

Caregiver rating bias in mild cognitive impairment and mild Alzheimer's disease: impact of caregiver burden and depression on dyadic rating discrepancy across domains

Livia Pfeifer,¹ Reinhard Drobotz,¹ Sonja Fankhauser,¹ Moyra E. Mortby,²
Andreas Maercker¹ and Simon Forstmeier¹

¹Institute of Psychology, Division Psychopathology and Clinical Intervention, University of Zurich, Binzmühlestr. 14/17, 8050 Zurich, Switzerland

²Centre for Research on Ageing, Health and Wellbeing, Australian National University, Eggleston Road, Canberra, ACT 0200, Australia

ABSTRACT

Background: Caregivers of individuals with dementia are biased in their rating of mental health measures of the care receiver. This study examines caregiver burden and depression as predictors of this bias for mild cognitive impairment and mild Alzheimer's disease in different domains.

Methods: The sample consisted of 202 persons: 60 with mild cognitive impairment, 41 with mild Alzheimer's disease, and 101 caregivers. Discrepancy scores were calculated by subtracting the mean caregiver score from the respective mean patient score on the following assessment instruments: the Geriatric Depression Scale, Apathy Evaluation Scale, Bayer-Activities of Daily Living Scale, and Quality of Life-AD scale. Caregiver burden and depression were assessed by the Zarit Burden Interview and the Center for Epidemiologic Studies Depression Scale.

Results: Intraclass correlation coefficients were low for apathy (0.38), daily functioning (0.38), and quality of life (0.30) and moderate for depression (0.49). These domains showed negative rating discrepancies, which indicates caregiver rating bias for all four domains. Regression analyses revealed that caregiver burden significantly contributed to explaining these discrepancies in the domains apathy, daily functioning, and quality of life.

Conclusion: Caregiver rating bias can be attributed to caregiver burden. When caregiver burden is present, data based on caregiver ratings should therefore be interpreted with caution.

Key words: Alzheimer's disease, mild cognitive impairment, caregiver, anosognosia, rating discrepancy, dyadic data

Introduction

Significant differences between the rating of mental health measures by individuals with dementia and the reports of their caregivers are well documented (Farias *et al.*, 2005; Leicht *et al.*, 2010; Schulz *et al.*, 2013). Discrepancy scores are often interpreted as a standard measure of how accurately patients with Alzheimer's disease (AD) are aware of cognitive impairment and deficits. The more the caregiver's rating exceeds the patient's rating, the more diminished is the patient's awareness (Clare, 2004). However, analyzing possible caregiver predictors of discrepancy scores leads to the assumption that

dyadic rating discrepancy is associated not only with patients' reduced awareness but also with caregiver characteristics. Hence, dyadic rating discrepancy might reflect a combination of patients' underrating their own suffering and caregivers' overestimating the patient's suffering – the caregiver rating bias. Several studies were indeed able to provide evidence for such a bias in various domains but show inconsistent results about which factors contribute to it (Ready *et al.*, 2004; Sands *et al.*, 2004; Conde-Sala *et al.*, 2008; Chang *et al.*, 2011; Clare *et al.*, 2011; Schulz *et al.*, 2013).

Discrepancies in reporting depressive symptoms have been found in several studies. Rosenberg *et al.* (2005) and Teri and Truax (1994) showed that caregiver burden and depression contribute to caregiver ratings of patients' depressive symptoms. Caregiver burden alone was found to be the only significant predictor of discrepancies in rating depression in Chang *et al.* (2011) and Burke *et al.*

Correspondence should be addressed to: Livia Pfeifer, MSc, Institute of Psychology, Division Psychopathology and Clinical Intervention, University of Zurich, Binzmühlestr. 14/17, 8050 Zurich, Switzerland. Phone: +41 44 635 73 26; Fax: +41 44 635 73 19. Email: lpfeifer@psychologie.uzh.ch. Received 6 Sep 2012; revision requested 14 Oct 2012; revised version received 19 Mar 2013; accepted 24 Mar 2013. First published online 7 May 2013.

(1998), but neither study addressed the influence of caregiver depression. Similar results were reported in studies on dyadic rating discrepancies in rating activities of daily living (ADL). Some researchers reported caregiver burden to be associated with the discrepancy scores (Schatzberg *et al.*, 1998; Zanetti *et al.*, 1999; Clare *et al.*, 2011), although only Zanetti *et al.* (1999) assessed both caregiver burden and depression. In contrast, Argüelles *et al.* (2001) showed that caregiver depression but not burden was related to caregiver rating bias. For quality of life, the domain with the most numerous and most current studies, the consensus is that caregiver burden is a strong predictor of dyadic rating discrepancy (Karlavish *et al.*, 2001; Sands *et al.*, 2004; Conde-Sala *et al.*, 2008; Schulz *et al.*, 2013). Caregiver depression, however, has been seen as tending to play a less significant role in the dyadic rating discrepancy of quality of life (Karlavish *et al.*, 2001; Schulz *et al.*, 2013).

To our knowledge, no study until now has evaluated caregiver burden or depression as predictors of dyadic rating discrepancy in apathy. Apathy, defined as a lack of motivation, i.e. behavioral, cognitive, and emotional concomitants of goal-directed behavior, is one of the most prevalent neuropsychiatric symptoms in AD and mild cognitive impairment (MCI); Clarke *et al.* (2011). Although definitions of depression overlap with it in terms of key symptoms, apathy may be viewed as distinguishable from depression in dementia (Mortby *et al.*, 2012). Apathy measurements are usually based on self- and informant reports. Because apathy is associated with caregiver burden, investigating the influence of burden on caregivers' apathy ratings is of major importance. Accurate assessment is crucial to improving understanding and management of apathy (Clarke *et al.*, 2011).

