Multifaceted shared care intervention for late life depression in residential care: randomised controlled trial

Robert H Llewellyn-Jones, Karen A Baikie, Heather Smithers, Jasmine Cohen, John Snowdon, Chris C Tennant

Editorial by Haynes

Department of Psychological Medicine, University of Sydney, New South Wales 2006, Australia

Robert H Llewellyn-Jones *lecturer*

Karen A Baikie *senior research officer*

Heather Smithers *research officer*

Jasmine Cohen *data manager*

John Snowdon *clinical associate professor*

Chris C Tennant *professor*

Correspondence to: R H Llewellyn-Jones, Healthy Aging Research Unit, Hornsby Ku-ring-gai Hospital, Hornsby, New South Wales 2077, Australia rljones@mail.usyd.edu.au

BMJ 1999;319:676–82

Abstract

Objective To evaluate the effectiveness of a population based, multifaceted shared care intervention for late life depression in residential care.
Design Randomised controlled trial, with control and intervention groups studied one after the other and blind follow up after 9.5 months.
Setting Population of residential facility in Sydney living in self care units and hostels.
Participants 220 depressed residents aged ⩾65 without severe cognitive impairment.
Intervention The shared care intervention included: *(a)* multidisciplinary consultation and collaboration, *(b)* training of general practitioners and carers in detection and management of depression, and *(c)* depression related health education and activity programmes for residents. The control group received routine care.
Main outcome measure Geriatric depression scale.
Results Intention to treat analysis was used. There was significantly more movement to "less depressed" levels of depression at follow up in the intervention than control group (Mantel-Haenszel stratification test, P = 0.0125). Multiple linear regression analysis found a significant intervention effect after controlling for possible confounders, with the intervention group showing an average improvement of 1.87 points on the geriatric depression scale compared with the control group (95% confidence interval 0.76 to 2.97, P = 0.0011).
Conclusions The outcome of depression among elderly people in residential care can be improved by multidisciplinary collaboration, by enhancing the clinical skills of general practitioners and care staff, and by providing depression related health education and activity programmes for residents.

website *e*xtra

The sample calculation and flow of participants through the trial appear on the BMJ's website

www.bmj.com

Introduction

Large numbers of depressed elderly people live in residential care. Their depression is often chronic,[1] unrecognised,[1] and associated with significant disability[2] and premature mortality.[3 4] Few depressed elderly people living in residential care receive appropriate management.[1 2 5]

Various specific interventions are effective for late life depression in residential care.[6–10] Effective models of delivering these interventions (which recognise the scarcity of psychogeriatric resources in many countries) are, however, lacking. Although there are descriptive accounts of psychogeriatric service provision in residential care,[11 12] only one study[1] focused on depression. It evaluated a service model in which a psychiatrist visited homes regularly and recommended a range of interventions to be carried out by general practitioners and care staff. But the interventions proved difficult to implement and the study was not a randomised controlled trial, therefore definite conclusions about efficacy were not possible.

We aimed to systematically overcome the following barriers to the care of late life depression: inadequate detection and management by general practitioners,[2 13] variable cooperation of general practitioners and care staff in intervention programmes,[1] the reluctance of elderly depressed people to seek help,[14] poor social environments,[1] and underutilisation of psychosocial interventions.[1] To achieve this aim we considered it necessary to change the care culture (including usual care practices) of the residential facility under study as well as the population's culture as a whole (including help seeking behaviour and compliance with treatment). We therefore chose a population based intervention because the health of individuals is profoundly influenced by the social characteristics and culture of the community in which they live.[15] In addition, we adopted a multifaceted shared care approach to deal with the complexity of late life depression[16 17] in the context of limited specialist services and to maximise the potential for synergy between different elements of the intervention.

Although multifaceted interventions for depression are more effective than routine general practitioner care for elderly people living in the community,[18 19] no randomised controlled trials of such interventions for late life depression in residential care have been reported. We evaluated the effectiveness of a population based, multifaceted shared care intervention for late life depression by comparing it with routine care in a randomised controlled trial, using multiple linear regression.

Participants and methods

Design and randomisation

Our study was conducted at a large residential facility in Sydney where residents live in self care units, hostels, and nursing homes. A computerised register of residents allowed us to randomly allocate the entire non-nursing home population (1466 people) into two groups, using computer generated random numbers. Randomisation was stratified to ensure that the groups were matched for the proportions of residents managed by specific general practitioners.

We conducted assessments at baseline and at 9.5 months follow up for each group. Because the intervention was implemented for the entire non-nursing home population and their carers, the groups were studied serially to ensure that the control group did not receive the intervention. The control group was monitored first while the entire population received routine care. The intervention was then implemented and the intervention group monitored, with their baseline measures being gathered just before implementation. Baseline measures for both groups were gathered in winter thereby controlling for seasonal factors.

