Li et al., 2001; Holtzer et al., 2003; Aalten et al., 2005; Cortes et al., 2005; Holtzer et al., 2005; Zhu et al., 2006; Froelich et al., 2009; Kunik et al., 2010; Gonfrier et al., 2012). Major reasons for excluding publications were retrospective data analysis, studies on samples of informal caregivers and major or unclear part of participants living in institutions.

Study characteristics

The quality score of the 23 studies ranged from 52% to 86% (see Table 2). Twenty of these have a score of 60% or higher. Selection bias was present in all studies. Response rate was given in only two studies and information about responders versus non-responders was given in only one other study. Confounding was present in 19 studies and in
Table 2. Results of the methodological quality assessment of prospective cohort studies on the course of NPS.

<table>
<thead>
<tr>
<th>QUALITY CRITERIA</th>
<th>COMPLETE</th>
<th>INFORMATION</th>
<th>CONFOUNDERING</th>
<th>DESCRIPTIVE ITEMS</th>
<th>QUALITY SCORE</th>
<th>SCORE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SELECTION BIAS</td>
<td>FOLLOW-UP</td>
<td>BIAS</td>
<td></td>
<td></td>
<td>(%)</td>
</tr>
<tr>
<td>Holtzer, et al. (2005)*</td>
<td>+ + + +</td>
<td>?</td>
<td>−</td>
<td>−</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Devanand, et al. (1997)*</td>
<td>+ + + +</td>
<td>?</td>
<td>−</td>
<td>−</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Kunik, et al. (2010)</td>
<td>+ + + +</td>
<td>−</td>
<td>+</td>
<td>−</td>
<td>+</td>
<td>−</td>
</tr>
<tr>
<td>Holtz, et al. (2003)*</td>
<td>+ + + +</td>
<td>?</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Paulsen, et al. (2000)</td>
<td>+ + + +</td>
<td>?</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>−</td>
</tr>
<tr>
<td>Wilson, et al. (2000)</td>
<td>+ + + +</td>
<td>+</td>
<td>−</td>
<td>+</td>
<td>−</td>
<td>+</td>
</tr>
<tr>
<td>Ballard, et al. (1997)</td>
<td>+ + + +</td>
<td>+</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>−</td>
</tr>
<tr>
<td>Swearer, et al. (1996)</td>
<td>+ + + +</td>
<td>?</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Rosen and Zienbenko (1991)</td>
<td>+ + + +</td>
<td>?</td>
<td>−</td>
<td>−</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Keene and Hope (1996)*</td>
<td>+ + + +</td>
<td>?</td>
<td>−</td>
<td>−</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>McShane, et al. (1998)*</td>
<td>+ + + +</td>
<td>?</td>
<td>−</td>
<td>+</td>
<td>−</td>
<td>+</td>
</tr>
<tr>
<td>McShane, et al. (1995)</td>
<td>− + + +</td>
<td>−</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>−</td>
</tr>
<tr>
<td>Lis, et al. (2001)</td>
<td>+ + + +</td>
<td>?</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Aalen, et al. (2005)</td>
<td>+ + + +</td>
<td>?</td>
<td>−</td>
<td>−</td>
<td>+</td>
<td>−</td>
</tr>
<tr>
<td>Cortes, et al. (2005)*</td>
<td>+ + + +</td>
<td>?</td>
<td>−</td>
<td>−</td>
<td>+</td>
<td>−</td>
</tr>
</tbody>
</table>

+, positive (design or conduct adequate); −, negative (design or conduct inadequate); ?, unclear (insufficient information).

*Total ‘+’.

aPredictors.

bPredictors 2.

cKeene/McShane.

dREAL.FR.
17 studies information bias was present. Follow-up in all studies was at least 12 months. In 17 studies loss to follow-up was less than 20% per year. After close inspection no direct association was found between setting/country, number enrolled in cohort, criteria for diagnosis, duration of follow-up, loss to follow-up, gender, age, living situation, year of publication and the total quality score.

Eight studies are performed in the United States of America (USA), five in the United Kingdom, two in Germany and two in France, one study in Japan, one study in Finland and one study in the Netherlands (see Table 3). Two studies are performed in 3 sites (USA, France, Greece) and 1 study is performed in 12 European countries. In total 7184 patients participated in the included studies. The mean number of study participants was 312 and the median number 170 (range 30–2288).

Dementia was diagnosed according to several diagnostic criteria and in one of them according to histopathological criteria (Rosen and Zubenko, 1991). NPS were assessed by using 15 different instruments. Five studies used the Columbia Scale for Psychopathology in Alzheimer’s disease (CUSPAD) of which three papers described the results of the Predictors study, four used the Neuropsychiatric Inventory (NPI) of which two papers described the results of the Reseau sur la Maladie d’Alzheimer Francais (REAL.FR) study, three studies used the Present Behavioural Examinations (PBE) of which two papers described the results of the same study and two studies used the Diagnostic and Statistical Manual of Mental Disorders (DSM) criteria for delusions and hallucinations. The Cohen Mansfield Agitation Inventory (CMAI), Cornell Scale for Depression in Dementia (CSDD), DSM criteria for depression, DSM criteria for psychosis, Burns’ Symptom checklist, Caretaker Obstreperous - Behavior Rating Assessment (COBRA) scale, Behavioural Abnormalities in Alzheimer's Disease (BEHAVE-AD), Troublesome Behavior Scale (TBS), and Hamilton Depression Rating Scale (HADRS) were each used in one study. One study used a semi-standardized carers’ interview and one study used a structured clinical interview.

Follow-up varied between 1 and 6 years with a mean of 3 years and median of 3.5 years. Twenty-two studies reported data on loss to follow-up: a total of 3024 patients (44%) were lost to follow-up with a mean of 38% (range 0–85.3%).

Nineteen studies reported data on gender. In these 19 studies 2256 of the participants were male (40%) with a mean of 119 per study and median 75 (range 13–857) and 3376 of the participants were female (60%) with a mean of 178 per study and median 102 (range 10–1431). In the 20 studies reporting data on age the mean age was 74.8 years (range 68.8–79.9). Two of these studies reported data on age per diagnosis. In one study for dementia with Lewy bodies (LBD) 76.5 ± 7.9 years and for Alzheimer’s disease (AD) 81.1 ± 6.6 years. In the other study for AD 76.5 ± 7.1 years and for vascular dementia (VaD) 71.4 ± 8.1 years.

