Effects of a physiotherapy and occupational therapy intervention on mobility and activity in care home residents: a cluster randomised controlled trial

Catherine M Sackley, Maayken E van den Berg, Karen Lett, Smitaa Patel, Kristen Hollands, Christine C Wright, Thomas J Hoppitt

School of Health and Population Sciences, Primary Care Clinical Sciences, University of Birmingham, Edgbaston, Birmingham B15 2TT

Correspondence to: T J Hoppitt
t.j.hoppitt@bham.ac.uk

Cite this as: *BMJ* 2009;339:b3123
doi: 10.1136/bmj.b3123This article

ABSTRACT

Objective To compare the clinical effectiveness of a programme of physiotherapy and occupational therapy with standard care in care home residents who have mobility limitations and are dependent in performing activities of daily living.

Design Cluster randomised controlled trial, with random allocation at the level of care home.

Setting Care homes within the NHS South Birmingham primary care trust and the NHS Birmingham East and North primary care trust that had more than five beds and provided for people in the care categories "physical disability" and "older people."

Participants Care home residents with mobility limitations, limitations in activities of daily living (as screened by the Barthel index), and not receiving end of life care were eligible to take part in the study.

Intervention A targeted three month occupational therapy and physiotherapy programme.

Main outcome measures Scores on the Barthel index and the Rivermead mobility index.

Results 24 of 77 nursing and residential homes that catered for residents with mobility limitations and dependency for activities of daily living were selected for study: 12 were randomly allocated to the intervention arm (128 residents, mean age 86 years) and 12 to the control arm (121 residents, mean age 84 years). Participants were evaluated by independent assessors blind to study arm allocation before randomisation (0 months), three months after randomisation (at the end of the treatment period for patients who received the intervention), and again at six months after randomisation. After adjusting for home effect and baseline characteristics, no significant differences were found in mean Barthel index scores at six months post-randomisation between treatment arms (mean effect 0.08, 95% confidence interval −1.14 to 1.30; P=0.90), across assessments (−0.01, −0.63 to 0.60; P=0.96), or in the interaction between assessment and intervention (0.42, −0.48 to 1.32; P=0.36). Similarly, no significant differences were found in the mean Rivermead mobility index scores between treatment arms (0.62, −0.51 to 1.76; P=0.28), across assessments (−0.15, −0.65 to 0.35; P=0.55), or interaction (0.71, −0.02 to 1.44; P=0.06).

Conclusions The three month occupational therapy and physiotherapy programme had no significant effect on mobility and independence. On the other hand, the variation in residents' functional ability, the prevalence of cognitive impairment, and the prevalence of depression were considerably higher in this sample than expected on the basis of previous work. Further research to clarify the efficacy of occupational therapy and physiotherapy is required if access to therapy services is to be recommended in this population.

Trial registration ISRCTN79859980

WHAT IS ALREADY KNOWN ON THIS TOPIC?

Care home residents have greater dependence in activities of daily living than do community dwelling elderly people

Care home residents have limited access to physiotherapy and occupational therapy

Research is inconclusive as to whether such therapies are beneficial in this population and, subsequently, a cost effective service

WHAT THIS STUDY ADDS

The three month physiotherapy and occupational therapy intervention delivered did not prove more beneficial than standard care in this sample

The prevalence of mood disorders and cognitive impairment was greater than was previously anticipated

The findings do not to support the argument that such services would be cost effective and reduce burden on care staff and society

INTRODUCTION

Several surveys have found that care home residents in the UK have limited access to rehabilitation services such as physiotherapy and occupational therapy. Evidence for the benefit of rehabilitation services in this population is conflicting and inconclusive.[1-6]

The main objective of this trial was to evaluate the clinical effectiveness of a programme of physiotherapy and occupational therapy against standard care in care home residents with mobility limitations who are dependent on carers in some activities of daily living.

METHODS

We selected 24 of 77 nursing and residential homes from the Birmingham area. Homes were purposefully chosen to encompass variations in geographical location, size, and funding sector. A cluster randomised controlled

This is an abridged version of a paper that was published on bmj.com. Cite this article as: *BMJ* 2009;339:b3123

design was used to randomly allocate care homes to either the intervention arm or the control arm.

Care home staff were asked to screen all residents with the Barthel index of activities of daily living and to provide information on cognitive status.[7] Residents who scored below 5 or over 16 on the Barthel index were excluded.

Residents in the intervention arm received a three month physiotherapy and occupational therapy intervention. The physiotherapy intervention was developed using a modified version of another protocol and the consensus of a steering group of physiotherapists.[8] Therapy was aimed at enhancing mobility and the ability to perform activities of daily living independently, and addressed components such as strength, flexibility, balance, and exercise tolerance. In addition, functional tasks such as bed to chair transfers, sit to stand, and walking or wheeling were practised. The intervention was delivered by two physiotherapists and was adjusted for each individual.

The occupational therapy intervention was developed using the consensus of an occupational therapy steering group.[9] Therapy was targeted towards improving independence in personal activities of daily living such as feeding, dressing, toileting, bathing, and transferring. The intervention was delivered by two occupational therapists who followed a client centred approach, including routine assessment, treatment, and reassessment. The dose, frequency, and duration of both physiotherapy and occupational therapy were dependent on the goals agreed by the individual participant and the therapists and on progress throughout the intervention.

The intervention arm also included an intervention delivered to the care home staff. This involved a programme of staff training to provide practice in promoting residents' independence and the use of therapeutic aids.

Residents in care homes allocated to the control arm continued to receive standard care equal to that received before recruitment to the trial. Occupational therapy was not used routinely by any of the homes and physiotherapy was accessed only via general practitioner referral. The control group received the therapy intervention after the trial had ended.

Outcome measures

Assessments were carried out by two independent assessors blinded to cluster allocation. Assessment were conducted before randomisation (baseline, between July 2004 and July 2005), at three months after randomisation, and at six months after randomisation. The primary outcomes were the scores on the Barthel index and the Rivermead mobility index.[10] Measures of activities of daily living and of mobility were also selected as main outcomes because the intervention was targeted at a population in whom limitations in these parameters are highly prevalent.

In addition, the mini mental state examination was used at the first assessment to determine the level of residents' cognitive impairment (not an exclusion criterion).

Mood was assessed using the hospital anxiety and depression scale depression subscale. Residents unable to complete this had the stroke aphasic depression questionnaire completed by a proxy. Medical history and comorbidity information were also collected.

Data analysis

A 2 point change on the Barthel index was considered to be a meaningful change in independence with respect to activities of daily living and a sample size of 300 was targeted to allow for participant withdrawal. Scores for the Barthel index and the Rivermead mobility index were summarised by treatment arm at each of the three assessments.