However, caregiver burden and depression are the most widely used caregiver variables for explaining dyadic rating discrepancy and appear to contribute to the caregiver rating bias in various domains. Care of cognitively impaired people is physically and mentally demanding and often associated with burden and depression. In addition, it may lead to distorted perception and exaggeration of the patient's symptoms (Schulz *et al.*, 2013). Caregiver burden and depression can be considered as two separate variables. Caregiver burden is defined as a reaction to the physical, emotional, economic, and social costs of the caregiving relationship, while caregiver depression is a mood disturbance triggered by the stress of providing care and manifested as feelings of loneliness, isolation, fearfulness, and being easily annoyed (Pinquart and Sörensen, 2003). Behavioral problems, limited ADL, and cognitive impairment are strongly related

to caregiver burden and depression (Pinquart and Sörensen, 2003; Schoenmakers *et al.*, 2010). In addition, caregiver depression seems to be mediated by caregiver burden (Clyburn *et al.*, 2000).

It remains unresolved, however, whether caregiver burden or depression predicts caregiver rating bias within specific domains in a similar way or if there are domain-specific differences. Clare *et al.* (2011) alone demonstrated that predictors of rating discrepancies in the domains of memory, ADL, and social functioning were not domain-specific. Caregiver burden was associated with discrepancies in rating memory and ADL but not in social functioning; caregiver depression was not addressed. On the basis of the results presented above, one might assume that caregiver burden is a factor that predicts rating discrepancies in various domains. But even if the study samples are similar (patients with mild to moderate dementia), the results are based on different methods and hence only comparable to a limited extent. To our knowledge, no study has analyzed the association between both caregiver burden and depression and caregiver rating bias simultaneously across various domains such as depression, apathy, daily functioning, and quality of life. These domains are of high relevance for dementia and are often assessed by informant reports.

Hence, the main purpose of the current study was to examine domain-specific relationships between caregiver burden or depression and dyadic rating discrepancies. We hypothesized that caregiver burden is more prominent than caregiver depression in predicting dyadic rating discrepancy. We also hypothesized the association between caregiver burden or depression and the dyadic rating discrepancy to be domain-specific. Furthermore, we explored the extent to which discrepancies were related to the severity of cognitive impairment and to patient-caregiver relationship, as it is not yet clear to what degree these variables influence the rating discrepancy (Farias *et al.*, 2005; Mougias *et al.*, 2011).

Methods

Participants

Participants were 101 elderly individuals, 60 of them with MCI and 41 with AD (see Table 1). The sample was derived from the Swiss longitudinal study "Motivational Reserve as a Protective Factor in Mild Alzheimer's Dementia and Mild Cognitive Impairment" (MoReA); only baseline data were used. Participants were recruited from 17 collaborating local hospitals or clinics in the German-speaking part of Switzerland. For inclusion, participants had to be diagnosed with

Table 1. Characteristics of study participants and group comparisons

	TOTAL	MCI	AD	<i>t/χ²</i>	p
Patients N	101	60	41		
Age, mean (years)	77.22 (8.19)	74.82 (7.51)	80.73 (7.95)	-3.79	0.000
Education, mean (years)	13.18 (10.89)	14.15 (13.64)	11.62 (3.02)	1.37	0.175
Gender (% female)	50.5	40	65.9	6.51	0.011
MMSE	25.26 (3.27)	26.98 (2.04)	22.64 (3.06)	7.80	0.000
NPI	7.4 (11.03)	6.68 (9.92)	8.44 (12.53)	-0.75	0.455
Caregivers N	101	60	41		
Age, mean (years)	65.90 (13.30)	67.25 (13.47)	63.95 (12.97)	1.23	0.221
Gender (% female)	75.2	76.7	73.2	0.16	0.689
Burden (ZBI)	15.24 (12.62)	12.67 (12.34)	19.01 (12.21)	-2.54	0.013
Depression (CES-D)	6.24 (6.29)	6.01 (6.47)	6.59 (6.09)	-0.46	0.650
Relationship N (%)				11.33	0.023
Partner	61.4	71.7	46.3		
Child	25.7	18.3	36.6		
Other	12.9	10.0	17.1		
Relationship, mean (years)	47.10 (14.79)	44.86 (15.67)	50.38 (12.90)	-1.89	0.062
Living situation N (%)				15.85	0.003
Alone	26.7	25.0	29.3		
Partner	60.4	71.7	43.9		
Other	5.0	1.7	9.8		
Nursing home	5.9	0	14.6		

t, χ^2 = values of statistics from the paired *t*-tests (with standard deviation) and χ^2 tests; NPI = Neuropsychiatric Inventory; ZBI = Zarit Burden Interview; CES-D = Center for Epidemiologic Studies Depression Scale; Relationship, mean (years): years of knowing each other.

either MCI or mild AD and had to be age 60 or older.

In each of the cooperating memory clinics, an interdisciplinary team assigned the diagnosis after thorough neurological, psychiatric, clinical, and neuropsychological assessments. Clinical dementia was diagnosed according to the criteria of the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV; American Psychiatric Association, 2000). To meet a diagnosis of AD, participants needed to demonstrate gradual onset and progressive deterioration of cognitive functioning and have all other specific causes of dementia excluded. Our clinical diagnosis of AD corresponds to the diagnosis of “probable Alzheimer’s disease” according to the National Institute of Neurological and Communicative Disorders and Stroke–Alzheimer’s Disease and Related Disorders Association (NINCDS-ADRDA) criteria (McKhann *et al.*, 2011). Only mild AD cases with a score of 1 on the Clinical Dementia Rating scale (CDR; Morris, 1997) and scores between 18 and 26 on the Mini-Mental State Examination (MMSE; Folstein *et al.*, 1975) were included. The MCI diagnosis was based on international consensus criteria (Winblad *et al.*, 2004), including the following criteria: absence of dementia as diagnosed by DSM-IV criteria (MMSE \geq 24); cognitive decline, i.e. self- and/or informant report and

impairment in completing objective tasks and/or evidence of decline over time on objective cognitive tasks; preserved basic ADL and not exceeding minimal impairment in complex instrumental functions (CDR \leq 0.5); at least mild impairment in one of the following cognitive domains: memory, language, praxis, executive function, or attention. Exclusion criteria of the present study were a history of a malignant disease, severe organ failure, metabolic or hematologic disorders, neurosurgery or neurological condition such as Parkinson’s disease, epilepsy, and postencephalitic and postconcussional syndrome.