Three studies reported an outpatient setting, but did not give specific information on the living situation of the participants at baseline (Haupt et al., 2000; Li et al., 2001; Holtzer et al., 2005). Two studies reported that a small part of the participants were not living at home. One study reported that 91.1% of the participants lived at home, 6.8% in nursing home, 1.3% in retirement home, 0.9% other living situations (Devanand et al., 1997). The other study reported that 5.9% of the participants were recruited from a long-term care facility (Holtzer et al., 2003).

After close inspection no direct association was found between total quality score and the frequency parameters of the studies.

Course of NPS in patient with dementia in primary care

In 22 studies, 2–12 assessments were conducted (see Table 4). In one study the number of assessments was not given (Kunik et al., 2010). Twelve studies presented data of PPA, ten studies presented data of completers and for one study this was not clear. In one study the results were presented in figures per stage of dementia using the Clinical Dementia Rating scale (CDR) and not in numbers (Asada et al., 1999). This study concluded that the patterns of NPS change depended on the baseline severity of AD. The NPS frequencies peaked in the middle stage (CDR 2) and followed a downward trend thereafter. Two studies reported on the psychotic, hyperactivity and affective subgroup as a whole, as well as on individual symptoms (Aalten et al., 2005; Gonfrier et al., 2012). The study of Aalten et al. reported point and cumulative prevalence rates, (cumulative) incidence and persistence rates on all individual NPI symptoms as well as on the subgroups (see Table 4).