Separate multilevel models were used to test the efficacy of the intervention according to each score. Respective centred pre-intervention scores were included in the model as a covariate. Assessment was defined as a repeated measures factor. Study group, assessment (three months post-randomisation and six months post-randomisation), and interaction between the two were modelled as fixed effects. Care home and participants were modelled as random effects. Models with different error structures were fitted. Estimated effects are reported from the model of best fit.

A further analysis was conducted using separate multilevel models to test time standardised area under the curve values for Barthel index and Rivermead mobility index scores across post-intervention and follow-up assessments. Respective centred pre-intervention scores were included in the model as a covariate, study group was modelled as a fixed effect, and care home and participants were modelled as random effects. The same error structure was used as in the "best fit" model above. The estimated value for the intracluster correlation coefficient was computed using pre-intervention scores on the Barthel index and the Rivermead mobility index.

RESULTS

Participants

A total of 24 homes and 249 participants were recruited to the trial. Twelve homes were randomised to each group, with 128 residents allocated to the intervention arm (median number per home=11) and 121 to the control arm (median number per home=8). By the time of the six month follow-up 62 participants had withdrawn (29 in the intervention group and 33 in the control group). See bmj.com.

The most common specific comorbidities were arthritis (56%), stroke (46%), dementia (40%), and diabetes (36%).

Intervention

Out of 128 participants randomised, 123 received physiotherapy and occupational therapy to some degree. The mean number of physiotherapist visits was 6.4 per resident, with an average total contact time of 2.21 hours per resident. The mean number of occupational therapist visits was 9.8 per resident, with an average total contact time of 3.6 hours per resident.

Primary analyses

Primary analyses were conducted on responses from 243 participants who completed pre-randomisation measures: 127 in the intervention group and 116 in the control group. Overall, participants exhibited a low level of independence before randomisation, with mean Barthel index scores of 11.1 and 12.5 in the intervention and control groups, respectively. Furthermore, very low levels of mobility were evident pre-randomisation, with mean Rivermead mobility index scores of 5.8 and 6.9 in the intervention and control groups, respectively.

No statistically significant differences were found between the study groups on mean scores for Barthel index or Rivermead mobility index (adjusting for clusters) at either three months post-randomisation or six months post-randomisation. Intracluster correlation coefficient values of 0.49 and 0.48 were calculated using pre-intervention scores on Barthel index and Rivermead mobility index, respectively.

Once adjusted for home effect and pre-intervention scores, the minimal important difference threshold of 2 index points was not reached for mean Barthel index scores at six months post-randomisation between study groups, across assessments, or for the interaction between intervention and assessment. No differences were statistically significant (table).

Similarly once the scores had been adjusted, the minimal important difference threshold of 3 index points was not reached for mean Rivermead mobility index scores at six months post-randomisation across study groups, across assessments, or for the interaction between intervention and assessment. Again no differences were statistically significant.

In analyses on area under the curve values, no significant differences were found across study groups on Barthel or Rivermead mobility index score.

DISCUSSION

Our results suggest that the three month occupational therapy and physiotherapy programme was not effective in promoting independent living and mobility among care home residents over and above that achieved with standard care. Evidence exists to support occupational therapy and particular aspects of physiotherapy after stroke, but little evidence is available to support more widespread use.[11]

Summary of effects at six months post-randomisation in the multilevel model for Barthel index and Rivermead mobility index, adjusted for home effect and pre-intervention scores

	Outcome measure			
	Barthel index		Rivermead mobility index	
	Estimate (95% CI)	P value	Estimate (95% CI)	P value
Repeated measures analysis				
Intervention	0.08 (−1.14 to 1.30)	0.90	0.62 (−0.51 to 1.76)	0.28
Assessment	−0.01 (−0.63 to 0.60)	0.96	−0.15 (−0.65 to 0.35)	0.55
Interaction	0.42 (−0.48 to 1.32)	0.36	0.71 (−0.02 to 1.44)	0.057
Covariate	0.71 (0.59 to 0.83)	<0.0001	0.61 (0.50 to 0.72)	<0.0001
Area under the curve analysis				
Intervention	0.54 (−0.69 to 1.77)	0.37	1.11 (−0.14 to 2.36)	0.078
Covariate	0.72 (0.59 to 0.84)	<0.0001	0.60 (0.49 to 0.71)	<0.0001

Comparison with other studies

Previous studies of physiotherapy and occupational therapy in this setting provide conflicting results.[1-6] Results of this trial seem to support those other studies, which concluded that similar functional rehabilitation interventions had minimal impact on elderly people in residential care.[12-15]

Limitations

Certain characteristics of the population were unexpected, which could suggest that a larger sample is required. The intracluster correlation coefficients of 0.49 and 0.48 for the Barthel index and Rivermead mobility index, respectively, were higher than was anticipated from the sample size calculation conducted; however, similar intracluster correlation coefficients were observed in a study with participants from an equivalent population.[16] In addition, an analysis of a large number of studies that used intracluster correlation coefficients concluded that the magnitude of between cluster variation for a given measure can rarely be estimated in advance.[17]

In this study, care home residents were included in the trial if they scored in the mid-range on the Barthel index. Initial Barthel index screening was carried out by care home staff, and some residents did not score within the inclusion parameters when subsequent pre-intervention assessments were done. We decided that these residents should remain in the trial. The inclusion of these residents could possibly have masked intervention benefits slightly because the intervention would be of insufficient intensity to be beneficial in these individuals.

All residents with the defined level of dependency were referred for physiotherapy and/or occupational therapy, yet in routine clinical practice only those with a specific problem would be referred. This meant the therapists were in some cases delivering interventions that maintained the physical abilities of the residents rather than actively rehabilitated them. In addition, it could be argued that the setting of this study does not lend itself to improvements in independence that could be recorded with the outcome measures selected because of the standard institutional risk policies in place, such as only allowing assisted bathing.

Conclusion

The physiotherapy and occupational therapy intervention administered in this study do not have an effect on independence and mobility when applied relatively unselectively and the results do not support the provision of such services.

We acknowledge the help of care home residents and staff, the trial steering group chaired by Professor Marion Walker and including Professor Jonathan Mant, and the occupational therapy and physiotherapy advisory groups. We also thank Professors David Mant and Derick Wade for contributing to the study design. Finally, we would like to thank the locum therapists: Nicola Brittle, Eric Morgan, and Pam Versveldt.

Contributors: See bmj.com

Funding: This study was funded by The Health Foundation and the NHS National Institute for Health Research. The University of Birmingham acted as sponsors. The funders and sponsor had no involvement in the study design, data collection, analysis and interpretation, the writing of the report, or the decision to submit the paper for publication.

Competing interests: None declared.
Ethical approval: Ethical approval was granted by the Oxfordshire Applied and Qualitative Research Ethics Committee. Further site specific approval was granted by the NHS South Birmingham and NHS Birmingham East and North Local Research Ethics Committees.
Data sharing: All authors had access to all the data in the trial and can take responsibility for the integrity of the data and the accuracy of the data analysis.