All measures were gathered by trained interviewers. To minimise observer bias, baseline and outcome ratings were made by different interviewers who were not associated with the intervention and were blind to baseline results. To minimise potential bias from treatment expectancy and the Hawthorne effect, participants, general practitioners, other carers, and research interviewers were unaware that it was an intervention study.

Recruitment

All English speaking residents aged 65 or over and cognitively able to provide accurate information (brief orientation-memory-concentration[20] error score less than 20) were screened for depression. Inclusion criteria for study participation were geriatric depression scale[21] score ⩾10 and the absence of pronounced cognitive impairment (mini mental state examination[22] score ⩾18). Exclusion criteria were severe physical illness, or current treatment from a mental health professional for depression or a serious mental illness. Exclusion was determined by interviewers who were blind to the participant's group assignment, and inclusion was independently determined with scores on standardised measures. All subjects gave written informed consent. The study was approved by the hospital ethics committee.

Measure of depression

The main outcome measure was the geriatric depression scale,[21] a valid reliable screening instrument[23 24] sensitive to change[25–27] and strongly recommended for use in geriatric care.[28] Residents scoring ⩾10 on this scale were defined as depressed.[29 30] To measure interrater reliability, all 23 interviewers scored the geriatric depression scale from video recordings of a sample of five interviews.

Other measures

To control for potential confounding variables the following measures were taken and included in the multiple linear regression analysis: cognitive function (brief orientation-memory-concentration test[20] and mini mental state examination[22]); physical health (adapted from Belloc et al[31]); frequency of general practitioner visits and admissions to hospital; demographic characteristics (age, sex, hostel versus independent unit accommodation, marital status, socioeconomic status[32]); social support (adapted from Henderson et al[33]), alcohol use, previous history of depression; functional status (instrumental activities of daily living[34] and physical self maintenance scale[35]); extroversion and neuroticism (Eysenck personality questionnaire-16; Jorm et al[36] and unpublished data) developed from the Eysenck personality inventory[37]); drug usage; acute and chronic adverse life events (life event and difficulties schedule[38]); help seeking behaviour; number of weeks between baseline and follow up geriatric depression scale; and level of exposure to the intervention.

Key elements of the shared care intervention

Removing barriers to care

- Promoted holistic coordinated health care through multidisciplinary collaboration
- Care primarily delivered by general practitioners and residential staff, with specialist help available
- General practitioner, resident, staff, local psychogeriatric service, and project team representatives met regularly to ensure programme feasibility and acceptability
- Improved general practitioner and staff communication through monthly liaison committee meetings

Carer education

- Practical, case based education enhanced pre-existing skills and promoted both psychosocial treatments and antidepressants at adequate dose and duration
- "Insights" interactive workshops for general practitioners on assessing and managing depression and related comorbid illness
- Depression education and support for staff from a specialist psychogeriatric nurse

Health education and health promotion

- Marketed as "healthy ageing" to minimise stigma
- Encouraged residents to recognise depression, seek help, and attend positive activities
- Bimonthly newsletter (*Bright Horizons*) combating misconceptions about depression and its treatment sent to all residents, general practitioners, and staff (also on audiotape)
- Activities included graded gentle exercise classes and talks on depression, chronic pain, relaxation and stress management, arthritis, osteoporosis, and prevention of falls
- Volunteer programme to provide emotional support and assist frail, isolated, depressed residents to participate in activities

Intervention

The box sets out the key elements of the intervention. The main issues that the intervention addressed were: *(a)* increasing the detection rate of depression by carers; *(b)* getting elderly people to accept that depression is treatable; and *(c)* providing accessible treatment programmes in residential care. This innovative intervention was unique in combining the elements of

Table 1 Baseline demographic and clinical characteristics of control and intervention groups. Values are numbers (percentages) of participants unless stated otherwise

Baseline measure	Control group (n=111)	Intervention group (n=109)
Demographic		
Women	95 (86)	90 (83)
Hostel dwelling	70 (63)	74 (68)
Currently married	12 (11)	10 (9)
Widowed	74 (67)	78 (72)
Mean (SD) age (years)	83.8 (5.7)	84.9 (5.9)
Clinical		
Mean (SD) geriatric depression scale score	13.5 (3.4)	13.5 (3.2)
Mean (SD) brief orientation-memory-concentration score	5.21 (4.6)	5.34 (4.8)
Mean (SD) mini mental state examination score	26.9 (2.7)	29.3 (2.7)
Taking antidepressants	19 (17)	9 (8)
Mean (SD) daily dose of antidepressant (mg)*	37.9 (16.2)	44.8 (44.6)
Mean (SD) Belloc score†	2.11 (1.1)	1.87 (1.0)
Mean (SD) pain score‡	1.89 (1.6)	2.01 (1.6)

*Expressed as equivalent daily dose of amitriptyline (mg), converted using drug dose equivalent tables.[40]
†Modified Belloc scale[31]: 1 (severe physical disability) to 4 (no disability).
‡Likert scale of frequency of recurring pain in past 6 months: 0 (not at all) to 4 (constantly).