Any symptom

Ten studies reported on NPI total scores or on NPS in general without symptom specification of which seven studies reported on PPA and three studies on completers (Forstl et al., 1993; Swearer et al., 1996; Ballard et al., 1997; Devanand et al., 1997; Juva et al., 1997; Aalten et al., 2005; Cortes et al., 2005; Froelich et al., 2009; Gonfrier et al., 2012; Zhu et al., 2006). The point prevalence rates
<table>
<thead>
<tr>
<th>FIRST AUTHOR</th>
<th>STUDY QUALITY (%)</th>
<th>SETTING/COUNTRY</th>
<th>NUMBER ENROLLED IN COHORT</th>
<th>CRITERIA FOR DIAGNOSIS</th>
<th>DURATION OF FOLLOW-UP [YEARS (RANGE)]</th>
<th>LOSS TO FOLLOW-UP [N, (%)]</th>
<th>GENDER (M/F) AND AGE (YEARS ± S.D.; RANGE) AT BASELINE</th>
<th>LIVING SITUATION OF PARTICIPANTS AT BASELINE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Holtzer, et al. (2005)a</td>
<td>86</td>
<td>Outpatient/ USA (3 sites), France, Greece</td>
<td>536</td>
<td>DSM-III-R, NINCDS-ADRDA</td>
<td>CUSPAD</td>
<td>5.0</td>
<td>406 (75.7)</td>
<td>220/316; 74.0 ± 8.7</td>
</tr>
<tr>
<td>Devanand, et al. (1997)a</td>
<td>86</td>
<td>Outpatient/ USA (3 sites)</td>
<td>235</td>
<td>DSM-III-R, NINCDS-ADRDA</td>
<td>CUSPAD</td>
<td>3.0 ± 2.5</td>
<td>98 (41.7)</td>
<td>96/139; 73.1 ± 8.9</td>
</tr>
<tr>
<td>Kunik, et al. (2010)</td>
<td>81</td>
<td>Outpatient/ USA</td>
<td>215</td>
<td>ICD-9-CM</td>
<td>CMAI</td>
<td>2.0</td>
<td>16 (7.4)</td>
<td>205/10; 76 ± 6.2</td>
</tr>
<tr>
<td>Holtzer, et al. (2003)b</td>
<td>81</td>
<td>Outpatient/ USA (3 sites)</td>
<td>236</td>
<td>DSM-III-R, NINCDS-ADRDA</td>
<td>CUSPAD</td>
<td>5.0</td>
<td>134 (56.8)</td>
<td>97/139; 72.7 ± 9.2</td>
</tr>
<tr>
<td>Ballard, et al. (2001)</td>
<td>81</td>
<td>Outpatient/ UK</td>
<td>244</td>
<td>CAMCOG</td>
<td>CUSPAD, CSDD, DSM-II-R depression</td>
<td>1.0</td>
<td>20 (8.2)</td>
<td>LBD: 36/46; AD: 30/93 LBD: 76.5 ± 7.9; AD 81.1 ± 6.6</td>
</tr>
<tr>
<td>Paulsen, et al. (2000)</td>
<td>81</td>
<td>Outpatient/ USA</td>
<td>329</td>
<td>DSM-III, NINCDS-ADRDA</td>
<td>DSM-III Psychosis</td>
<td>5.0</td>
<td>Not given</td>
<td>Never psychotic (n = 194) 26/168; 72.4 ± 6.9; Psychotic at baseline (n = 75) 41/34; 73.4 ± 7.7; Psychotic at follow-up visit (n = 60) 31/29; 72.1 ± 6.4</td>
</tr>
<tr>
<td>Wilson, et al. (2000)</td>
<td>81</td>
<td>Outpatient/ USA</td>
<td>410</td>
<td>NINCDS-ADRDA</td>
<td>DSM-III-R subtypes of delusions</td>
<td>4.0</td>
<td>141 (34.4)</td>
<td>136/274; 75.5 ± 7.3 (45 – 95)</td>
</tr>
<tr>
<td>FIRST AUTHOR</td>
<td>STUDY QUALITY (%)</td>
<td>SETTING/COUNTRY</td>
<td>NUMBER ENROLLED IN COHORT</td>
<td>CRITERIA FOR DIAGNOSIS</td>
<td>DURATION OF FOLLOW-UP [YEARS (RANGE)]</td>
<td>LOSS TO FOLLOW-UP [N, (%)]</td>
<td>GENDER (M/F) AND AGE (YEARS ± S.D.; RANGE) AT BASELINE</td>
<td>LIVING SITUATION OF PARTICIPANTS AT BASELINE</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------------</td>
<td>-----------------</td>
<td>---------------------------</td>
<td>------------------------</td>
<td>----------------------------------------</td>
<td>----------------------------</td>
<td>-----------------------------------------------------</td>
<td>---------------------------------------------</td>
</tr>
<tr>
<td>Ballard, et al. (1997)</td>
<td>81</td>
<td>Outpatient/ UK</td>
<td>124</td>
<td>DSM-III-R, NINCDS-ADRDA Burns’s Symptom Checklist</td>
<td>1.0</td>
<td>37 (29.8)</td>
<td>38/102; 79.9</td>
<td>Consecutive referrals to old-age psychiatry services</td>
</tr>
<tr>
<td>Swearer, et al. (1996)</td>
<td>76</td>
<td>Outpatient/ USA</td>
<td>30</td>
<td>NINCDS-ADRDA COBRA Scale</td>
<td>17.83 ± 9.9 months (range 0.5 - 3)</td>
<td>16 (53.3)</td>
<td>13/17; 72.7 ± 6.5</td>
<td>Community-dwelling</td>
</tr>
<tr>
<td>Haupt, et al. (2000)</td>
<td>71</td>
<td>Outpatient/ Germany</td>
<td>90</td>
<td>NINCDS-ADRDA BEHAVE-AD</td>
<td>2.0</td>
<td>30 (33.3)</td>
<td>Not given</td>
<td>Not given</td>
</tr>
<tr>
<td>Förstl, et al. (1993)</td>
<td>71</td>
<td>Psychiatric University hospital/ Germany</td>
<td>50</td>
<td>NINCDS-ADRDA semi-standardized carers’ interview</td>
<td>2.0</td>
<td>7 (14.0)</td>
<td>20/30; 68.8 (49 – 92)</td>
<td>Living in the community independently or with their families</td>
</tr>
<tr>
<td>Rosen and Zubenko (1991)</td>
<td>71</td>
<td>Outpatient/ USA</td>
<td>32</td>
<td>Histopathological criteria for Alzheimer’s disease DSM-III Delusions and hallucinations</td>
<td>6.0</td>
<td>0 (0)</td>
<td>17/15; 70.3 ± 7.9</td>
<td>Ambulatory patients, living in the community</td>
</tr>
<tr>
<td>Asada, et al. (1999)</td>
<td>67</td>
<td>Outpatient/ voluntary/service providers/ Japan</td>
<td>103</td>
<td>NINCDS-ADRDA TBS</td>
<td>5.0</td>
<td>76 (73.8)</td>
<td>36/67; 79.4 ± 8.7</td>
<td>Living in a private residence with responsible caregivers</td>
</tr>
<tr>
<td>Keene and Hope (1998)</td>
<td>67</td>
<td>Outpatient/ UK</td>
<td>104</td>
<td>CAMDEX, NINCDS-ADRDA PBE</td>
<td>1.0</td>
<td>5 (4.8)</td>
<td>Not given</td>
<td>Patients living at home</td>
</tr>
<tr>
<td>McShane, et al. (1998)</td>
<td>67</td>
<td>Outpatient/ UK</td>
<td>104</td>
<td>DSM-III-R, CERAD PBE</td>
<td>4.0</td>
<td>18 (17.3)</td>
<td>43/43; 77 (IQR 8)</td>
<td>Patients living at home</td>
</tr>
<tr>
<td>Juva, et al. (1997)</td>
<td>67</td>
<td>Outpatient/ Finland</td>
<td>100</td>
<td>DSM-III structured clinical interview</td>
<td>1.0</td>
<td>9 (9.0)</td>
<td>48/52; 69.7 (48.3 – 89.0)</td>
<td>Living at home at first interview</td>
</tr>
<tr>
<td>Gonfrier, et al. (2012)</td>
<td>62</td>
<td>Outpatient/ France (multicenter; 16 sites)</td>
<td>686</td>
<td>DSM-IV, NINCDS-ADRDA NPI</td>
<td>4.0</td>
<td>535 (78.0)</td>
<td>4-year follow-up (n = 151) 40/111; 76.1 ± 6.4 Others (n = 479) 150/329; 78.4 ± 6.8</td>
<td>Home with spouse 403 (58.7%), home alone 183 (26.7%), home with family 80 (11.7%), group home/other 20 (2.9%)</td>
</tr>
<tr>
<td>FIRST AUTHOR</td>
<td>STUDY QUALITY (%)</td>
<td>SETTING/COUNTRY</td>
<td>NUMBER ENROLLED IN COHORT</td>
<td>CRITERIA FOR DIAGNOSIS</td>
<td>DURATION OF FOLLOW-UP [YEARS (RANGE)]</td>
<td>LOSS TO FOLLOW-UP [N, (%)]</td>
<td>GENDER (M/F) AND AGE (YEARS ± S.D.; RANGE) AT BASELINE</td>
<td>LIVING SITUATION OF PARTICIPANTS AT BASELINE</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------</td>
<td>----------------</td>
<td>---------------------------</td>
<td>-----------------------</td>
<td>--------------------------------------</td>
<td>--------------------------</td>
<td>------------------------------------------------------</td>
<td>---------------------------------------------</td>
</tr>
<tr>
<td>Froehlich, et al. (2009)</td>
<td>62</td>
<td>Outpatient/ 12 European countries</td>
<td>2288</td>
<td>DSM-IV</td>
<td>NPI</td>
<td>2.0</td>
<td>906 (39.6)</td>
<td>857/1431; 77.0 (30–100)</td>
</tr>
<tr>
<td>Zhu, et al. (2006)b</td>
<td>62</td>
<td>Outpatient/ USA (3 sites)</td>
<td>170</td>
<td>DSM-III-R, NINCDS-ADRDA</td>
<td>CUSPAD</td>
<td>4.0 (median 2.5; maximum 6.0]</td>
<td>145 (85.3)</td>
<td>76/94; 75.0 ± 7.6</td>
</tr>
<tr>
<td>McShane, et al. (1995)</td>
<td>62</td>
<td>Not given/UK</td>
<td>98</td>
<td>DSM-III-R, CERAD</td>
<td>PBE</td>
<td>5.0</td>
<td>57 (58.2)</td>
<td>Not given</td>
</tr>
<tr>
<td>Li, et al. (2001)</td>
<td>57</td>
<td>Outpatient/ USA</td>
<td>108</td>
<td>DSM-III-R</td>
<td>HDRS</td>
<td>3.5 (range 0.8 – 7.8)</td>
<td>74 (68.5)</td>
<td>AD 34/37; 76.5 ± 7.1 VaD 24/13; 71.4 ± 8.1</td>
</tr>
<tr>
<td>Aalten, et al. (2005)</td>
<td>52</td>
<td>Outpatient/ the Netherlands (2 sites)</td>
<td>199</td>
<td>DSM-IV, NINCDS-ADRDA, NINCDS-AIREN</td>
<td>NPI</td>
<td>2.0</td>
<td>99 (49.7)</td>
<td>83/116; 76.4 ± 8.0 (53 – 96)</td>
</tr>
<tr>
<td>Cortes, et al. (2005)d</td>
<td>52</td>
<td>Outpatient/ France (multicenter; 16 sites)</td>
<td>693</td>
<td>DSM-IV, NINCDS-ADRDA</td>
<td>NPI</td>
<td>4.0</td>
<td>195 (28.1)</td>
<td>Gender not given No discontinuation (n = 544); 77.2 ± 6.9 Discontinuation (n = 121); 79.4 ± 6.9</td>
</tr>
</tbody>
</table>