Each participant was required to be accompanied by a reliable collateral source, typically the partner; a child or a close friend was also accepted. Although level of impairment of the patients in the present study was relatively low, it can be assumed that the family members and friends performed a caregiving function (Garand *et al.*, 2005). Written informed consent was obtained from all participants and caregivers prior to inclusion.

Measures

General cognitive functioning of the patients was assessed with the MMSE (Folstein *et al.*, 1975). For the assessment of neuropsychiatric symptoms, the Neuropsychiatric Inventory (NPI; Cummings,

1997), a reliable, informant-based rating scale, was used. Severity and frequency of 12 neuropsychiatric symptoms that are typical of dementia are scored on the basis of structured questions. Scores range from 0 to 144, with higher scores indicating more severe symptoms.

The following assessment instruments to identify the patient's abilities or difficulties were administered with parallel forms to patient and caregiver separately. Depression was measured with the Geriatric Depression Scale (GDS; Yesavage *et al.*, 1983), a frequently used instrument for screening depression in the elderly people. The short version consists of a series of 15 yes/no self-referent statements to identify the presence of depression. The Apathy Evaluation Scale (AES; Marin *et al.*, 1991) was developed for multiple rater sources (clinician, informant, and self) to quantify and characterize apathy in adult patients. The scale comprises 18 items, which are rated on a four-point scale with the following categories: not at all characteristic, slightly characteristic, somewhat characteristic, and very characteristic. We assessed daily functioning with the Bayer-Activities of Daily Living Scale (B-ADL; Hindmarch *et al.*, 1998). The scale was developed within an international research project to assess deficits in performance of everyday activities in patients with mild to moderate dementia. The B-ADL consists of 25 items concerning frequency of difficulties, with a ten-point response scale ranging from "never" to "always." The Quality of Life-AD (QoL-AD; Logsdon *et al.*, 1999) is based on direct interviews with AD patients and a questionnaire consisting of 13 items on a four-point scale. Responses are rated from 1 (poor) to 4 (excellent).

Caregivers completed the Zarit Burden Interview (ZBI; Zarit *et al.*, 1980), a 22-item self-report inventory of perceived caregiver strain. Guidelines suggest interpreting severity of burden as follows: 61–88 as severe, 41–60 as moderate, 21–40 as weak, and scores under 21 as hardly at all (Braun *et al.*, 2010). To assess depression in caregivers, we used the short form of the Allgemeine Depressionsskale (ADS-K; Hautzinger and Bailer, 1993), the German version of the Center for Epidemiologic Studies Depression Scale (CES-D; Radloff, 1977). The critical cut-off point for this 15-item scale in the German version is a summary score of 18.

Data analysis

All statistical analyses were performed at an α level of 0.05 (two-tailed), using IBM SPSS Statistics 20.0. Descriptive statistics included frequencies for categorical variables (gender and relationship) and means and standard deviations (SD) for continuous

measures (age of patient, age of caregiver, years of education, MMSE, neuropsychiatric symptoms, caregiver burden, and years of relationship) for the entire sample as well as for the MCI and the AD group separately. χ^2 analyses were used to assess associations between categorical variables, and independent *t*-tests to compare continuous measures between groups.

The level of agreement between patients' and their caregivers' ratings was determined using intraclass correlation coefficients (ICC). ICC < 0.40 indicated weak agreement, 0.40–0.75 moderate to good agreement, and ICC > 0.75 outstanding agreement (Lee *et al.*, 1989). Paired *t*-tests were run to determine whether patient and caregiver ratings differed significantly. Effect sizes were calculated using Cohen's *d*: $d = 0.2$ was taken to indicate a small, $d = 0.5$ a moderate, and $d = 0.8$ a large effect size (Cohen, 1988).

Raw discrepancy scores for depression, apathy, daily functioning, and quality of life were calculated by subtracting the mean caregiver score from the respective mean patient score. Reversing the polarity of the QoL-AD scores made the values more comparable to those of the other scales, i.e. a higher score indicated a poorer status. Discrepancy scores differing significantly from zero provide evidence for a systematic bias, which was proved with a dependent *t*-test. A mean score difference lower than zero indicates that status of the patients was either underestimated by themselves or overestimated by the caregivers.

To explore the relationship between all variables, bivariate correlations were calculated. Furthermore, domain-specific stepwise multiple regression equations were conducted to determine the extent to which caregiver burden and depression were associated with the occurrence of rating discrepancies. Potential multicollinearities of the independent variables were evaluated. Because our sample size was relatively small, only a small number of variables could be considered. On the basis of the literature, we entered patients' age, sex, education, and neuropsychiatric symptoms in the first step of the model as control variables. By integrating the diagnosis (MCI/AD) in the second step, effects caused by the cognitive state could be controlled. For the same reason, the type of relationship (spouse versus others) was entered into a third step of the model. The final variables entered were caregiver burden and depression.