Table 2 Number (percentage) of participants classified at each geriatric depression scale depression level at follow up by that at baseline for control and intervention groups

Follow up depression level	Baseline depression level			
	Control group (n=83)		Intervention group (n=86)	
	Mild†	Moderate-severe‡	Mild†	Moderate-severe‡
Non-depressed*	19 (34)	1 (4)	23 (48)	6 (16)
Mild†	20 (36)	6 (22)	18 (37)	10 (26)
Moderate-severe‡	17 (30)	20 (74)	7 (15)	22 (58)
Total	56 (100)	27 (100)	48 (100)	38 (100)

Geriatric depression scale score: *0-9[29 42 43]; †10-13; ‡≥14.

carer education with health education and health promotion for the entire population of a large residential facility. The control group received routine care. Those carrying out the intervention were not told which residents were depressed and being evaluated, and depressed residents were not informed that they had been identified.

Statistical analysis

We carried out intention to treat analyses. A Mantel-Haenszel stratification test[39] compared the level of depression of the control and intervention groups at follow up after taking into account baseline levels, using the geriatric depression scale. An independent two sample *t* test was used to compare change in geriatric depression scale score between the groups. We used multiple linear regression analysis to evaluate the effect of the intervention on geriatric depression scale score at follow up, while controlling for the other independent variables measured. Baseline geriatric depression scale score and then group status (control *v* intervention) were forced in to the model first, followed by all other independent variables using forward stepwise entry.

The effect of the intervention on other clinical outcome measures was assessed with analysis of covariance for continuous variables and logistic regression for categorical variables, in each case looking at the effect of group membership on the follow up measure while controlling for the baseline measure as a covariate.

All analyses were planned a priori. All statistical tests used an α level of 0.05 and two sided hypothesis testing, and 95% confidence intervals were calculated for differences in change of scores or proportions. Analyses were carried out with SPSS for Windows (release 6.0).

Table 3 Mean (SD) geriatric depression scale scores and geriatric depression scale change scores for control and intervention groups

Variable	Control group (n=83)	Intervention group (n=86)
Geriatric depression scale score		
Baseline	13.18 (3.3)	13.38 (3.0)
Follow up	12.57 (4.1)	11.81 (4.7)
Change*	–0.61 (3.6)	–1.57 (3.7)

*Follow up score minus baseline score (negative score indicates improvement).

Results

Participant flow and follow up

Details of participant flow through the study are included on the *BMJ*'s website. The mean interval between baseline and follow up was 40.9 (SD 3.3) weeks. Interrater reliability of the geriatric depression scale was very high (intraclass correlation coefficient 0.996).

Sample characteristics

Table 1 shows the key baseline characteristics of the groups. We found no significant differences in the distribution of these variables between groups, although more control than intervention residents were taking antidepressants at baseline ($\chi^2 = 3.89$, df = 1, P = 0.049). As recommended,[41] the possible prognostic effect of any imbalance in baseline variables was taken into account in the multiple linear regression analysis. We found no significant differences on any key baseline characteristics between participants who completed the geriatric depression scale at follow up and those who dropped out. The 169 residents who completed follow up were cared for by 34 general practitioners of whom 26 (76%) were male and 27 (79%) worked in group practices.

Depression outcome

A Mantel-Haenszel stratification test[39] showed a significant difference between control and intervention groups ($\chi^2 = 6.37$, df = 1, P = 0.012) indicating significantly more movement to "less depressed" geriatric depression scale levels in the intervention group (table 2).

Table 3 shows the mean (SD) geriatric depression scale scores. Before adjusting for any covariates, the intervention group improved more than the control group, with the change in mean geriatric depression scale score between baseline and follow up approaching significance: mean difference 0.96 points (−0.15 to 2.06, $t = 1.70$, df = 167, P = 0.090).

We performed multiple linear regression analysis on 133 participants who had completed all measures at baseline and follow up. The multiple linear regression model was significant (F = 24.9, df = 5, 127, P < 0.001) and explained almost 50% of the variance in outcome scores on the geriatric depression scale ($r^2 = 0.495$, adjusted $r^2 = 0.475$). The model satisfied all necessary assumptions of multiple linear

regression: normality; homoscedasticity; linearity; and independence of error terms. Lower geriatric depression scale score at follow up was significantly associated with: lower geriatric depression scale score at baseline; being in the intervention group; lower neuroticism score; higher basic functional ability at baseline; and younger age (table 4).

We included interaction terms for each independent variable by group status in the regression, but none entered the final model. This suggests that the intervention did not have a differential effect on any subgroups of depressed residents when all other independent variables were controlled, although the power of the study to detect such effects is low.