1 Kerse N, Peri K, Robinson E, Wilkinson T, von Randow M, Kiata L, et al. Does a functional activity programme improve function, quality of life, and falls for residents in long term care? Cluster randomised controlled trial. *BMJ* 2008;337:a1445.
2 Mozley CG, Schneider J, Cordingley L, Molineux M, Duggan S, Hart C, et al. The care home activity project: does introducing an occupational therapy programme reduce depression in care homes? *Aging Ment Health* 2007;11:99-107.
3 Mulrow CD, Gerety MB, Kanten D, Cornell JE, DeNino LA, Chiodo L, et al. A randomized trial of physical rehabilitation for very frail nursing home residents. *JAMA* 1994;271:519-24.
4 Sackley CM, Wade DT, Mant D, Copley Atkinson J, Yudkin P, Cardoso K, et al. Cluster randomized pilot controlled trial of an occupational therapy intervention for residents with stroke in UK care homes. *Stroke* 2006;37:2336-21.
5 Rydwik E, Frandin K, Akner G. Effects of physical training on physical performance in institutionalised elderly patients (70+) with multiple diagnoses. *Age Ageing* 2004;33:13-23.
6 Helbostad JL. Fysisk trening av sykehjemsbeboere – har det noen hensikt? [Physical training for nursing home residents—has it any effect?] *Tidsskr Nor Laegeforen* 2005 4;125:1195-7 [Norwegian].
7 Collin C, Wade DT, Davies S, Horne V. The Barthel ADL Index: a reliability study. *Int Disabil Stud* 1988;10:61-3.
8 O'Neil MB, Woodard M, Sosa V, Hunter L, Mulrow CD, Gerety MB, et al. Physical therapy assessment and treatment protocol for nursing home residents. *Phys Ther* 1992;72:596-604.
9 Sackley CM, Atkinson JC, Walker MF. Occupational therapy in nursing and residential care settings: a description of a randomised controlled trial intervention. *Brit J Occupat Ther* 2004;67:104-10.
10 Collen FM, Wade DT, Robb GF, Bradshaw CM. The Rivermead Mobility Index: a further development of the Rivermead Motor Assessment. *Int Disabil Stud* 1991;13:50-4.
11 Legg L, Drummond A, Leonardi-Bee J, Gladman JRF, Corr S, Donkervoort M, et al. Occupational therapy for patients with problems in personal activities of daily living after stroke: systematic review of randomised trials. *BMJ* 2007;335:922-5.
12 Schnelle JF, Alessi CA, Simmons SF, Al-Samarrai NR, Beck JC, Ouslander JG. Translating clinical research into practice: a randomized controlled trial of exercise and incontinence care with nursing home residents. *J Am Geriatr Soc* 2002;50:1476-83.
13 Schnelle JF, Cruise PA, Rahman A, Ouslander JG. Developing rehabilitative behavioural interventions for long-term care: Technology transfer, acceptance and maintenance issues. *J Am Geriatr Soc* 1998;46:771-7.
14 Shaw C, Williams KS, Assassa RP. Patients views of a new nurse-led continence service. *J Clin Nurs* 2000;9:574-82.
15 Edwards H, Chapman H, Forster E, Gaskill D, Morrison P, Sanders F. Challenges associated with implementing an education program in a residential aged care setting. *Aust Health Rev* 2003;26:107-15.
16 Sackley CM, Rodriguez NA, Badger F, Wright CC, van den Berg MEL, Besemer J, et al. A phase II exploratory cluster randomised controlled trial of a group mobility training and staff education intervention to promote urinary continence in UK care homes. *Clin Rehabil* 2008;22:714-21.
17 Adams G, Gulliford MC, Ukoumunne OC, Eldridge S, Chinn S, Campbell MJ. Patterns of intra-cluster correlation from primary care research to inform study design and analysis. *J Clin Epidemiol* 2004;57:785-94.

Accepted: 1 March 2009

Equity, waiting times, and NHS reforms: retrospective study

Zachary N Cooper,[1] Alistair McGuire,[1] S Jones,[2] J Le Grand[3]

[1]Cowdray House, London School of Economics, London WC2A 2AE
[2]Dr Foster Intelligence, London EC1A 9LA
[3]Department of Social Policy, London School of Economics
Correspondence to: Z Cooper
z.cooper@lse.ac.uk

Cite this as: *BMJ* 2009;339:b3264
doi: 10.1136/bmj.b3264

ABSTRACT

Objective To determine whether observable changes in waiting times occurred for certain key elective procedures between 1997 and 2007 in the English National Health Service and to analyse the distribution of those changes between socioeconomic groups as an indicator of equity.
Design Retrospective study of population-wide, patient level data using ordinary least squares regression to investigate the statistical relation between waiting times and patients' socioeconomic status.
Setting English NHS from 1997 to 2007.
Participants 427 277 patients who had elective knee replacement, 406 253 who had elective hip replacement, and 2 568 318 who had elective cataract repair.
Main outcome measures Days waited from referral for surgery to surgery itself; socioeconomic status based on Carstairs index of deprivation.
Results Mean and median waiting times rose initially and then fell steadily over time. By 2007 variation in waiting times across the population tended to be lower. In 1997 waiting times and deprivation tended to be positively related. By 2007 the relation between deprivation and waiting time was less pronounced, and, in some cases, patients from the most deprived fifth were waiting less time than patients from the most advantaged fifth.
Conclusions Between 1997 and 2007 waiting times for patients having elective hip replacement, knee replacement, and cataract repair in England went down and the variation in waiting times for those procedures across socioeconomic groups was reduced. Many people feared that the government's NHS reforms would lead to inequity, but inequity with respect to waiting times did not increase; if anything, it decreased. Although proving that the later stages of those reforms, which included patient choice, provider competition, and expanded capacity, was a catalyst for improvements in equity is impossible, the data show that these reforms, at a minimum, did not harm equity.

WHAT IS ALREADY KNOWN ON THIS TOPIC

Little recent evidence exists on the association between waiting times and individual patients' socioeconomic status in England
The impact of the government's recent reforms on equity and quality is also not well documented