Results

Patient and caregiver characteristics are presented in Table 1. The overall study group comprised

Table 2. Dyadic discrepancy in rating depression, apathy, daily functioning, and quality of life ($N = 101$)

SCALE	PATIENT SELF-RATING	CAREGIVER RATING OF THE PATIENT	DISCREPANCY	<i>t</i>	COHEN'S <i>d</i>	ICC
GDS	3.37 (2.65)	4.39 (3.52)	-1.03 (3.16)	-3.26*	0.33	0.49
AES	13.93 (6.61)	19.97 (11.36)	-6.04 (10.37)	-5.86**	0.65	0.38
B-ADL	69.14 (37.52)	92.86 (57.73)	-23.35 (54.42)	-4.29**	0.49	0.38
QoL-AD	-39.69 (5.00)	-36.76 (5.82)	-2.93 (6.44)	-4.58**	0.54	0.30

GDS = Geriatric Depression Scale; AES = Apathy Evaluation Scale; B-ADL = Bayer-Activities of Daily Living Scale; QoL-AD = Quality of Life-AD; ratings and discrepancy = mean with standard deviation; *t* = value of *t*-statistic from the paired *t*-test; ICC: intraclass correlation coefficient.

* $p < 0.01$, ** $p < 0.001$.

60 patients with MCI and 41 with AD. These two groups differed significantly in patients' age, gender, and cognitive status (MMSE), as well as in caregiver burden, the relationship, and their living arrangements. Only 14.6% of the AD patients lived in nursing homes. MCI patients were younger than the AD patients ($t(99) = -3.79$, $p < 0.001$) and had higher MMSE scores ($t(60) = 7.80$, $p < 0.001$). A lower percentage of the MCI patients were female ($\chi^2(1) = 6.51$, $p < 0.05$). Furthermore, the MCI caregivers were less burdened than the AD caregivers ($t(99) = -2.54$, $p < 0.05$) and more often spouses ($\chi^2(4) = 11.33$, $p < 0.05$). Group differences in caregivers' age were explained by the varying types of relationship between patients and caregivers. Of the MCI patients, 71.7% were accompanied by spouses, in contrast to only 46.3% of the AD patients, who were more often accompanied by other family members, such as their children. In the present study, caregivers were on average neither burdened nor depressed. Even so, according to the cumulative frequencies, 29.7% of the caregivers had a ZBI value over the critical cut-off score of 21 and 4% a CES-D summary score above 18.

To assess the dyadic rating agreement, we calculated ICCs (see Table 2). ICCs for apathy (0.38), daily functioning (0.38), and quality of life (0.30) were low, and those for depression moderate (0.49). Patient and caregiver ratings differed significantly in all four domains (GDS: $t(100) = -3.26$, $p < 0.01$; AES: $t(100) = -5.86$, $p < 0.000$; B-ADL: $t(100) < -4.29$, $p < 0.000$; QoL-AD: $t(100) = -4.58$, $p < 0.000$). All discrepancy scores showed a negative value, differing significantly from zero. The statistical magnitude of this bias, defined by the effect size *d*, was low for depression and moderate for apathy, daily functioning, and quality of life.

Bivariate correlations among discrepancy scores are shown in Table 3. All discrepancy scores were significantly positively related to each other and negatively to caregiver burden. In addition,

depression discrepancy was related to patient's age and to caregiver depression, while apathy discrepancy was associated with the patient's neuropsychiatric symptoms. Discrepancies in rating daily functioning were related to patient's age and gender, MMSE, caregiver's age, and the relationship. Discrepancies in rating quality of life were related to the duration of the patient-caregiver relationship. The correlation coefficient between caregiver burden and depression was 0.50 ($p < 0.01$).

Table 4 displays results of the multivariate linear regression analyses. Caregiver burden significantly contributed to explaining dyadic rating discrepancies in apathy, daily functioning, and quality of life. Depression discrepancy was predicted by caregiver burden and depression with a similar non-significant β value. However, we observed a significant change in R^2 in step four with regard to the depression discrepancy. The type of relationship (spouses vs. others) did not predict any dyadic rating discrepancy. Only discrepancies in rating daily functioning were predicted by diagnosis; these were smaller for patients with MCI than with AD. Moreover, patient's age was a significant predictor for discrepancies in rating depression and daily functioning. These results reflect the unique contribution of caregiver burden in predictions for each domain, over and above the effects of covariates.

Discussion

The aim of the present study was to investigate to what extent caregiver burden and depression are related to caregiver rating bias regarding the patient's depression, apathy, daily functioning, and quality of life.

In accordance with previous literature, dyadic rating agreement on apathy, daily functioning, and quality of life was poor (Ready *et al.*, 2004; Sands *et al.*, 2004; Farias *et al.*, 2005; Leicht

Table 3. Bivariate correlations among study variables

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Patient															
1. Age	1														
2. Gender (1 = m; 2 = f)	0.05	1													
3. Education (years)	-0.04	-0.21*	1												
4. NPI	-0.11	-0.06	-0.03	1											
5. MMSE	-0.43***	-0.14	0.07	0.10	1										
Caregiver															
6. Age	0.13	-0.21*	0.02	-0.03	0.09	1									
7. Gender (1 = m; 2 = f)	0.31**	-0.52***	0.11	-0.05	-0.06	-0.11	1								
8. Relationship	-0.29**	-0.54**	0.12	0.10	0.20*	0.59***	-0.13	1							
9. Relationship (years)	0.44***	0.21*	-0.06	-0.04	-0.11	0.37***	-0.01	-0.13	1						
10. Burden	0.14	-0.05	-0.00	0.15	-0.33**	-0.00	0.06	-0.12	0.03	1					
11. Depression	0.03	-0.16	0.10	0.02	-0.18	0.13	-0.03	0.04	0.10	0.50***	1				
Discrepancy															
12. GDS	-0.22*	-0.08	0.04	-0.17	0.04	-0.12	-0.01	0.00	-0.19	-0.31**	-0.25*	1			
13. AES	-0.10	0.04	0.17	-0.20*	0.06	-0.13	0.04	-0.03	-0.12	-0.37***	-0.18	0.32**	1		
14. B-ADL	-0.37***	-0.21*	0.05	-0.08	0.40***	0.20*	-0.05	0.32**	-0.08	-0.39***	-0.14	0.39***	0.36***	1	
15. QoL-AD	-0.01	0.06	0.10	0.17	0.02	-0.04	0.04	0.02	-0.21*	-0.40***	-0.20	0.39***	0.43***	0.28**	1