After controlling for geriatric depression scale score at baseline, group status had the most significant effect on geriatric depression scale score at follow up even after taking into account all other variables thought to affect depression outcome. The measure of the intervention effect, after allowing for geriatric depression scale score at baseline and other significant independent variables, was an average improvement of 1.87 points on geriatric depression scale score at follow up compared with the control group. The final model was rerun with only the five significant variables on all 159 participants for whom these data were available. All five variables remained significant with comparable orders of magnitude, and the conclusions remain the same.

The multivariate and univariate analyses produce discrepant estimates of the intervention effect. The univariate analysis was based on all 169 participants who completed the geriatric depression scale at follow up whereas the multivariate analysis excluded 36 participants who had incomplete data on some independent variables. The difference between the analyses is therefore due to a combination of controlling for other variables in the regression and the reduced sample size. Although there were no significant differences on any key baseline characteristics between the 133 participants with complete data and the 36 participants who completed the geriatric depression scale but not all measures, we repeated the *t* test on the 133 participants to adjust for the possibility that they differed in some undetected way (mean difference 1.41, 0.27 to 2.56, $t=2.44$, df = 131, P = 0.016). Therefore, if there is any bias due to assuming that the 133 participants are representative of the whole group, this bias is estimated as 0.45 (1.41 – 0.96). If this is removed from the multivariate estimate of effect the adjusted intervention effect is likely to be around 1.42 (1.87 – 0.45), which remains significant (P = 0.012). Given this, the evidence for a significant intervention effect is robust.

Table 4 Factors affecting geriatric depression scale score at follow up, determined by multiple linear regression analysis

Variable	Regression coefficient (95% CI)	Standardised regression coefficient (β)	P value
Baseline geriatric depression scale score	0.73 (0.56 to 0.91)	0.56	0.0000
Group status*	–1.87 (–2.97 to –0.76)	–0.22	0.0011
Neuroticism†	0.55 (0.20 to 0.90)	0.21	0.0021
Physical maintenance scale score‡	–0.54 (–0.99 to –0.09)	–0.15	0.0202
Age (years)	0.10 (0.00 to 0.19)	0.13	0.0395

*Control *v* intervention group.
†Scale from Eysenck personality questionnaire-16 (Jorm et al[36] and unpublished data) developed from Eysenck personality inventory[37] (scored 0-8: higher score indicates higher neuroticism).
‡Measure of basic functional ability, for example, to dress or feed oneself independently (scored 0-8: higher score indicates higher independence).

Table 5 Adjusted means for control and intervention groups and adjusted difference* for other clinical outcome measures

Outcome measure	Control group (n=64)	Intervention group (n=69)	Adjusted difference (95% CI)	P value
Social support†	5.38	5.92	–0.54 (–1.01 to –0.06)	0.03
Daily dose of antidepressant (mg)‡	37.9	57.1	–19.2 (–51.2 to 12.8)	0.20
Number of depressogenic drugs§	0.44	0.47	–0.03 (–0.19 to 0.13)	0.69

*Means at follow up and differences adjusted for baseline measure using analysis of covariance.[44]
†Scale adapted from Henderson et al[33] (0-8: higher score indicates more social support).
‡Dose expressed as equivalent daily dose of amitriptyline (mg), converted using drug dose equivalent tables.[40] Mean dose only for 11 participants taking antidepressants at both baseline and follow up.
§Potentially depressogenic drugs, as defined by Callahan et al.[45]

Other clinical outcomes

Social support increased significantly in the intervention compared with the control group. Despite general practitioner education the intervention did not significantly increase the mean daily dose of antidepressants or reduce the number of depressogenic drugs taken compared with routine care (table 5). Logistic regression showed that intervention participants were more likely to be taking antidepressants at follow up than controls, taking into account whether participants were taking antidepressants at baseline (odds ratio 3.1, 0.9 to 10.2, P = 0.066).

We investigated whether the intervention effect was due to residents not taking antidepressants at baseline being prescribed antidepressants by follow up. We used analysis of covariance on geriatric depression scale score at follow up, with the score at baseline as covariate, then antidepressant use at follow up and group membership was entered hierarchically. The antidepressant by group interaction was not significant (F = 0.18, df = 1, 111, P = 0.70) nor was follow up antidepressant use (F = 0.48, df = 1, 111, P = 0.49), but the group main effect remained significant (F = 8.17, df = 1, 111, P = 0.005). The intervention effect is not simply an antidepressant effect.

Discussion

Shared care was statistically more effective than routine care, after controlling for possible confounders, with an average improvement of 1.87 points on the geriatric depression scale in the intervention compared with the control group. Significantly more movement to "less depressed" geriatric depression scale levels at outcome was found in the intervention group and evidence that the intervention helped prevent mild depression from becoming worse.