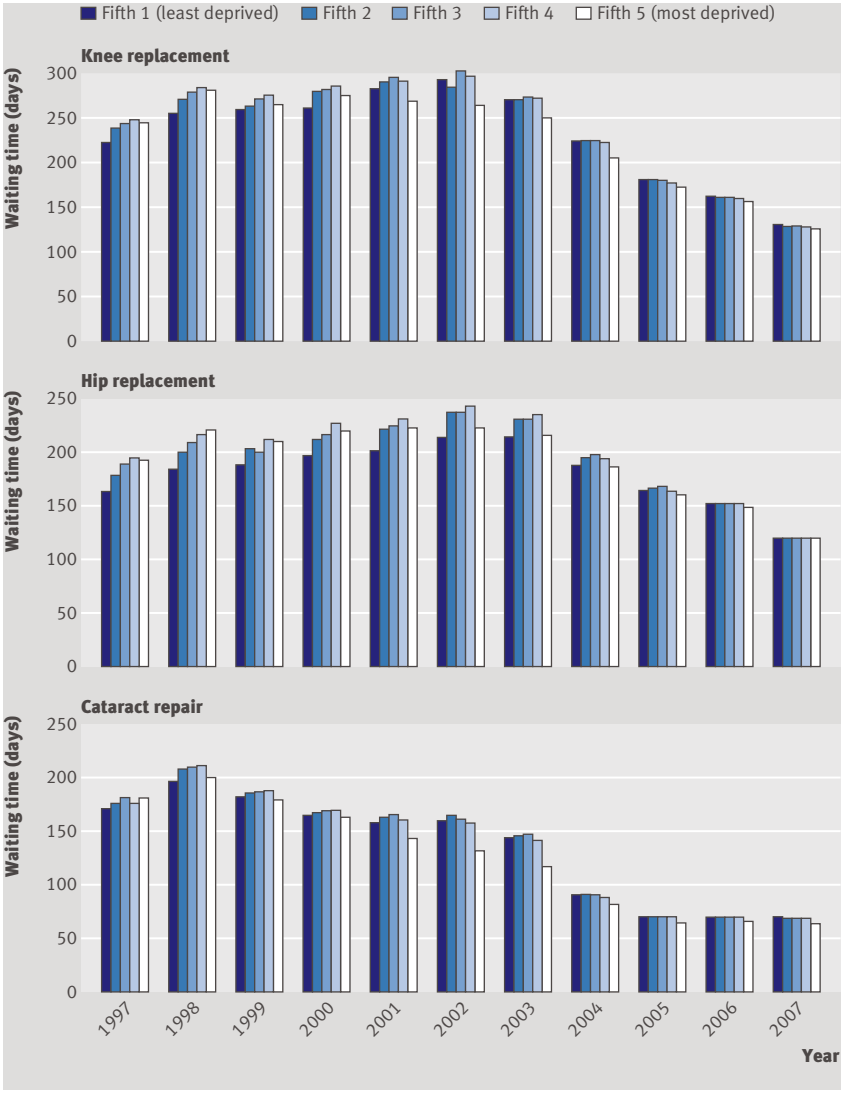
WHAT THIS PAPER ADDS

The reforms have not had a deleterious impact on the equity of waiting times for elective surgery in England

This article is an abridged version of a paper that was published on bmj.com. Cite this article as: *BMJ* 2009;339:b3264

INTRODUCTION

Hospital waiting times have dropped considerably over the past 10 years as the government in England increased the supply of doctors, increased funding for the health service, set rigid waiting time targets, and introduced market based reforms. Yet, whereas



Mean waiting times for knee replacement, hip replacement, and cataract repair, by deprivation fifth

NHS waiting times are widely accepted to have fallen between 1997 and 2007, little is known about whether the drop in waiting times has been equitably distributed with respect to socioeconomic status. One of the main fears associated with the market based elements of the reforms was that any improvements in quality or drops in waiting times would come at the expense of equity.[1-3]

We used population-wide, patient level data to examine the extent of changes in waiting times for key elective procedures between 1997 and 2007 and to analyse the distribution of those changes between socioeconomic groups. We relate these changes to three distinct periods in government policy: a period from 1997 to 2000 when the government focused on waiting lists, not waiting times, and moderately increased funding; a period from 2001 to 2004 when funding increased dramatically and the government focused on targets and performance management; and a period from 2005 to 2007 when the government expanded supply and introduced increased choice for patients and competition between providers.[4 5]

METHODS

We examined patient level, national hospital activity data for day cases and inpatient cases in England from 1 January 1997 to 31 December 2007. We examined three common, high volume elective surgical procedures that all had chronically long waiting times: knee replacement, hip replacement, and cataract repair. We looked at non-revision cases.

We excluded observations with any missing data fields or for which a patient's waiting time was greater than three years. We linked observations to local area characteristics from patients' postcodes. We calculated deprivation by using the 2001 Carstairs index of deprivation at the output area level and broke the level of deprivation into population fifths. The Carstairs index of deprivation is a composite deprivation index based on car ownership, unemployment, overcrowding, and social class within output areas.[6] In our study, deprivation 1 was the least deprived fifth and deprivation 5 was the most deprived fifth.

Waiting times were measured as the time from when the patient was referred by a specialist for surgery to the time the patient actually had surgery. We calculated mean and median waiting times for each year and each procedure. We tested differences in mean and median waiting times for the periods of 1997-2000, 2001-4, and 2005-7. We determined whether intra-year variation in waiting times existed between fifths of deprivation.

We used regression to determine whether patients' deprivation level was associated with waiting times, controlling for the patients' age, sex, area type, and the year of procedure. We ran regressions on data from three periods that corresponded to changing government policy (1997-2000, 2001-4, and 2005-7) independently.[4]

RESULTS

The total number of observations comprised 444 867 knee replacements, 423 203 hip replacements, and 2 647 235 cataract repairs. In total, 17 590 knee replacement observations, 16 950 hip replacement observations, and 78 917 cataract repair observations were excluded for missing data or because they had waiting times greater than three years. After these amendments, the observations were limited to 427 227 knee replacements, 406 253 hip replacements, and 2 568 318 cataract repairs.

For all three procedures, mean and median waiting times rose initially and then steadily fell. The figure shows mean waiting times for all three procedures, broken down by deprivation. In 1997 deprivation and waiting time tended to be positively related—the greater the degree of deprivation, the longer the waiting time. By 2007 waiting times were much more uniformly distributed across the spectrum of deprivation; for cataract repair and knee replacement, the distribution had actually reversed to show a negative relation between waiting time and deprivation.

We found a statistically significant difference in waiting times for each procedure between each of our three periods ($P<0.001$) determined by t tests and Wilcoxon sign rank tests. We found a statistically significant intra-year variation in waiting times between deprivation

groups for all three procedures for all years, except hip replacements in 2005 and 2006, measured with analysis of variance and Kruskal-Wallis rank tests ($P<0.05$).

The relation between waiting times and the deprivation fifths also changed over time. For all three procedures, each successive time period was associated with a statistically significant reduction in waiting times (see bmj.com). More interestingly, less variation in waiting times existed across socioeconomic groups over time. For example, for hip replacement surgery in 1997-2000 each successive increase in deprivation fifth was associated with a statistically significant increase in waiting time of between one and two weeks compared with the least deprived fifth ($P<0.001$). In 2001-4 variations in waiting times between deprivation fifths tended to be large and significant. Each procedure showed a modified U shaped distribution, with the middle fifths waiting the longest for care. In 2005-7 very little difference existed in days waited depending on patients' deprivation fifth. In fact, patients from the most deprived fifth having either a knee replacement or a cataract repair waited less time than patients from the least deprived fifth ($P<0.05$).