Relationship = spouses versus others; GDS = Geriatric Depression Scale; AES = Apathy Evaluation Scale; B-ADL = Bayer-Activities of Daily Living Scale; QoL-AD = Quality of Life-AD; MMSE = Mini-Mental State Examination; NPI = Neuropsychiatric Inventory.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

The values represent Pearson correlations (between two continuous variables), point-biserial correlations (between a continuous and a dichotomous variable), or ϕ coefficients (between two dichotomous variables).

Table 4. Results of stepwise regression analyses predicting dyadic discrepancies in rating depression, apathy, daily functioning, and quality of life (N = 101)

	GDS DISCREPANCY			AES DISCREPANCY			B-ADL DISCREPANCY			QoL-AD DISCREPANCY		
	B	SE B	β	ΔR^2	B	SE B	β	ΔR^2	B	SE B	β	ΔR^2
Step 1. Patient				0.11*				0.08				0.21***
Age	-0.11	0.04	-0.27**		-0.16	0.13	-0.12		-2.64	0.62	-0.40***	
Sex	-0.59	0.65	-0.09		1.63	2.13	0.08		-21.46	10.22	-0.20*	
Education	0.00	0.03	0.00		0.17	0.10	0.17		-0.03	0.47	-0.01	
NPI	-0.06	0.03	-0.20		-0.19	0.10	-0.19		-0.65	0.47	-0.13	
Step 2. Diagnosis				0.01				0.00				0.16***
MCI vs. AD	-0.68	0.72	-0.10		-0.65	2.39	-0.03		-47.77	10.25	-0.43***	
Step 3. Relationship				0.01				0.00				0.02
Spouses vs. others	-0.87	0.80	-0.13		-1.40	2.67	-0.07		18.10	11.30	0.16	
Step 4. Caregiver				0.09*				0.11**				0.07**
Burden (ZBI)	-0.04	0.03	-0.18		-0.28	0.10	-0.35**		-1.14	0.41	-0.27**	
Depression (CES-D)	-0.09	0.06	-0.17		-0.02	0.19	-0.01		0.005	0.79	0.00	
												0.14**
												0.02
												0.06
												0.12
												0.02

GDS = Geriatric Depression Scale ($R^2 = 0.22$); AES = Apathy Evaluation Scale ($R^2 = 0.19$); B-ADL = Bayer-Activities of Daily Living Scale ($R^2 = 0.45$); QoL-AD = Quality of Life-AD ($R^2 = 0.17$); NPI = Neuropsychiatric Inventory; ZBI = Zarit Burden Interview; CES-D = Center for Epidemiologic Studies Depression Scale. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

et al., 2010; Schulz et al., 2013). The ICC for depression was moderate; one explanation for this divergence could be that the GDS is based on a two-point scale, which increases the probability of agreement. Negative rating discrepancies were found in all four domains, which indicate that either the patients underrated their own suffering, or that the caregivers overestimated it, or both.

With respect to the main objective of our study, our data indicated that caregiver burden is a more important predictor of caregiver rating bias than is caregiver depression, which tended to play a less significant role in dyadic rating discrepancies. Our findings thus confirm the first hypothesis that caregiver rating bias can be attributed to caregiver burden. Consistent with our second hypothesis, our findings show domain-specific patterns. Because the predictor structure differs in relation to the domain, the influence of caregiver burden and depression on rating discrepancies cannot be generalized across domains. In addition, there is empirical support to show that dyadic rating discrepancy is not associated only with reduced awareness of the patient but also with caregiver characteristics.

After controlling for demographic characteristics and neuropsychiatric symptoms of the patient, our regression analysis revealed that daily functioning was the only domain predicted by the level of cognitive impairment. Discrepancy scores were higher for the AD group than for the MCI group. These findings are consistent with several previous studies reporting existent or non-existent associations between cognitive impairment and rating discrepancies (Ready et al., 2004; Sands et al., 2004; Farias et al., 2005; Clare et al., 2011). Farias et al. (2005) argued that the group difference reflects a decreased awareness in patients with dementia. Our data do not support this assumption, given that a reduction of awareness would influence ratings of all domains, not just daily functioning. We presume that the difference is instead due to a ceiling effect because, in line with the diagnostic criteria, patients with MCI show few limitations in daily functioning (Winblad et al., 2004). Hence, it might be easier to rate the absence of difficulties than to rank them on a ten-point response scale, making discrepancies smaller. Furthermore, our data showed no associations between patient-caregiver relationship and rating discrepancies, which is consistent with other studies (Sands et al., 2004; Farias et al., 2005). Quality of relationship might be a better mechanism than type of relationship for explaining rating discrepancies (Farias et al., 2005).