Clinical significance

Although the intervention effect was modest we believe that it is clinically significant. The intervention was population based: general practitioners and other carers were not told which residents had been identified as depressed, so the positive result may reflect an improvement in both detection and management of depression. The intervention was naturalistic: although all residents and their carers were invited to fully participate, participation was variable. Only 53 (62%) of 86 intervention group participants had general

practitioners who attended the education programme, and only 21 (28%) of the 74 participants for whom data were available attended exercise classes. The intervention's full implementation was also delayed by management. Given these factors, the results are probably conservative, and a fully implemented intervention may provide more substantial outcomes.

Our study has methodological strengths, which increase the generalisability of the findings. It is the first randomised controlled trial in a priority research area.[13 46 47] We used a reliable and valid measure of depression, blinded assessments, and a large sample. We used intention to treat analysis in that the outcome of all participants was examined regardless of the extent to which they or their carers participated in the intervention. The positive finding remained after controlling for an extensive list of possible confounders. The only pertinent previous trial yielded a negative result.[1]

Possible limitations

The design of this population based randomised controlled trial is unusual in certain respects. Because the intervention was population based, it was not possible to conduct a classic randomised controlled trial with concurrent controls within the single facility as there was no way to prevent contamination of the control group. Although studying the populations of two separate facilities (control *v* intervention) would have enabled the use of concurrent controls, it would have been difficult to adequately control for differences between the facilities in available resources, care cultures, and the characteristics of the populations of residents and carers. Such bias may have been eliminated in a multicentre study using the residential facility as the unit of randomisation but this was beyond our resources. We decided against a classic "before and after" evaluation with participants as their own historical controls because it meant more interviews for participants over a longer time period, likely to result in greater loss to follow up. We therefore adopted a single site design, and studied control and intervention groups one after the other, given the difficulties associated with alternative designs. This serial design necessitated first randomising the population and then selecting the participants, but this is very nearly equivalent to the conventional method of first selecting the participants and then randomising them. Although the design is not entirely typical of a randomised controlled trial, both a control group and the principle of randomisation were used.

We acknowledge that this design has certain limitations. Its serial nature introduces the possibility of secular confounding, since the groups are studied over a different period of time. The delay between randomisation and the start of treatment in the intervention group does not, however, seem to have introduced bias as there are no differences at baseline on any key variables between participants in the control and intervention groups. Baseline measures for both groups were collected in winter thereby reducing the impact of seasonal factors. The age difference between the control and intervention groups caused by the serial design was not statistically significant. The possible confounding effect of age differences was addressed by including age in the regression analysis.

Key messages

- Large numbers of depressed elderly people live in residential care but few receive appropriate management
- A population based, multifaceted shared care intervention for late life depression was more effective than routine care in improving depression outcome
- The outcome of late life depression can be improved by enhancing the clinical skills of general practitioners and care staff and by providing depression related health education and activity programmes for residents
- The intervention needs further refining and evaluation to improve its effectiveness and to determine how best to implement it in other residential care settings

Generalisability may be limited because only one large residential facility was studied. However, although the facility's size was atypical the population and available resources were not. Psychogeriatric resources were scarce, staff to resident ratios were low, and the needs of residents were great. Since this was not a highly expensive intervention, replete with mental healthcare resources, the results have applicability to other settings. However, since severely cognitively impaired people were excluded the results are not applicable to patients with depression and significant dementia.

It is difficult to know how far generalisability is limited by non-response bias at the initial intake interviews and by losing participants to follow up. Refusal rates were low (21% or less), and although follow up rates were only moderately high (at least 75% for the geriatric depression scale and 58% for all measures), there were no significant differences between those who completed the study and those who dropped out. Although fewer intervention than control group residents were eligible to participate (see website), due in part to their greater attrition over the longer time period between randomisation and intake measures, this does not seem to have introduced bias since there are no differences between participants in the control and intervention groups at baseline. The influence of key baseline prognostic factors was taken into account in the regression analysis. Therefore, the extent of such bias is probably limited.

Implications

The intervention was primarily delivered by busy general practitioners and overstretched residential care staff. But by using existing resources in a more effective fashion a significant result was obtained.

By design, we evaluated the model of care as a whole giving limited consideration to the impact of the intervention's individual elements. The result, however, does not seem to be a function of increasing the use or dose of antidepressants, reducing the use of depressogenic drugs, or increasing social support. Further research is needed to determine the relative impact of different elements of the intervention and whether

improved outcomes would occur with increased participation.

Late life depression is common in residential care and its prognosis is usually poor.[1] The results of this trial argue in favour of further refining and evaluating this form of intervention to improve its effectiveness and to determine how best to implement it in other residential settings. The promising results of our trial also have wider implications. Population based interventions to improve mental health are uncommon but their potential benefits are great.[15] They have the potential to shift the population's entire distribution of depression scores in a favourable direction and hence exert a very powerful effect on reducing the burden of disease due to depression.[15] Our results provide encouragement for further research to determine the impact of such interventions not only on those identified as depressed at the outset but also on the population as a whole.