DISCUSSION

Between 1997 and 2007, waiting times for elective knee replacements, hip replacements, and cataract repairs dropped significantly and equity, measured as the variation in waiting times according to socioeconomic status, improved. Previous research has shown that greater deprivation is associated with longer waits in Scotland, and a small scale study of 4309 patients in England from April 2000 to 2001 found some inequity in the distribution of waiting times.[7 8]

Given the plethora of reforms aimed at reducing waiting times introduced between 1997 and 2007 in England, ascribing the drop in waiting times that occurred after 2000 to one policy reform is difficult. The rise in funding, the rigid government targets, and increased choice and competition are all likely to have played a role in shortening patients' waits. In addition to reducing waits and improving quality, the government argued that the reforms, especially those associated with choice for patients, would improve equity.[9-11] They argued that in health services without formalised choice, some form of privilege always exists for middle class and upper class users, who use their "voice" to negotiate for better services within the publicly funded service.[10] Creating formal choice in the health service would give all patients greater power to affect their use of resources. In contrast, several analysts argued that the expansion of choice for patients and competition between providers would not only not improve equity but would harm it.[1-3] Critics of choice and competition have argued that better off people were better equipped to choose and that the reforms would produce incentives to focus on wealthy people to the detriment of the poor.

Our results show that during the period after the reforms were introduced, waiting times for knee replacement, hip replacement, and cataract repair continued to fall and the variation in waiting times between fifths of deprivation was reduced. In certain circumstances, by 2007 patients in more deprived areas were waiting less time than patients from less deprived areas. We can assert with confidence that the introduction of choice and competition, as well as the other post-2000 government reforms, did not lead to inequitable distribution of waiting times across socioeconomic groups.

Limitations

This study has several limitations. Firstly, the Carstairs index of deprivation is one of several ways of measuring socioeconomic status. It cannot pick up deprivation of individual patients but rather the deprivation level of the area in which the patient lives. Secondly, these data are cross sectional, and the patients' characteristics varied from year to year depending on who was referred for care. Therefore, variation exists in our samples over time. Likewise, this study looks at equity with reference to use of services not with reference to access. Thirdly, this analysis is focused on three particular surgical procedures. Together, the three procedures account for between 6% and 7% of total elective surgical activity. We have no means of knowing that waiting times for all other elective surgical procedures have followed the same trends.

We thank the Dr Foster Unit at Imperial College London for processing the data. We also thank Gordon Hart for his contribution to constructing the original dataset and comments on earlier drafts of this manuscript.

Contributions: See bmj.com.

Funding: This research was funded through an LSE doctoral studentship and a Morris Finer PhD fellowship.

Competing interests: JLG worked part time in the Policy Directorate at No 10 Downing Street from October 2003 to June 2004 and full time from June 2004 until August 2005. He was seconded from his position at the London School of Economics and continued to receive the same salary and pension contributions as at the LSE. His roles at No 10 included assessing inequities in use of services within the NHS and discussing the possible impact on equity of the government's reforms on choice. He also advised on the rolling-out of the government' reform policies; the policies themselves pre-dated his appointment. The work consisted of advising the prime minister and other members of the government on health service issues, assembling research evidence as required, discussions with stakeholders, helping with the intellectual content of speeches, and working with civil servants on implementing the reforms.

Ethical approval: Not needed.

1 Appleby J, Harrison A, Devlin N. *What is the real cost of more patient choice*? London: King's Fund, 2003.
2 Oliver A, Evans J. The paradox of promoting choice in a collectivist system. *J Med Ethics* 2005;31:187.
3 Barr D, Fenton L, Blane D. The claim for patient choice and equity. *J Med Ethics* 2008;34:271-4.
4 Appleby J, Harrison A. *The war on waiting for hospital treatment*. London: King's Fund, 2005.
5 Klein R. The troubled transformation of Britain's National Health Service. *N Engl J Med* 2006;355:409-15.
6 Morgan O, Baker A. Measuring deprivation in England and Wales using 2001 Carstairs scores. *Health Stat Q* 2006;(31):28-33.
7 Pell J, Pell A, Norrie J, Ford I, Cobbe S. Effect of socioeconomic deprivation on waiting time for cardiac surgery: retrospective cohort study. *BMJ* 2000;321:15-8.
8 Siciliani L, Martin S. An empirical analysis of the impact of choice on waiting times. *Health Econ* 2007;16:763-79.
9 Department of Health. *Building on the best: choice, responsiveness and equity in the NHS*. Norwich: HMSO, 2003.
10 Dixon A, Le Grand J. Is greater patient choice consistent with equity? The case of the English NHS. *J Health Serv Res Policy* 2006;11:162-6.
11 Stevens S. Equity and choice: can the NHS offer both? In: Oliver AJ, ed. *Equity in health and health care: views from ethics, economics and political science proceedings from the meeting of the Healthcare Equity Network*. London: Nuffield Trust, 2003.

Accepted: 11 March 2009

Comparisons between geographies of mortality and deprivation from the 1900s and 2001: spatial analysis of census and mortality statistics

Ian N Gregory

Faculty of Arts and Social Sciences, Furness College, Lancaster University, Lancaster LA1 4YG
I.Gregory@lancaster.ac.uk

Cite this as: *BMJ* 2009;339:b3454
doi: 10.1136/bmj.b3454

ABSTRACT

Objectives To examine the geographical relation between mortality and deprivation in England and Wales at the start of the 20th and 21st centuries. To explore the evidence for a strengthening or weakening of this relation over the century and test for relations between the mortality and deprivation patterns of a century ago and modern mortality and causes of death.

Design Census and mortality data for 634 districts from the 1900s directly compared with interpolated ward level data from 2001.

Setting Census data and national statistics for England and Wales in the 1900s and 2001.

Population Entire population in both periods.

Main outcome measures Standardised mortality ratios for all districts for both periods with additional cause specific ratios calculated for 2001. Deprivation (Carstairs) scores for each district in 2001, with comparable measure created for the 1900s. Correlations and partial correlations between deprivation scores and standardised mortality ratios in the 1900s and 2001 for the 614 districts for which all data were available.

Results There was no evidence of a significant change in the strength of the relation between deprivation and mortality between the start and end of the 20th century. Modern patterns of mortality and deprivation remain closely related to the patterns of a century ago. Even after adjustment for modern deprivation, standardised mortality ratios for the 1900s show a significant correlation with modern mortality and most causes of death. Conversely, however, there was no significant relation between deprivation in the 1900s and modern mortality for most causes of death after adjustment for modern deprivation.

Conclusions Despite all the medical, public health, social, economic, and political changes over the 20th century, patterns of poverty and mortality and the relations between them remain firmly entrenched. There is a strong relation between the mortality levels of a century ago and those of today. This goes beyond what would have been expected from the continuing relation between deprivation and mortality and holds true for most major modern causes of death.