**AD**: Alzheimer’s disease; **BEHAVE-AD**: behavioural abnormalities in Alzheimer’s disease rating scale; **CERAD**: consortium to establish a registry for Alzheimer’s disease; **CMAI**: cohen-mansfield agitation inventory; **COBRA**: caretaker obstreperous behavior rating assessment scale; **CSDD**: cornell scale for depression in dementia; **CUSPAD**: columbia scale for psychopathology in Alzheimer’s disease; **DSM**: diagnostic and statistical manual of mental disorders; **HDRS**: hamilton depression rating scale; **IQR**: interquartile range; **LBD**: dementia with Lewy bodies; **NINCDS-ADRDA**: national institute of neurological and communicative disorders and stroke - Alzheimer’s disease and related disorders association; **NPI**: neuropsychiatric inventory; **PBE**: present behavioural examination; **TBS**: troublesome behavior scale; **VaD**: vascular dementia.

aPredictors.
bPredictors 2.
cKeene/McShane.
dREAL.FR.
Table 4. Course of NPS – first revision.

<table>
<thead>
<tr>
<th>FIRST AUTHOR</th>
<th>STUDY QUALITY (%)</th>
<th>STUDY COHORT</th>
<th>OUTCOME MEASURES (N Y, N ASM) COMPLETERS VERSUS PATIENTS PER ASM</th>
<th>PSYCHOTIC SUBGROUP (HALLUCINATIONS, DELUSIONS, DELUSIONAL MISIDENTIFICATION, PARANOIA)</th>
<th>HYPERACTIVITY SUBGROUP (AGITATION, AGGRESSIVE, EUPHORIA, DISINHIBITION, IRRITABILITY, ABERRANT MOTOR BEHAVIOR)</th>
<th>AFFECTIVE SUBGROUP (DEPRESSION, ANXIETY, APATHY, NIGHT-TIME BEHAVIOR DISTURBANCES, EATING ABNORMALITIES)</th>
<th>ANY SYMPTOM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Holtzer, et al. (2005)*</td>
<td>86</td>
<td>CUSPAD</td>
<td>(5y fu, 6 asm) Patients per asm</td>
<td>Hallucinations PP: 8%–8%–12%–17%–11%–20%–12% Delusions (any type) (&lt;0.05) PP: 24%–28%–32%–33%–33%–35%–31% ↑</td>
<td>Behavior disturbance (&lt;0.001) PP: 52%–52%–62%–61%–67%–71%–66% ↑ Wandering or agitation (&lt;0.001) PP: 39%–40%–47%–51%–52%–62%–57% ↑ Physical aggression (&lt;0.001) PP: 6%–9%–10%–11%–20%–21%–19% ↑</td>
<td>Depression PP: 40%–42%–41%–39%–28%–24% ↓ Completers PP: 39%–43%–42%–40%–30%–29% ↓ Depressed mood PP: 25%–20%–27%–23%–22%–20%–23%</td>
<td>PP: 64% at baseline 8.5% remained free of all NPS CP: 91.5% CI: 27%</td>
</tr>
<tr>
<td>Devanand, et al. (1997)*</td>
<td>86</td>
<td>CUSPAD</td>
<td>(3y fu, 7 asm) Patients per asm</td>
<td>Hallucinations PP: 8%–16%–16%–17%–11%–20%–12% Delusions (any type) (&lt;0.05) PP: 24%–28%–32%–33%–33%–35%–31% ↑</td>
<td>Behavior disturbance (&lt;0.001) PP: 52%–52%–62%–61%–67%–71%–66% ↑ Wandering or agitation (&lt;0.001) PP: 39%–40%–47%–51%–52%–62%–57% ↑ Physical aggression (&lt;0.001) PP: 6%–9%–10%–11%–20%–21%–19% ↑</td>
<td>Depression PP: 40%–42%–41%–39%–28%–24% ↓ Completers PP: 39%–43%–42%–40%–30%–29% ↓ Depressed mood PP: 25%–20%–27%–23%–22%–20%–23%</td>
<td>PP: 64% at baseline 8.5% remained free of all NPS CP: 91.5% CI: 27%</td>
</tr>
<tr>
<td>Kunik, et al. (2010)</td>
<td>81</td>
<td>CMAI (2y fu)</td>
<td>Patients per asm</td>
<td>Hallucinations PP: 8%–16%–16%–17%–11%–13% Delusions PP: 40%–48%–49%–45%–37%–34% Similar prevalences and changes over time in completers</td>
<td>Behavior disturbance (&lt;0.001) PP: 52%–52%–62%–61%–67%–71%–66% ↑ Wandering or agitation (&lt;0.001) PP: 39%–40%–47%–51%–52%–62%–57% ↑ Physical aggression (&lt;0.001) PP: 6%–9%–10%–11%–20%–21%–19% ↑</td>
<td>Depression PP: 40%–42%–41%–39%–28%–24% ↓ Completers PP: 39%–43%–42%–40%–30%–29% ↓ Depressed mood PP: 25%–20%–27%–23%–22%–20%–23%</td>
<td>PP: 64% at baseline 8.5% remained free of all NPS CP: 91.5% CI: 27%</td>
</tr>
<tr>
<td>Holtzer, et al. (2003)*</td>
<td>81</td>
<td>CUSPAD</td>
<td>(5y fu, 6 asm) Patients per asm</td>
<td>Hallucinations PP: 8%–16%–16%–17%–11%–13% Delusions PP: 40%–48%–49%–45%–37%–34% Similar prevalences and changes over time in completers</td>
<td>Behavior disturbance (&lt;0.001) PP: 52%–52%–62%–61%–67%–71%–66% ↑ Wandering or agitation (&lt;0.001) PP: 39%–40%–47%–51%–52%–62%–57% ↑ Physical aggression (&lt;0.001) PP: 6%–9%–10%–11%–20%–21%–19% ↑</td>
<td>Depression PP: 40%–42%–41%–39%–28%–24% ↓ Completers PP: 39%–43%–42%–40%–30%–29% ↓ Depressed mood PP: 25%–20%–27%–23%–22%–20%–23%</td>
<td>PP: 64% at baseline 8.5% remained free of all NPS CP: 91.5% CI: 27%</td>
</tr>
</tbody>
</table>
### Table 4. Continued