We thank the residents, general practitioners, residential care management and staff, and the research interviewers and staff of the Aged Care Psychiatry Service of Hornsby Ku-ring-gai Hospital. Dr Dimity Pond and Ms Carol Andrews helped initiate the study and, together with Dr Simon Willcock, assisted in the design and implementation of the intervention and commented on the research design. Mrs Sally Castell and Mrs Anne Baikie played key roles in developing the health education and promotion elements of the intervention and in coordinating the implementation of the intervention as a whole. Dr Janet Beaurepaire advised on and scored life events measures. Professors Geoffrey Berry and Stewart Dunn provided statistical advice, and associate professor David Ames provided comments on earlier versions of this manuscript.

Contributors: RHL-J, the principal investigator, initiated and designed the study, designed the multifaceted intervention and supervised its implementation, contributed to data analysis and interpretation, and oversaw the study as a whole. KAB contributed to the study design, discussed core ideas, designed data collection protocols, assisted in coordinating data collection, and guided and conducted the statistical analysis and interpretation of the data. HS contributed to the study design, discussed core ideas, designed data collection protocols, coordinated data collection, commented on data analysis, and assisted with data interpretation. JC helped design data collection protocols, discussed core ideas, was the data manager, analysed data, and contributed to data interpretation. JS assisted with the study design, contributed to the design of the multifaceted intervention, and discussed core ideas. CCT advised on and assisted with the study design, advised on research measures, and discussed core ideas. RHL-J and KAB jointly wrote, revised, and edited the paper. HS, JC, JS, and CCT made contributions to earlier drafts. All authors gave final approval for the paper. RHL-J and KAB will act as guarantors for the paper.

Funding: The project was initially conducted while the first author was a research fellow of the New South Wales Institute of Psychiatry. The study received grant aided support from the Commonwealth Department of Health and Family Services, General Practice Evaluation Program and its Aged Care Program Support Program, the University of Sydney McGeorge bequest, the New South Wales Health Department Health Outcomes Program, and a Roche Products clinical research grant. An educational grant from Roche Products was used to devise and produce the materials for the "Insights" general practitioner education workshop on late life depression. This workshop was the third in a series of six general practitioner workshops on depression developed by the "Insights" Depression Education Advisory Committee chaired by associate professor John Tiller. However, only "Insights workshop 3: depression in the elderly" was part of the intervention. The content of this workshop was devised by a project group (chaired by RHL-J) of the advisory committee in association with both Oxford Clinical Communications Sydney and the whole committee.

Competing interests: RHL-J and JS received small honoraria from Roche Products for attending "Insights" advisory committee meetings. RHL-J was also reimbursed by Roche Products for attending a symposium. KAB's work on the study was part funded by a Roche Products clinical research grant. Neither the "Insights" general practitioner education meetings nor the multifaceted intervention as a whole promoted the use of any specific brand of pharmaceutical product.