The last step of the regression analysis revealed a significant association between caregiver burden and rating discrepancies in the domains apathy,

daily functioning, and quality of life. Caregiver depression was not related to any rating discrepancy in the three domains. These findings appear to bolster the argument that caregiver burden and depression, albeit related, are indeed separate variables. Our findings are also consistent with several studies that reported caregiver burden as being a significant predictor of discrepancies in rating daily functioning (Zanetti *et al.*, 1999; Clare *et al.*, 2011) and quality of life (Sands *et al.*, 2004; Conde-Sala *et al.*, 2008; Mougias *et al.*, 2011; Schulz *et al.*, 2013). However, Schulz *et al.* (2013) reported an association between caregiver burden and depression with the discrepancy in rating quality of life. Their finding deviates from our results, which showed a significant bivariate correlation between QoL-AD discrepancy and caregiver burden but not between QoL-AD discrepancy and caregiver depression. This deviance could have been caused by the sample. Caregivers in the study by Schulz *et al.* (2013) were comparably burdened but significantly more depressed than those in the present study. Schulz *et al.*'s multiple regression analysis reveals that caregiver burden alone predicts the QoL-AD discrepancy score significantly, which is in line with our results.

To our knowledge, predictors of apathy rating discrepancy have yet to be investigated. Given that apathy is associated with caregiver burden, it is meaningful to know that caregiver burden might lead to a bias in rating apathy. In addition, the different predictor structure for apathy and depression adds to the evidence that apathy may be a distinguishable syndrome from depression in dementia (Mortby *et al.*, 2012). Our findings regarding depression rating discrepancy, compared to the other domains, show a different pattern. β -values for caregiver burden and depression were similar and not significant. The results of the regression analysis may have been ambiguous because caregiver burden and depression share a substantial portion of variance ($r = 0.50$). However, we conducted a *post hoc* analysis in which caregiver depression and burden were included separately as a single predictor in the fourth step of the multiple regression analysis. These findings confirm our assumption that both caregiver burden ($\beta = -0.27$, $p < 0.01$) and caregiver depression ($\beta = -0.26$, $p < 0.01$) predict discrepancies in rating depression. These results are also in accordance with Rosenberg *et al.* (2005), who also concluded that caregiver depression and burden affect rating discrepancy in depression and that a depressed informant is likely to report more depressive symptoms than the patient does. Similar to our findings, the effects by Rosenberg *et al.* (2005) were notable but not large

in magnitude, and caregivers in their study were minimally burdened and depressed.

As described above, caregiver burden seems to be more frequently associated with caregiver rating bias than is caregiver depression. We argue that this bias can be attributed to the demanding situation to which caregivers are exposed; caring for a person with dementia is associated with physical, emotional, economical, and social costs. According to Zanetti *et al.* (1999), caregivers' judgments were especially influenced by demands and restrictions on their time. There are several possible reasons for an association between caregiver rating bias and burden. Caregivers' suffering may foster a growing intolerance to the patients' symptoms, or caregivers may transfer their own feelings of worry, frustration, fear, etc. into their ratings, inducing an overestimation of the patient's condition (Ready *et al.*, 2004; Sands *et al.*, 2004). Furthermore, observers have been said to give more weight to negative than to positive information, which influences their ratings (Farias *et al.*, 2005). We assume that this effect is more pronounced in burdened caregivers. According to the Attribution Bias Context (ABC) model (De Los Reyes and Kazdin, 2005), rating discrepancies in childhood psychopathology were caused by disparities among informants' attributions of the causes of the child's behavior. The ABC model could explain caregiver bias in rating dementia as well. Caregiver burden might lead to a differential weighting between environmental and dispositional causes of the patient's behavior. Further longitudinal studies are needed to test the potential explanations.

At the same time, several limitations of our study should be taken into consideration. First, caregivers were on average neither burdened nor depressed. This is not entirely surprising in light of the fact that limited ADL, cognitive impairment, and behavioral problems – variables that are strongly related to caregiver burden and depression – were relatively low in the present sample (Pinquart and Sörensen, 2003; Schoenmakers *et al.*, 2010). Given that solely caregivers who feel strongly burdened tend to develop a depression, it is not surprising that caregivers of patients with MCI and mild AD are only marginally depressed and do not differ with respect to their depression scores (Clyburn *et al.*, 2000; Ready *et al.*, 2004; Schoenmakers *et al.*, 2010). Consequently, it can be assumed that in samples with a higher degree of cognitive impairment, results might be more pronounced. Second, it would have been useful to assess the level of patients' awareness to analyze to what extent variance of the rating discrepancy is explained by the patient's awareness and to what extent by caregiver burden. In addition,

an association between impaired awareness and caregiver burden has also been consistently reported in the literature (Clare *et al.*, 2011). Unfortunately, the present study does not explain the relationship between rating discrepancy, caregiver burden, and patients' awareness. Third, the patients' neuropsychiatric symptoms were conducted with the NPI (Cummings *et al.*, 1997), which is based on caregiver reports. We realize that integrating caregiver reports as predictors for caregiver rating bias could be confounding. Nevertheless, neuropsychiatric symptoms seem to contribute to rating discrepancy and should therefore be controlled in the regression analysis (Clare *et al.*, 2011). Finally, a further potential limitation of this study includes the modest sample in relation to the large number of statistical tests performed. As such, the findings of the present study should be viewed cautiously and replicated in a larger sample.

Despite these limitations, the results of this study have important clinical implications for awareness research and clinical trials that include caregiver ratings. On the one hand, a dyadic rating discrepancy cannot be directly attributed to reduced patient awareness without controlling for caregiver state. This may explain some of the apparently contradictory findings reported in the awareness literature (Clare, 2004). On the other hand, although caregiver reports are a frequently used approach in dementia research (Schulz *et al.*, 2013), using them without controlling for caregiver burden can therefore lead to substantial distortions, which might influence diagnosis and therefore type and frequency of treatment. This, in turn, can have negative effects on the patient's health and well-being as well as a considerable impact on healthcare. Due to the strong implications of the caregiver rating for patients, it is of major importance to invest resources in caregiver intervention programs.