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Ballard, <em>et al.</em> (2001)</td>
<td>81</td>
<td>CUSPAD, CSDD, DSM-II-R depression (1y fu, 2 asm) Completers</td>
<td>Visual hallucinations IN: 16%; 30% (LBD), 13% (AD) PE: 64%; 77% (LBD), 26% (AD) Auditory hallucinations IN: 13%; 28% (LBD), 7% (AD) PE: 42%; 41% (LBD), 45% (AD) Delusions IN: 25%; 30% (LBD), 24% (AD) PE: 42%; 40% (LBD), 44% (AD) Delusional misidentification IN: 16%; 30% (LBD), 12% (AD) PE: 25%; 30% (LBD), 18% (AD)</td>
<td></td>
<td>Depression IN: 8%; 12% (LBD), 6% (AD) PE: 35%; 38% (LBD), 31% (AD)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paulsen, <em>et al.</em> (2000)</td>
<td>81</td>
<td>DSM-III Psychosis (5y fu, 6 asm) Patients Per asm</td>
<td>Hallucinations or delusions CI: 20%–36%–50%–51% (1–4y)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table continued...
### Table 4. Continued

<table>
<thead>
<tr>
<th>First Author</th>
<th>Study Quality (%)</th>
<th>Outcome Measures (N Y, N ASM)</th>
<th>Psychotic Subgroup (Hallucinations, Delusions, Delusional Misidentification, Paranoia)</th>
<th>Hyperactivity Subgroup (Agitation, Aggression, Euphoria, Disinhibition, Irritability, Aberrant Motor Behavior)</th>
<th>Affective Subgroup (Depression, Anxiety, Apathy, Night-Time Behavior, Disturbances, Eating Abnormalities)</th>
<th>Any Symptom</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wilson, et al. (2000)</td>
<td>81</td>
<td>DSM-III-R subtypes of delusions (4y fu, 5 asm) Patients per asm</td>
<td>Hallucinations PP: 41%–40%–34%–31% CP: 70% Delusions PP: 55%–46%–34%–30% ↓ CP: 80%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ballard, et al. (1997)</td>
<td>81</td>
<td>Burns’s Symptom Checklist (1y fu, 2 asm) Completers</td>
<td></td>
<td>Hallucinations IN: 20% RE: 61% Delusions IN: 30% RE: 73% Delusional misidentification: IN: 17% RE: 65%</td>
<td></td>
<td>IN: 47% PE: 28% RE: 53%</td>
</tr>
<tr>
<td>Swearer, et al. (1996)</td>
<td>81</td>
<td>COBRA Scale (mean 1.5y fu, 4 asm) Patients per asm</td>
<td>Hallucinations PP: 13%–10%–21%–31% PP: 0%–7%–0%–0% (severe) Delusions PP: 23%–33%–42%–61% ↑ PP: 7%–10%–5%–15% (severe) Paranoia PP: 20%–27%–37%–46% ↑ PP: 13%–3%–9%–8% (severe) Disordered ideation CP: 70%</td>
<td>Aggressive/assaultive PP: 17–7–26–23% PP: 13–7–9–8% (severe) CP: 40% Mechanical and motor abnormalities PP: 10%–10%–47%–61% ↑ PP: 7%–3%–16%–0% (severe) CP: 50%</td>
<td></td>
<td>CP: 83%</td>
</tr>
</tbody>
</table>
Table 4. Continued