1 Ames D. Depression among elderly residents of local-authority residential homes: its nature and the efficacy of intervention. *Br J Psychiatry* 1990;156:667-75.
2 Ames D. Depressive disorders among elderly people in long-term institutional care. *Aust NZ J Psychiatry* 1993;27:379-91.
3 Rovner BW, German PS, Brant LJ, Clark R, Burton L, Folstein MF. Depression and mortality in nursing homes. *JAMA* 1991;265:993-6.
4 Murphy E, Smith R, Lindesay J, Slattery J. Increased mortality rates in late-life depression. *Br J Psychiatry* 1988;152:347-53.
5 Heston LL, Garrard J, Makris L, Kane RL, Cooper S, Dunham T, et al. Inadequate treatment of depressed nursing home elderly. *J Am Geriatr Soc* 1992;40:1117-22.
6 Katz IR, Simpson GM, Curlik SM, Parmelee PA, Muhly C. Pharmacologic treatment of major depression for elderly patients in residential care settings. *J Clin Psychiatry* 1990;51:41-7.
7 McMurdo MET, Rennie L. A controlled trial of exercise by residents of old people's homes. *Age Ageing* 1993;22:11-5.
8 Rattenbury C, Stones MJ. A controlled evaluation of reminiscence and current topics discussion groups in a nursing home context. *Gerontologist* 1989;29:768-71.
9 Fernandez-Ballesteros R, Izal M, Diaz P, Gonzalez JL, Souto E. Training of conversational skills with institutionalized elderly: a preliminary study. *Percept Mot Skills* 1988;66:923-6.
10 Rosen J, Rogers JC, Martin RS, Mulsant BH, Shahar A, Reynolds CF. Control-relevant intervention in the treatment of minor and major depression in a long-term care facility. *Am J Geriatr Psychiatry* 1997;5:247-57.
11 Stolee P, Kessler L, Le Clair JK. A community development and outreach program in geriatric mental health: four years' experience. *J Am Geriatr Soc* 1996;44:314-20.
12 Green S, Girling DM, Lough S, Ng AMN, Whitcher SK. Service provision for elderly people with long-term functional illness. *Psychiatr Bull* 1997;21:353-7.
13 Lebowitz BD, Pearson JL, Schneider LS, Reynolds CFI, Alexopoulos GS, Bruce ML, et al. Diagnosis and treatment of depression in late life: consensus statement update. *JAMA* 1997;278:1186-90.
14 Waxman HM, Carner EA, Klein M. Underutilization of mental health professionals by community elderly. *Gerontologist* 1984;24:23-30.
15 Rose G. *The strategy of preventive medicine*. Oxford: Oxford University Press, 1992.
16 Morris RG, Morris LW. Cognitive and behavioural approaches with the depressed elderly. *Int J Geriatr Psychiatry* 1991;6:407-13.
17 Callahan CM, Hui SL, Nienaber NA, Musick BS, Tierney WM. Longitudinal study of depression and health services use among elderly primary care patients. *J Am Geriatr Soc* 1994;42:833-8.
18 Blanchard MR, Waterreus A, Mann AH. The effect of primary care nurse intervention upon older people screened as depressed. *Int J Geriatr Psychiatry* 1995;10:289-98.
19 Banerjee S, Shamash K, Macdonald AJD, Mann AH. Randomised controlled trial of effect of intervention by psychogeriatric team on depression in frail elderly people at home. *BMJ* 1996;313:1058-61.
20 Katzman R, Brown T, Fuld P, Peck A, Schechter R, Schimmel H. Validation of a short orientation-memory-concentration test of cognitive impairment. *Am J Psychiatry* 1983;140:734-9.
21 Yesavage JA, Brink TL, Rose TL, Lum O, Huang V, Adey M, et al. Development and validation of a geriatric depression screening scale: a preliminary report. *J Psychiatr Res* 1983;17:37-49.
22 Folstein MF, Folstein SE, McHugh PR. "Mini-mental state:" a practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975;12:189-98.
23 Montorio I, Izal M. The geriatric depression scale: a review of its development and utility. *Int Psychogeriatr* 1996;8:103-12.
24 Parmelee PA, Lawton MP, Katz IR. Psychometric properties of the geriatric depression scale among the institutionalized aged. *Psychol Assess* 1989;1:331-8.
25 Schifano F, Garbin A, Renesto V, De Dominicis MG, Triniarelli G, Silvestri A, et al. A double-blind comparison of mianserin and maprotiline in depressed medically ill elderly people. *Acta Psychiatr Scand* 1990;81:289-94.
26 Mossey JM, Knott KA, Higgins M, Talerico K. Effectiveness of a psychosocial intervention, interpersonal counseling, for subdysthymic depression in medically ill elderly. *J Gerontol A Biol Sci Med Sci* 1996;51A:172-8M.
27 Hooper SC, Vaughan KJ, Tennant CC, Perz JM. Major depression and refusal of life-sustaining medical treatment in the elderly. *Med J Aust* 1996;165:416-9.
28 Katona CLE. *Depression in old age*. Chichester: Wiley, 1994.
29 Spreen O, Strauss E. *A compendium of neuropsychological tests: administration, norms and commentary*. New York: Oxford University Press, 1991.
30 McGivney SA, Mulvihill M, Taylor B. Validating the GDS depression screen in the nursing home. *J Am Geriatr Soc* 1994;42:490-2.
31 Belloc NB, Breslow L, Hochstim JR. Measurement of physical health in a general population survey. *Am J Epidemiol* 1971;93:328-36.
32 Daniel A. *Power, privilege and prestige: occupations in Australia*. Melbourne: Longman Cheshire, 1983.
33 Henderson S, Byrne DG, Duncan-Jones P. *Neurosis and the social environment*. Sydney: Academic Press, 1981.
34 Lawton MP, Moss M, Fulcomer M, Kleban MH. A research and service oriented multilevel assessment instrument. *J Gerontol* 1982;37:91-9.