In summary, this study has provided a comprehensive examination of caregiver rating bias in relation to caregiver burden in MCI and AD. Our findings indicate that dyadic discrepancy in rating depression, apathy, daily functioning, and quality of life seems to be associated with caregiver burden. When caregiver burden is present, data based on caregiver ratings might lead to a substantial caregiver rating bias and thus should be interpreted with caution. Because caregiver ratings are of high clinical relevance, particularly in MCI and AD populations, caregiver burden assessment should be included in clinical practice.

Conflict of interest

None.

Description of authors' roles

L. Pfeifer was responsible for formulating research questions, conducting the statistical analyses, and drafting the paper. S. Forstmeier and A. Maercker designed and carried out the background study (MoReA) and contributed to data analysis. R. Drobetz, S. Fankhauser, and M.E. Mortby contributed to the acquisition and generation of the data. Each co-author contributed to data interpretation, critically discussed all versions of the manuscript, and revised and finally approved the version to be published.

Acknowledgments

This project was funded by the Swiss National Science Foundation SNSF, the Velux Foundation, Tropos Foundation, Kurt Fries Foundation, and the Swiss Alzheimer's Association.

References

- American Psychiatric Association** (2000). *Diagnostic and statistical manual of mental disorders*, DSM-IV-TR. Washington, DC: American Psychiatric Publishing, Inc.
- Argüelles, S., Loewenstein, D. A., Eisdorfer, C. and Argüelles, T.** (2001). Caregivers' judgments of the functional abilities of the Alzheimer's disease patient: impact of caregivers' depression and perceived burden. *Journal of Geriatric Psychiatry and Neurology*, 14, 91–98. doi:10.1177/089198870001300402.
- Braun, M., Scholz, U., Hornung, R. and Martin, M.** (2010). Caregiver burden with dementia patients. A validation study of the German language version of the Zarit Burden Interview. *Zeitschrift für Gerontologie und Geriatrie*, 43, 111–119. doi:10.1080/13607860902879441.
- Burke, W. J., Roccaforte, W. H., Wengel, S. P., McArthur-Miller, D., Folks, D. G. and Potter, J. F.** (1998). Disagreement in the reporting of depressive symptoms between patients with dementia of the Alzheimer type and their collateral sources. *American Journal of Geriatric Psychiatry*, 6, 308–319.
- Chang, Y. P., Edwards, D. F. and Lach, H. W.** (2011). The collateral source version of the Geriatric Depression Scale: evaluation of psychometric properties and discrepancy between collateral sources and patients with dementia in reporting depression. *International Psychogeriatrics*, 23, 961–968. doi:10.1017/S1041610211000147.
- Clare, L.** (2004). Awareness in early-stage Alzheimer's disease: a review of methods and evidence. *British Journal of Clinical Psychology*, 43, 177–196. doi:10.1348/014466504323088042.
- Clare, L. et al.** (2011). The influence of psychological, social and contextual factors on the expression and measurement of awareness in early-stage dementia: testing a biopsychosocial model. *International Journal of Geriatric Psychiatry*, 27, 167–177. doi:10.1002/gps.2705.