<table>
<thead>
<tr>
<th>FIRST AUTHOR</th>
<th>STUDY QUALITY (%)</th>
<th>OUTCOME MEASURES (N, N ASM) COMPLETERS VERSUS PATIENTS PER ASM</th>
<th>PSYCHOTIC SUBGROUP (HALLUCINATIONS, DELUSIONS, DELUSIONAL MISIDENTIFICATION, PARANOIA)</th>
<th>HYPERACTIVITY SUBGROUP (AGITATION, AGGRESSION, EUPHORIA, DISINHIBITION, IRRITABILITY, ABBERRANT MOTOR BEHAVIOR)</th>
<th>AFFECTIVE SUBGROUP (DEPRESSION, ANXIETY, APATHY, NIGHT-TIME BEHAVIOR, DISTURBANCES, EATING ABNORMALITIES)</th>
<th>ANY SYMPTOM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Haupt, et al. (2000)</td>
<td>76</td>
<td>BEHAVE-AD (2y fu, 3 asm) Completers</td>
<td>Hallucinations CP: 35% PE: 0% RE: 27%</td>
<td>Agitation CP: 100% PE: 67% RE: 2%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Förstl, et al. (1993)</td>
<td>76</td>
<td>semi-standardized carers’ interview (2y fu, 2 asm) Patients per asm</td>
<td>Hallucinations CP: 34% Delusions CP: 46% Tended to be persistent and non-elaborate Delusional misidentification CP: 34%</td>
<td></td>
<td></td>
<td>CP: 62%</td>
</tr>
<tr>
<td>Rosen and Zubenko (1991)</td>
<td>76</td>
<td>DSM-III Delusions and hallucinations (6y fu, 7 asm) Completers</td>
<td>Hallucinations and delusions CP: 47% PE: 87% RE: 13%</td>
<td>Depression CP: 22%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asada, et al. (1999)</td>
<td>67</td>
<td>TBS per CDR stage (5y fu, 6 asm) Completers</td>
<td>Overall, the patterns of change in each of the three factors for each CDR stage group were similar; data given in graphics not in numbers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Keene and Hope (1998)*</td>
<td>67</td>
<td>PBE (&gt;1y fu, 4 monthly asm) Patients per asm</td>
<td></td>
<td>Hyperphagia: CP: 26% Eating change: Hyperphagia CP: 23% Hypophagia CP: 66% Duration: Hyperphagia 16 months Hypophagia: 16 months</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FIRST AUTHOR</td>
<td>STUDY QUALITY (%)</td>
<td>STUDY QUALITY (%)</td>
<td>OUTCOME MEASURES (N Y, N ASM) COMPLETERS VERSUS PATIENTS PER ASM</td>
<td>PSYCHOTIC SUBGROUP (HALLUCINATIONS, DELUSIONS, DELUSIONAL MISIDENTIFICATION, PARANOIA)</td>
<td>HYPERACTIVITY SUBGROUP (AGITATION, AGGRESSION, EUPHORIA, DISINHIBITION, IRRITABILITY, ABERRANT MOTOR BEHAVIOR)</td>
<td>AFFECTIVE SUBGROUP (DEPRESSION, ANXIETY, APATHY, NIGHT-TIME BEHAVIOR DISTURBANCES, EATING ABNORMALITIES)</td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>-------------------</td>
<td>-------------------</td>
<td>-----------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>McShane, et al. (1998)$^c$</td>
<td>67</td>
<td>PBE</td>
<td>(4y fu, 4 monthly asm) Completers</td>
<td>Physical aggression: IN: 36%</td>
<td>Motor hyperactivity: IN: 20%</td>
<td></td>
</tr>
<tr>
<td>Gonfrier, et al. (2012)$^d$</td>
<td>62</td>
<td>NPI</td>
<td>(4y fu, 5 asm) Completers</td>
<td>Psychotic symptoms PP: 31%–34%–37%–49%–36% ↑</td>
<td>Apathy subgroup: ↑ PP: 49%–56%–51%–60%–65% Apathy PP: 43%–63% ↑</td>
<td></td>
</tr>
<tr>
<td>Froehlich, et al. (2009)</td>
<td>62</td>
<td>NPI</td>
<td>(2y fu, 5 asm) Patients per asm</td>
<td>Psychotic symptoms PP: 31%–34%–37%–49%–36% ↑</td>
<td>Depressive symptoms PP: 19%–26%–17%–10%–20% CP: 20%</td>
<td></td>
</tr>
<tr>
<td>Zhu, et al. (2006)$^b$</td>
<td>62</td>
<td>CUSPAD</td>
<td>(mean 2.5y fu, semiannually asm) Patients per asm</td>
<td>Hallucinations CP: 32%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>McShane, et al. (1995)</td>
<td>62</td>
<td>PBE</td>
<td>(5y fu, median 8 times over mean of 2.7y) Completers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First Author</td>
<td>Study Quality (%)</td>
<td>Outcome Measures (N Y, N ASM)</td>
<td>Completers Versus Patients Per ASM</td>
<td>Outcome Subgroup</td>
<td>Hyperactivity Subgroup</td>
<td>Affective Subgroup</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------------</td>
<td>--------------------------------</td>
<td>-----------------------------------</td>
<td>------------------</td>
<td>------------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Li, et al. (2001)</td>
<td>57</td>
<td>HDRS (mean 3.5 y fu, &gt; 3 asm)</td>
<td>Not given</td>
<td>Psychotic subgroup PP: 25%-24%-32%-345-23% CP: 53% CI: 37% PE: 12%-65%-3%-6% (2,3,4 times)</td>
<td>Hyperactivity subgroup PP: 46%-595-65%-64%-64% ↑</td>
<td>Affective subgroup PP: 67%-68%-80%-71%-75%</td>
</tr>
<tr>
<td>Aalten, et al. (2005)</td>
<td>52</td>
<td>NPI (2 y fu, 5 asm)</td>
<td>Patients per asm</td>
<td>Psychotic subgroup PP: 25%-24%-32%-345-23% CP: 53% CI: 37% PE: 12%-65%-3%-6% (2,3,4 times)</td>
<td>Hyperactivity subgroup PP: 46%-595-65%-64%-64% ↑</td>
<td>Affective subgroup PP: 67%-68%-80%-71%-75%</td>
</tr>
<tr>
<td>Cortes, et al. (2005)</td>
<td>52</td>
<td>NPI (1 y fu, 2 asm)</td>
<td>Completers</td>
<td>Psychotic subgroup PP: 25%-24%-32%-345-23% CP: 53% CI: 37% PE: 12%-65%-3%-6% (2,3,4 times)</td>
<td>Hyperactivity subgroup PP: 46%-595-65%-64%-64% ↑</td>
<td>Affective subgroup PP: 67%-68%-80%-71%-75%</td>
</tr>
</tbody>
</table>

Asm: assessments; AD: Alzheimer’s disease; BEHAVE-AD: behavioural abnormalities in Alzheimer’s disease rating scale; CERAD: consortium to establish a registry for Alzheimer’s disease; CI: cumulative incidence; CMAI: cohen-mansfield agitation inventory; COBRA: caretaker obstreperous behavior rating assessment scale; CP: cumulative prevalence; CSDD: cornell scale for depression in dementia; CUSPAD: Columbia scale for psychopathology in Alzheimer’s disease; FL: fluctuating; fu: follow-up; HDRS: Hamilton depression rating scale; IN: incidence; LBD: dementia with Lewy bodies; MS: mean scores; NPI: neuropsychiatric inventory; NPS: neuropsychiatric symptoms; PBE: present behavioral examination; PE: persistence; PP: point prevalence; RE: resolution; TBS: troublesome behavior scale; VaD: vascular dementia; y: year(s).

*Predictors.
*Predictors 2.
Keene/McShane.
REAL.FR.
ranged between 11% and 90% (PPA 11%–88% and completers 42%–90%) and cumulative prevalence rates ranged between 49% and 95% (PPA 49%–95% and completers 66%–88%). Incidence rate was 47% (PPA), cumulative incidence rates ranged between 27% and 74% (PPA), persistence rates ranged between 8% and 65% (PPA) and resolution rate was 53% (PPA). Two studies presented NPI scores per assessment and not frequency parameters and these increased during follow-up (one PPA and one completers) (Cortes et al., 2005; Froelich et al., 2009). One of these two studies (on completers) reported that NPI scores were stable in 10%, worsened in 51% and improved in 38% of the participants (Cortes et al., 2005).