35 Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. *Gerontologist* 1969;9:179-86.
36 Jorm AF, Duncan-Jones P, Scott R. An analysis of the re-test artefact in longitudinal studies of psychiatric symptoms and personality. *Psychol Med* 1989;19:487-93.
37 Eysenck HJ, Eysenck SBG. *Manual of the Eysenck personality inventory*. London: London University Press, 1964.
38 Brown GW, Harris TO, eds. *Life events and illness*. New York: Guildford Press, 1989.
39 Armitage P, Berry G. *Statistical methods in medical research*, 3rd ed. Oxford: Blackwell Scientific, 1994.
40 Victorian Medical Postgraduate Foundation. *Psychotropic drug guidelines 1996-1997*, 3rd ed. Melbourne: VMPF Therapeutics Committee, 1995.
41 Altman DG. Comparability of randomised groups. *Statistician* 1985;34:125-36.
42 Ouslander JG, Osterweil D, Morley J. *Medical care in the nursing home*. New York: McGraw-Hill, 1991.
43 Flaherty JH, McBride M, Marzouk S, Miller DK, Chien N, Hanchett M, et al. Decreasing hospitalization rates for older home care patients with symptoms of depression. *J Am Geriatr Soc* 1998;46:31-8.
44 Howell DC. *Statistical methods for psychology*, 2nd ed. Boston: PWS-KENT Publishing, 1987.
45 Callahan CM, Hendrie HC, Dittus RS, Brater DC, Hui SL, Tierney WM. Improving treatment of late life depression in primary care: a randomized clinical trial. *J Am Geriatr Soc* 1994;42:839-46.
46 Cassel CK. Researching the health needs of elderly people. [Editorial.] *BMJ* 1994;308:1655-6.
47 NIH Consensus Development Panel on Depression in Late Life. Diagnosis and treatment of depression in late life. *JAMA* 1992;268:1018-24.

(Accepted 4 August 1999)

Commentary: Beyond the boundary for a randomised controlled trial?

Jonathan J Deeks, Edmund Juszczak

Imperial Cancer Research Fund and NHS Centre for Statistics in Medicine, Institute of Health Sciences, Headington, Oxford OX3 7LF
Jonathan J Deeks *medical statistician*
Edmund Juszczak *medical statistician*
Correspondence to: J J Deeks J.Deeks@icrf.icnet.uk

From the outset, Llewellyn-Jones et al abandoned several key principles of randomised controlled trials in their evaluation of multifaceted shared care for later life depression. They randomised before assessing eligibility, there were delays between allocation and starting interventions, both delivery of interventions and assessment of outcomes in control and experimental groups were not concurrent, and only a subgroup of participants who received the intervention were included in the final analysis. The authors defend these study features on grounds of practicality, but how may these deficiencies have affected their findings?

The non-concurrent assessment of control and intervention groups is of most concern. During the duration of a trial there can be notable variations in factors external and intrinsic to the trial. For example, weather and national events can alter psychological state; staff changes and learning curves may cause temporal variations in the assessments of eligibility and outcome. Such factors may be difficult to assess (or even identify) but can introduce systematic patterns in outcome with time. When groups are not treated concurrently any such differences will lead to bias. Trials with non-concurrent enrolment and assessment have been described as uncontrolled, as only the fortuitous absence of these temporally related factors will ensure the comparability of the groups.[1]

To consider an analogy, followers of test cricket (where the competing teams alternate between fielding and batting over several days) will be aware that matches are occasionally lost or won by the random allocation of the batting order when the ground, weather, and light conditions change during the match. Football (or any sport where the two teams compete at the same time) seems to be a fairer game.

The trial's focus was on residents assessed as depressed when screened, who were the only individuals included in the follow up. However, all residents would have been exposed to the intervention at some level since it was delivered to the population as a whole. Although it can be surmised that the effect of the intervention will be most concentrated among those identified as depressed, we cannot be certain whether the intervention has had a similar, more beneficial, or even harmful effect among the rest of the residents.

It seems unlikely that delays in assessments of eligibility and commencement of the interventions will have directly introduced bias. Postponing assessments of eligibility until after randomisation may lead to unequal group sizes (reducing statistical power) but should not introduce systematic bias providing the allocations remain concealed[2] (which is the case reported for this study). In some trials time delays between allocation and commencing interventions cause problems when the eligibility of the participants changes in the intervening period. Delaying the assessment of eligibility will have circumvented this issue in this instance.

The aim of randomisation is to ensure an unbiased assessment of the effect of treatment by making "study groups equivalent in all respects other than the treatment itself."[3] Readers and trialists alike should be aware that randomisation can only produce groups that are comparable at the start of a study; other aspects of good trial design are required to retain the comparability to the end.

Competing interests: None declared.

1 Meinert CL. Clinical trials, overview. In: Armitage P, Colton T, eds. *Encyclopedia of biostatistics*, Vol 1. Chichester: Wiley, 1999;698-713.
2 Schulz K, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled clinical trials. *JAMA* 1995;273:408-12.
3 Wittes J. Randomized treatment assignment. In: Armitage P, Colton T, eds. *Encyclopedia of biostatistics*, Vol 5. Chichester: Wiley, 1999;3703-11.

Endpiece

The doctor's fault

The safest way to health, say what you will,
Is never to suppose we shall be ill;
Most of those ills we poor mortals know
From doctors and imagination flow.

Said by W Dale, MD, London, to be the gist of psychotherapy (*Lancet* 1892;ii:416)

Submitted by Ann Dally, Wellcome Institute for the History of Medicine