- Clarke, D. E., Ko, J. Y., Kuhl, E. A., van Reekum, R., Salvador, R. and Marin, R. S.** (2011). Are the available apathy measures reliable and valid? A review of the psychometric evidence. *Journal of Psychosomatic Research*, 70, 73–97. doi:10.1016/j.jpsychores.2010.01.012.
- Clyburn, L. D., Stones, M. J., Hadjistavropoulos, T. and Tuokko, H.** (2000). Predicting caregiver burden and depression in Alzheimer's disease. *Journal of Gerontology: Psychological Sciences*, 55B, 2–13. doi:10.1093/geronb/55.1.S2.
- Cohen, J.** (1988). *Statistical Power Analysis for the Social Sciences*. Hillsdale, NJ: Erlbaum.
- Conde-Sala, J. L., Garre-Olmo, J., Turro-Garriga, O., Lopez-Pousa, S. and Vilalta-Franch, J.** (2008). Factors related to perceived quality of life in patients with Alzheimer's disease: the patient's perception compared with that of caregivers. *International Journal of Geriatric Psychiatry*, 24, 585–594. doi:10.1002/gps.2161.
- Cummings, J. L.** (1997). The Neuropsychiatric Inventory: assessing psychopathology in dementia patients. *Neurology*, 48, 10–16.
- De Los Reyes, A. and Kazdin, A. E.** (2005). Informant discrepancies in the assessment of childhood psychopathology: a critical review, theoretical framework, and recommendations for further study. *Psychological Bulletin*, 131, 483–509. doi:10.1037/0033-2909.131.4.483.
- Farias, S. T., Mungas, D. and Jagust, W.** (2005). Degree of discrepancy between self and other-reported everyday functioning by cognitive status: dementia, mild cognitive impairment, and healthy elders. *International Journal of Geriatric Psychiatry*, 20, 827–834. doi:10.1002/gps.1367.
- Folstein, M. F., Folstein, S. E. and McHugh, P. R.** (1975). "Mini Mental State": a practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12, 189–198. doi:10.1016/0022-3956(75)90026-6.
- Garand, L., Dew, M. A., Eazor, L. R., DeKosky, S. T. and Reynolds, C. F.** (2005). Caregiving burden and psychiatric morbidity in spouses of persons with mild cognitive impairment. *International Journal of Geriatric Psychiatry*, 20, 512–522. doi:10.1002/gps.1318.
- Hautzinger, M. and Bailer, M.** (1993). *Allgemeine Depressions Skala*. Göttingen: Beltz Test GmbH.
- Hindmarch, I., Leheld, H., Jongh, P. and Erzigkeit, H.** (1998). The Bayer Activities of Daily Living Scale (B-ADL Scale). *Dementia and Geriatric Cognitive Disorders*, 9, 20–26. doi:10.1159/000051195.
- Karlawish, J. H. T., Casarett, D., Klocinski, J. and Clark, C. M.** (2001). The relationship between caregivers' global ratings of Alzheimer's disease patients' quality of life, disease severity, and the caregiving experience. *Journal of the American Geriatrics Society*, 49, 1066–1070. doi:10.1046/j.1532-5415.2001.49210.x.
- Lee, J., Koh, D. and Ong, C. N.** (1989). Statistical evaluation of agreement between two methods for measuring a quantitative variable. *Computers in Biology and Medicine*, 19, 61–70. doi:10.1016/0010-4825(89)90036-X.
- Leicht, H., Berwig, M. and Gertz, H.-J.** (2010). Anosognosia in Alzheimer's disease: the role of impairment levels in assessment of insight across domains. *Journal of the International Neuropsychological Society*, 16, 463–473. doi:10.1017/S1355617710000056.
- Logsdon, R. G., Gibbons, L. E., McCurry, S. M. and Teri, L.** (1999). Quality of life in Alzheimer's disease: patient and caregiver reports. *Journal of Mental Health and Aging*, 5, 21–32. doi:10.1037/t00930-000.
- Marin, R. S., Biedrzycki, R. C. and Firinciogullari, S.** (1991). Reliability and validity of the apathy evaluation scale. *Psychiatry Research*, 38, 143–162. doi:10.1016/0165-1781(91)90040-V.
- McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D. and Stadlan, E. M.** (2011). Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's disease. *Neurology*, 77, 939–939. doi:10.1212/01.wnl.0000400650.92875.cf.
- Morris, J. C.** (1997). Clinical Dementia Rating: a reliable and valid diagnostic and staging measure for dementia of the Alzheimer type. *International Psychogeriatrics*, 9, 173–176. doi:10.1017/S1041610297004870.
- Mortby, M. E., Maercker, A. and Forstmeier, S.** (2012). Apathy: a separate syndrome from depression in dementia? A critical review. *Aging Clinical and Experimental Research*, 24, 305–316. doi:10.3275/8105.
- Mougiatis, A. A., Politis, A., Lyketos, C. G. and Mavreas, V. G.** (2011). Quality of life in dementia patients in Athens, Greece: predictive factors and the role of caregiver-related factors. *International Psychogeriatrics*, 23, 395–403. doi:10.1017/s1041610210001262.
- Pinquart, M. and Sörensen, S.** (2003). Associations of stressors and uplifts of caregiving with caregiver burden and depressive mood: a meta-analysis. *Journal of Gerontology: Psychological Sciences*, 58B, 112–128. doi:10.1093/geronb/58.2.P112.
- Radloff, L. S.** (1977). The CES-D scale. *Applied Psychological Measurement*, 1, 385–401. doi:10.1177/014662167700100306.
- Ready, R. E., Ott, B. R. and Grace, J.** (2004). Patient versus informant perspectives of Quality of life in mild cognitive impairment and Alzheimer's disease. *International Journal of Geriatric Psychiatry*, 19, 256–265. doi:10.1002/gps.1075.
- Rosenberg, P. B., Mielke, M. M. and Lyketos, C. G.** (2005). Caregiver assessment of patients' depression in Alzheimer disease: longitudinal analysis in a drug treatment study. *American Journal of Geriatric Psychiatry*, 13, 822–826.
- Sands, L., Ferreira, P., Stewart, A., Brod, M. and Yaffe, K.** (2004). What explains differences between dementia patients' and their caregivers' ratings of patients' quality of life? *American Journal of Geriatric Psychiatry*, 12, 272–280.
- Schatzberg, K. L., Sudha, S. and Mutran, E. J.** (1998). Elder-proxy agreement concerning the functional status and medical history of the older person: the impact of caregiver burden and depressive symptomatology. *Journal of the American Geriatrics Society*, 46, 1103–1111.
- Schoenmakers, B., Buntinx, F. and Delepeleire, J.** (2010). Factors determining the impact of care-giving on caregivers of elderly patients with dementia. A systematic literature review. *Maturitas*, 66, 191–200. doi:10.1016/j.maturitas.2010.02.009.

- Schulz, R. et al.** (2013). Magnitude and causes of bias among family caregivers rating Alzheimer disease patients. *American Journal of Geriatric Psychiatry*, 21, 14–25. doi:10.1016/j.jagp.2012.10.002.
- Teri, L. and Truax, P.** (1994). Assessment of depression in dementia patients: Association of caregiver mood with depression ratings. *The Gerontologist*, 34, 231–234. doi:10.1093/geront/34.2.231.
- Winblad, B. et al.** (2004). Mild cognitive impairment – beyond controversies, towards a consensus: report of the International Working Group on Mild Cognitive Impairment. *Journal of Internal Medicine*, 256, 240–246. doi:10.1111/j.1365-2796.2004.01380.x.
- Yesavage, J. A. et al.** (1983). Development and validation of a geriatric depression screening scale: a preliminary report. *Journal of Psychiatry Research*, 17, 37–49. doi:10.1016/0022-3956(82)90033-4.
- Zanetti, O. et al.** (1999). Contrasting results between caregiver's report and direct assessment of activities of daily living in patients affected by mild and very mild dementia: the contribution of the caregiver's personal characteristics. *Journal of the American Geriatrics Society*, 47, 196–202.
- Zarit, S. H., Reever, K. E. and Bach-Peterson, J.** (1980). Relatives of the impaired elderly: correlates of feelings of burden. *Gerontologist*, 20, 649–655. doi:10.1093/geront/20.6.649.