**Psychotic subgroup**

In the two studies (one on PPA and one on completers) that reported on the psychotic subgroup as a whole the point prevalence rates ranged between 13% and 34% (PPA 23%–34% and completers 13%–21%). Cumulative prevalence rate and cumulative incidence rate were 53% and 37%, respectively (PPA) and persistence rates ranged between 3% and 12% (PPA). Point prevalence rates on psychotic symptoms (not specified) presented in one study ranged between 31% and 49% (PPA) (Zhu et al., 2006).

Hallucination point prevalence rates ranged between 0% (severe symptoms) and 43% (PPA) and cumulative prevalence rates ranged between 32% and 70% (PPA 34%–70% and completers 32%–47%). In one study there was a trend of an increasing rate of hallucinations during follow-up (completers) (Gonfrier et al., 2012). One study reported an incidence rate of 20% (completers) (Ballard et al., 1997) and another study (completers) reported incidence rates separately for visual hallucinations 16% (LBD 30% and AD 13%) and for auditory hallucinations 13% (LBD 28% and AD 7%) (Ballard et al., 2001). Persistence rate ranged between 0% and 64% (completers). One study (completers) reported persistence rates separately for visual hallucinations 64% (LBD 77% and AD 26%) and for auditory hallucinations 42% (LBD 41% and AD 45%) (Ballard et al., 2001). Resolution rates ranged between 27% and 61% (completers).

Delusion/delusional misidentification point prevalence rates ranged between 5% (severe symptoms) and 61% (PPA). In two studies (PPA) there was a trend that delusions/delusional misidentification increased during follow-up (Swearer et al., 1996; Kunik et al., 2010), but in one study (PPA) there was a trend that these symptoms decreased (Wilson et al., 2000). Cumulative prevalence rates ranged between 34% and 80% (one study completers 53%). One study (completers) reported an incidence rate of 30% for delusions and 17% for delusional misidentification (Ballard et al., 1997) and another study (completers) reported incidence rates for delusions 25% (LBD 30% and AD 24%) and for delusional misidentification 16% (LBD 30% and AD 12%) (Ballard et al., 2001). Cumulative incidence rate was 34% (PPA). Persistence rate was 0% for delusions in one study (Haupt et al., 2000) and in another study 42% (LBD 40% and AD 44%) for delusions and 25% (LBD 30% and AD 18%) for delusional misidentification (Ballard et al., 2001). One study reported a resolution rate of 57% (Haupt et al., 2000) and another study 73% for delusions and 65% for delusional misidentification (all completers) (Ballard et al., 1997).

For hallucinations and/or delusions together (PPA) cumulative prevalence rate was 47%, persistence rate was 87% and resolution rate was 13% (Rosen and Zubenko, 1991; Paulsen et al., 2000). Increasing cumulative incidence rates ranged between 20% and 51% (PPA) (Paulsen et al., 2000).

On paranoia (PPA) increasing point prevalence rates were reported and ranged between 20% and 46% (Swearer et al., 1996). For severe symptoms there was no trend and point prevalence rates ranged between 3% and 13% (Swearer et al., 1996).

**Hyperactivity subgroup**

In the two studies that reported on the hyperactivity subgroup as a whole the point prevalence rates ranged between 34% and 65% (PPA 46%–65%, completers 34%–54%) and in both studies there was a trend that hyperactivity increased during follow-up.

(Physical) aggression point prevalence rates ranged between 6% and 26% (all three studies PPA) (Swearer et al., 1996; Holtzer et al., 2003; Kunik et al., 2010). In two studies there was a trend that (physical) aggression increased during follow-up (Holtzer et al., 2003; Kunik et al., 2010). Cumulative prevalence rates ranged between 40% and 74% (PPA 40% and 41% and completers 74%) and incidence rate was 36% (completers), persistence rate was 22% (completers) and resolution rate was 25% (completers).

Wandering or agitation point prevalence rates ranged between 18% and 62%. In all three studies (PPA) there was a trend that wandering or agitation increased during follow-up (Holtzer et al., 2003; Kunik et al., 2010; Gonfrier et al., 2012). Cumulative prevalence rates ranged between
40% and 100%, persistence rate was 0%–67% and resolution rate 2%.

On mechanical and motor abnormalities/motor hyperactivity two studies reported increasing point prevalence rates that ranged between 10% and 61% (PPA 10%–61% and completers 14%–29%) (Swearer et al., 1996; Gonfrier et al., 2012). For severe symptoms (PPA), there was no trend and these point prevalence rates ranged between 0% and 16%. Cumulative prevalence rate was 50% (PPA) and incidence rate was 20% (completers). One study (completers) reported an increase in aberrant motor activity during follow-up (Cortes et al., 2005).

**Affective Subgroup**

In the two studies that reported on the affective subgroup as a whole the point prevalence rates ranged between 23% and 80% (PPA 67%–80% and completers 23%–26%). Cumulative prevalence rate was 86%, incidence rate was 57% and persistence rates ranged between 5% and 37% (all for PPA).

Depression point prevalence ranged between 10% and 42% (PPA 10%–42% and completers 29%–43%) and two studies (one PPA and one completers) reported a decrease in prevalence of depression during follow-up (Cortes et al., 2005; Holtzer et al., 2005). Cumulative prevalence rates ranged between 20% and 78% (PPA 20% and completers 22%–78%) and incidence was 8% (12% LBD and 6% AD) (completers). Persistence rates ranged between 33% and 35% (38% LBD and 31% AD) (completers) and resolution rate was reported in one study (completers) and was 15% (Haupt et al., 2000). One study (PPA or completers not given) reported cumulative incidence, persistence, resolution and fluctuating rates separately for AD and VaD (Li et al., 2001). For AD these rates were 15%, 27%, 67% and 7%, respectively. For VaD these rates were 27%, 67%, 22% and 11%, respectively.

Anxiety point prevalence rate ranged between 18% and 42% (PPA) and one study (completers) reported a decreasing number of affected patients with anxiety during follow-up (Cortes et al., 2005). Cumulative prevalence rates ranged between 43% (PPA) and 66% (completers) and cumulative incidence rate was 28% (PPA). One study (completers) reported a persistence rate of 12% (Haupt et al., 2000) and another study presented a decreasing trend in persistence rates that ranged from 22% to 1% (PPA) (Aalten et al., 2005). Resolution rate was reported in only one study (completers) and was 38% (Haupt et al., 2000).