WHAT IS ALREADY KNOWN ON THIS TOPIC

There has been a strong relation between deprivation and poverty in recent decades

There is a clear relation between poverty and mortality in inner London in the 1890s and poverty and mortality in the 1990s

There is some evidence of a direct link between the social conditions in which an individual's mother or grandparents lived and their adult heath

WHAT THIS STUDY ADDS

Despite the fact that inequalities in mortality have narrowed, the relation between poverty and mortality across the whole of England and Wales seems as strong today as it was at the start of the 20th century

Mortality and deprivation patterns of 100 years ago are strong predictors of these patterns today; in particular areas with high rates of mortality or deprivation in the past still tend to have high rates of mortality today

Even when the effects of modern deprivation are taken into account, mortality patterns from the 1900s still have a significant relation with mortality today and this affects most major modern causes of death

INTRODUCTION

The 20th century saw dramatic improvements in patterns of mortality in England and Wales. Age and sex specific mortality rates declined across all ages but particularly among the young. Life expectancy has risen from 46 for men and 50 for women in the 1900s[1] to 77 for men and 81 for women in 2001.[2] These changes are linked to a major change in causes of death. Modern causes are dominated by cancers, which contributed 25.6% of all deaths in 2001, ischaemic heart diseases (19.9%), and stroke (11.0%). In the 1900s classification was less well organised, resulting in over half of all deaths being assigned to "other causes." Respiratory diseases and cancer were important. The major change is that in the 1900s infectious and parasitic diseases accounted for nearly a fifth of deaths.

The experience of poverty has also changed over the century. In the 1900s poverty was usually an absolute concept, which meant that the income of an individual was not "sufficient to obtain the minimum necessaries for the maintenance of mere physical efficiency."[3] Advances in standards of living and the growth of the welfare state now mean poverty tends to be seen as a relative concept. Relative poverty is usually expressed by comparing the individual's income or deprivation with that experienced by the society as a whole.

There is a well known relation between poverty and mortality today,[4 5] but few studies have looked at whether this has changed over the long term. Dorling et al compared mortality and poverty in inner London in the 1890s and the 1990s and showed clear links between an area's modern mortality and its poverty in the past.[6] I explored the link between deprivation and mortality for areas in the 1900s and today for the whole of England and Wales.

METHODS

The Office for National Statistics published modern mortality and deprivation data for England and Wales

This article is an abridged version of a paper that was published on bmj.com. Cite this article as: *BMJ* 2009;339:b3454

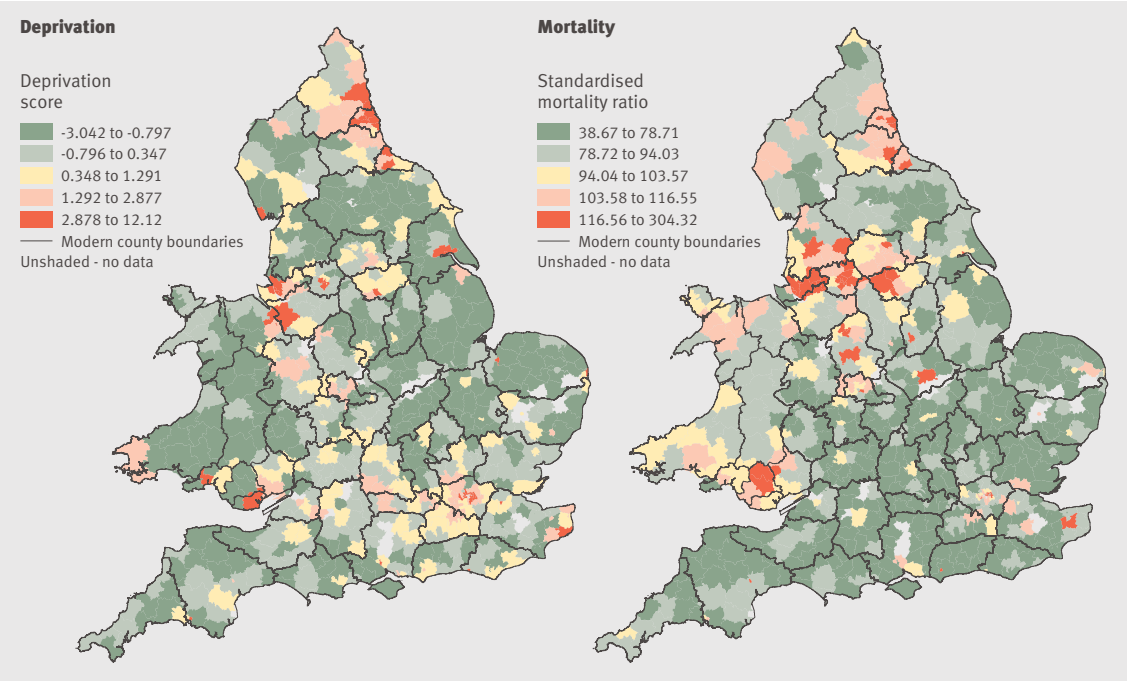


Fig 1 | Deprivation and mortality in the 1900s. Class intervals determined from population fifths from 1901. Modern county boundaries superimposed on 1900s registration districts to assist orientation

in 2001. Similar data are available for the decade from 1901 to 1910.[7-9] I developed a deprivation index for the 1900s and interpolated 2001 ward-level data onto the registration districts of the 1900s data to allow direct comparisons.

The Carstairs index of deprivation is an aggregate of standardised data from four census variables: overcrowded housing, low social class, male unemployment, and households without a car.[10] I calculated a deprivation index using variables similar to those used in the Carstairs index for the whole of England and Wales in the 1900s for 614 of the 634 registration districts.

I used a geographical information system that stores the statistical data along with the boundaries of the administrative units they refer to. Having the data in this form allowed calculation of the intersection between the 2001 census area statistics wards and the registration districts and poor law unions of the 1900s. In this way I constructed directly comparable measures of area based mortality and deprivation from the beginning of the 20th and 21st centuries.

RESULTS

In the 1900s the highest rates of deprivation and mortality were found in urban and industrial areas such as inner London, south Wales, Birmingham, Liverpool, Manchester, Sheffield, and the north east. Low rates are a primarily rural phenomenon (fig 1). Calculation of a Pearson product moment correlation coefficient between the deprivation index and the standardised mortality ratios for the 614 registration districts for which all data were available gave a coefficient of $r=0.503$ ($P<0.001$). This tells us that there was a strong positive relation between mortality and deprivation across England and Wales in the 1900s.

Figure 2 shows Carstairs scores and standardised mortality ratios for 2001 using 1900s registration districts. The pattern of deprivation seems similar to the 1900s pattern: high rates are concentrated in urban and industrial areas, except that these have spread out, particularly in what is now the M62 motorway corridor running from Liverpool through Manchester and Sheffield to Hull. The modern mortality map seems superficially to be noticeably different from the 1900s map. On closer inspection, however, urban and industrial areas are high at both dates. A difference is that there are several rural areas that have high standardised mortality ratios in 2001. The Pearson's product moment correlation coefficient between mortality and deprivation with these units is $r=0.497$ ($P<0.001$). Given the changes in the deprivation measures, the similarity in the correlations, $r=0.503$ and $r=0.497$ in the 1900s and 2001, respectively, suggest that there has been little change in the strength of this relation over the course of the century.