Apathy point prevalence rates with increasing trend were reported in one study (PPA) and ranged from 43% to 63% (Gonfrier et al., 2012). In another study (completers) also an increasing number of affected patients with apathy during follow-up was reported (Cortes et al., 2005).

Hyperphagia, eating change in hyperphagia and eating change in hypophagia cumulative prevalence rates were reported in one study (PPA) (Keene and Hope, 1998). These rates were 26% for hyperphagia, 23% for eating change in hyperphagia and 66% for eating change in hypophagia.

One study (PPA) reported point prevalence rates on sleeping disturbance 9%–18%, cumulative prevalence 34%, cumulative incidence 31% and persistence 1%–20% (Aalten et al., 2005). One study (completers) reported on sleep disturbances and an increasing number of affected patients with sleep disturbances during follow-up was reported (Cortes et al., 2005).

**Discussion**

We found 23 prospective cohort studies with 7184 community-dwelling patients with dementia. NPS in general are highly prevalent, incident, and persistent although frequency parameters vary considerably across studies. Results presented for PPA compared to completers are diverse but not conclusive. Virtually all patients with dementia show any NPS during a period of 1–6 years. The overall quality of the studies was rather good (20 of 23 studies have a score of 60% or higher) although selection bias was present in all studies and confounding and information bias in the majority (19 of 23) of the studies.

Delusions/delusional misidentification, wandering/agitation, aberrant motor behavior/motor hyperactivity, and apathy are the most common NPS. For hallucinations, delusions/delusional misidentification, paranoia, aggression, wandering/agitation, aberrant motor behavior/motor hyperactivity, disinhibition, apathy, and sleep disturbance increasing trends in point prevalence rates have been found. Decreasing trends in depression and anxiety have also been found in some studies. For VaD compared to AD there were higher cumulative incidence and persistence rates. Resolution rate is higher for AD compared to VaD. NPS in the hyperactivity subgroup have higher point prevalence rates than in the psychotic subgroup and NPS in the affective subgroup have higher point prevalence rates than in the hyperactivity subgroup.

The variance in frequency parameters may partly be explained by different assessment instruments used, different intervals between assessments and different follow-up periods. The increasing trends for hallucinations, delusions/delusional
misidentification, aggression, wandering/agitation, aberrant motor behavior/motor hyperactivity, and apathy are in line with findings of studies on the course of NPS in nursing home patients with dementia, especially for wandering/agitation and apathy (Wetzels et al., 2010; Selbaek et al., 2013). Apparently a universal course of NPS exists regardless of the setting of patients with dementia. Although the development and course of NPS is a result of complex interactions between psychological, environmental and biological factors (de Vugt et al., 2004; Robert et al., 2005; Cipriani et al., 2011). Sleep disturbance, agitation/aggression and depression/dysphoria are the symptoms that causes caregivers the most emotional distress (Craig et al., 2005). Apathy is the symptom that gives a negative impact on the quality of the relationship between patient with dementia and caregiver and is a predictor for relationship change (de Vugt et al., 2003).

Strengths and limitations
In this systematic review, we used an extensive search strategy to identify relevant studies. We pre-tested the search strategy in a pilot assessment and we searched all relevant databases without language restriction. Finally, we independently extracted data and assessed the quality of included studies with a validated checklist of predefined criteria, which has been used in previous prognostic reviews (Kuijpers et al., 2004; Olde Hartman et al., 2009). We presented our results together with a quality score of each study to visualize the susceptibility of each study for bias, because the quality of the individual study influences outcomes.

There are also limitations to this review that should be acknowledged. Pooling of data was impossible due to the heterogeneity of the characteristics of the community-dwelling patients studied and the methods used. The presence of selection bias in all studies limits the generalizability of the results for the individual patients, but general trends in NPS have been found for the overall population. In two studies a small part (5.9% and 8.9%) of the sample was institutionalized. This could also influence the generalizability of the results. Furthermore, we included 22 studies from the USA and Europe and 1 study from Japan. Unfortunately, we have not been able to include any studies from Africa or South America. It is quite likely that the pattern of NPS and caregivers’ response may differ among various ethnic and cultural groups, especially between developing and developed countries. This also limits the generalizability of the results. The median number of participants enrolled in the cohorts of the included studies in this review is 170 with a wide range of 30–2288 study participants. Five studies enrolled less than 100 patients into the cohort (Rosen and Zubenko, 1991; Forstl et al., 1993; McShane et al., 1995; Sweer et al., 1996; Haupt et al., 2000). These low numbers of participants in the cohorts limit the strength of the evidence concerning outcomes.

In the included 23 studies several diagnostic criteria were used and even more different assessment instruments. There was a wide range in duration of follow-up and in six studies loss to follow-up was more than 20% per year. Therefore, the interpretation and comparability of these studies has been difficult. Further studies on the course of NPS which are large enough to follow up a substantial amount of patients over longer periods of time and which will use established assessment instruments will improve comparability and will give more information on the persistence and resolution of NPS. Together with important prognostic factors such as type and stage of dementia this will give important recommendations for professional long-term care for community-dwelling patients with dementia.

Conclusions
Delusions/delusional misidentification, wandering/agitation, aberrant motor behavior/motor hyperactivity, and apathy are the most common NPS. For hallucinations, delusions/delusional misidentification, paranoia, aggression, wandering/agitation, aberrant motor behavior/motor hyperactivity, disinhibition, apathy, and sleep disturbance increasing trends in point prevalence rates have been found. The increasing trends of several NPS in the course of dementia require a preventive approach of professional caretakers. For such an approach, a timely diagnosis and adequate professional support to prevent or diminish these problems is necessary.

Conflict of interest
None.

Description of authors’ roles
PB is the primary investigator of this systematic review and wrote the study design and the manuscript. PB, RBW and PLL independently extracted the information from the selected publications. All authors reviewed this systematic review and have given final approval of the version to be published.
Supplementary material

To view supplementary material for this article, please visit http://dx.doi.org/10.1017/S1041610214002282

Acknowledgments

The researchers of this systematic review would like to acknowledge Mrs. A. Tillema of the Radboud University Nijmegen Library of Medical Sciences for her help in completing the search strategies for the five databases and her help in aggregation of the databases in Endnote, finding duplicates and deleting them.

References