In 2001 the average standardised mortality ratio of the population 10th with the lowest mortality rates was 75.6 while the average for the highest 10th was 135.5. Thus in 2001 the population 10% of the population living in the highest mortality areas had mortality rates that were 1.79 times higher than the 10% living in areas with the lowest rates. In the 1900s the average for the highest 10th was 2.05 times higher than the lowest 10th. This shows that the morality gap between the best and worst areas had narrowed over the century. The worst deprivation 10th in 2001 had an average standardised mortality ratio 1.36 times higher than the best, in the 1900s this ratio was 1.39. This shows that, while the mortality gap might have narrowed over the century, the relation between the extremes of deprivation and mortality is as strong today as it was a century ago.

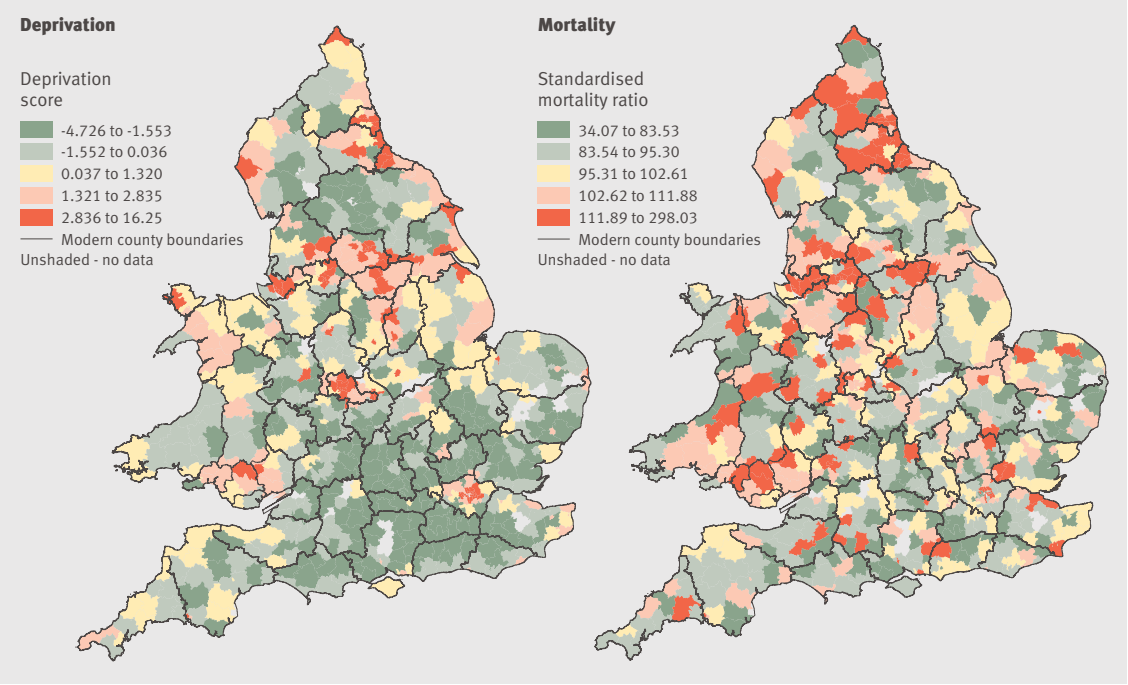


Fig 2 | **Deprivation and mortality in 2001. Data interpolated from census area statistics wards onto registration districts for the 1900s. Class intervals determined from population fifths from 2001. Modern county boundaries superimposed on 1900s registration districts to assist orientation**

The next question is whether the spatial pattern of mortality and deprivation in the 1900s can be used to predict modern patterns. Comparing standardised mortality ratios in the 1900s with those in 2001 gives a correlation coefficient of r=0.414 (P<0.001), while comparing the deprivation scores gives a correlation coefficient of r=0.578 (P<0.001). This confirms the patterns suggested by the figures—namely, that strong relations between patterns of mortality and deprivation have persisted over the course of the century.

The areas with the lowest and highest 10ths of mortality in the 1900s still have low and high mortality rates in 2001. In the 1900s the high mortality 10th had rates 2.05 times higher than the low mortality 10th. The mortality ratio between these areas had narrowed to being a more modest 1.38 by 2001, but this is still a large difference. The impact of 1900s deprivation is also interesting because areas with the lowest deprivation scores in the 1900s have an average standardised mortality ratio of 97.2 in 2001, suggesting that areas that were affluent 100 years ago do not now have mortality rates that are significantly better than other areas. By contrast, areas with the highest deprivation scores in the 1900s still have high standardised mortality ratios today, although at 116.6, this is less pronounced than the impact of high 1900s mortality at 125.4.

I used partial correlation coefficients to explore whether modern standardised mortality ratios are related to either mortality or deprivation in the 1900s, controlling for the impact of modern deprivation. This gave partial correlation coefficients of r=0.160 (P<0.001) for 1900s mortality, but only −0.016 for 1900s deprivation (P=0.690) (table). The table summarises the correlation coefficients. These suggest that mortality and deprivation continue to be closely linked and, in addition, that the mortality patterns of a century ago are still related to today's mortality patterns in a way that cannot simply be explained by inertia in deprivation patterns. Deprivation in the 1900s does not seem to be related to modern mortality once modern deprivation is controlled for (table).

DISCUSSION

The 20th century has seen a dramatic decline in mortality but the link between mortality and deprivation across England and Wales remains as strong today as it was a century ago. Geographical inequalities in mortality have declined slightly but there is no evidence that inequalities in deprivation have declined or that the relation between mortality and deprivation has lessened to any significant degree. Patterns of mortality and deprivation are deeply entrenched such that in both cases the patterns

Pearson's product moment correlation coefficients between variables. All data have been standardised on 1900s registration districts unless otherwise stated

Variable 1	Variable 2	Controlled for	r
Deprivation 1900s	Mortality 1900s	—	0.503*
Deprivation 2001	Mortality 2001	—	0.497*
Deprivation 2001 (ward level)	Mortality 2001 (ward level)	—	0.466*
Mortality 1900s	Mortality 2001	—	0.414*
Deprivation 1900s	Deprivation 2001	—	0.578*
Mortality 1900s	Deprivation 2001	—	0.612*
Deprivation 1900s	Mortality 2001	—	0.276*
Mortality 1900s	Mortality 2001	Deprivation 2001	0.160*
Deprivation 1900s	Mortality 2001	Deprivation 2001	−0.016
Deprivation 1900s	Deprivation 2001	Mortality 2001	0.528*
Mortality 1900s	Deprivation 2001	Mortality 2001	0.514*

*Significant at P<0.01.

of a century ago are strong predictors of today's patterns. This is not simply because of inertia in socioeconomic conditions because mortality in the 1900s is significantly related to modern mortality even when modern deprivation is taken into account.

I focused on areas rather than individuals, which is both a strength and a weakness. It is a strength because the study covers the entire population; it is a limitation because it is unable to explain the patterns and, in particular, say whether they are caused by area effects or individual level behaviours.

The 20th century has seen widescale reforms aimed at improving living conditions for society in general and the poor in particular. These include the formation of the National Health Service (NHS) and the welfare state. These have undoubtedly led to large increases in life expectancy but seem to have failed to reduce the impact that poverty has on mortality. This is not to say that these policies have been a failure as it is possible that without them health inequalities might have become far worse over time.

A major unanswered question is whether the increases in mortality in deprived areas that can be ascribed to conditions in the past are caused by area effects or operate more directly at the individual level. If it is an area effect then the long term effects of the physical, economic, or social environment still seem to have a relation to modern conditions beyond those that the Carstairs index is able to measure. Individual effects are more problematic. There have been suggestions that socioeconomic conditions can have hereditary effects—for example, it has been claimed that there is a relation between mortality among Swedish men and the food supply of their paternal grandfathers.[11] On a shorter time scale, the Barker hypothesis claims that poor fetal nutrition leads to heart disease, diabetes, and respiratory disorders in later life.[12 13]

Digital boundary data for 2001 census area statistics wards were provided through EDINA (http://edina.ac.uk/) UKBorders with the support of the ESRC (Economic and Social Research Council) and JISC (Joint Information Services Committee) and use boundary material that is copyright of the Crown. The 2001 VS data were created by the Office for National Statistics and distributed by the UK Data Archive, University of Essex. Crown copyright material is reproduced with the permission of the controller of HMSO. The original data creators, depositors, or copyright holders, the funders of the data collections (if different), and the UK Data Archive bear no responsibility for their further analysis or interpretation. Census output is Crown Copyright and is reproduced with the permission of the controller of HMSO. The ArcGIS software package produced by ESRI (www.esri.com) was used to standardise the geographical units used in this study. It also assisted the analysis and produced the maps.

Funding: This study was funded by the Leverhulme Trust under their early career fellowship scheme (ECF/40115).

Competing interests: None declared.

Ethical approval: Not required.

1 Kinsella K. Changes in life expectancy, 1900-1990. *Am J Clin Nutr* 1992;55:1196-202.
2 Rasulo D, Bajekal M, Yar M. Inequalities in health expectancies in England and Wales—small area analysis from the 2001 Census. *Health Stat Q* 2007;34:35-45.
3 Rowntree BS. *Poverty. A study of town life*. London: MacMillan, 1901:86.
4 Shelton N, Birkin M, Dorling D. Where not to live: a geo-demographic classification of mortality for England and Wales, 1981-2000. *Health Place* 2006;12:557-69.
5 Shaw M, Thomas B, Davey Smith G, Dorling D. *The grim reaper's road map*. Bristol: Policy Press, 2008.
6 Dorling D, Mitchell R, Shaw M, Orford S, Davey Smith G. The ghost of Christmas past: health effects of poverty in London in 1896 and 1991. *BMJ* 2000;321:1547-51.
7 Registrar General's decennial supplement to the seventy-fifth annual report. Part III: Registration summary tables, 1901-1910. British Parliamentary Papers 1914-16/VIII (Cd 8002) .
8 County Reports from the 1901 Census, British Parliamentary Papers: 1902 CXVIII to CXXI and 1903 LXXXIV to LXXXVI.
9 1911 Census Report. *Volume VII. Ages and condition as to marriage*. London: National Archives (British Parliamentary Papers 1912-13 CXIII.)
10 Morgan O, Baker A. Measuring deprivation in England and Wales using 2001 Carstairs scores. *Health Stat Q* 2006;31:28-33.
11 Pembrey ME, Bygren LO, Kaati G, Edvinsson S, Northstone K, Sjostrom M, et al. Sex-specific, male-line transgenerational responses in humans. *Eur J Hum Genet* 2006;14:159-66.
12 Barker DJB, ed. *Fetal and infant origins of adult disease*. London: BMJ Publishing, 1992.
13 Barker DJB. *Mothers, babies, and diseases in later life*. London: BMJ Publishing, 1994.

Accepted: 24 March 2009

The effects of excluding patients from the analysis in randomised controlled trials: meta-epidemiological study

Eveline Nüesch,[12] Sven Trelle,[12] Stephan Reichenbach,[13] Anne W S Rutjes,[14] Elizabeth Bürgi,[5] Martin Scherer,[67] Douglas G Altman,[8] Peter Jüni[12]

ABSTRACT

Objective To examine whether excluding patients from the analysis of randomised trials are associated with biased estimates of treatment effects and higher heterogeneity between trials.

Design Meta-epidemiological study based on a collection of meta-analyses of randomised trials.

Data sources 14 meta-analyses including 167 trials that compared therapeutic interventions with placebo or non-intervention control in patients with osteoarthritis of the hip or knee and used patient reported pain as an outcome.

Methods Effect sizes were calculated from differences in means of pain intensity between groups at the end of follow-up, divided by the pooled standard deviation. Trials were combined by using random effects meta-analysis. Estimates of treatment effects were compared between trials with and trials without exclusions from the analysis, and the impact of restricting meta-analyses to trials without exclusions was assessed.

Results 39 trials (23%) had included all patients in the analysis. In 128 trials (77%) some patients were excluded from the analysis. Effect sizes from trials with exclusions tended to be more beneficial than those from trials without exclusions (difference –0.13, 95% confidence interval –0.29 to 0.04). However, estimates of bias between individual meta-analyses varied considerably (τ^2=0.07). Tests of interaction between exclusions from the analysis and estimates of treatment effects were positive in five meta-analyses. Stratified analyses indicated that differences in effect sizes between trials with and trials without exclusions

This article is an abridged version of a paper that was published on bmj.com. Cite this article as: *BMJ* 2009;339:b3244

[1]Institute of Social and Preventive Medicine, University of Bern, Switzerland
[2]CTU Bern, Bern University Hospital, Switzerland
[3]Department of Rheumatology, Immunology and Allergology, Bern University Hospital, Switzerland
[4]Department of Clinical Pharmacology and Epidemiology, Consorzio Mario Negri Sud, Santa Maria Imbaro, Chieti, Italy
[5]Department of Internal Medicine, Bern University Hospital, Switzerland
[6]Department of General Practice, University of Göttingen, Germany
[7]Institute of Social Medicine, University of Luebeck, Germany
[8]Centre for Statistics in Medicine, University of Oxford, Oxford
Correspondence to: P Jüni
juni@ispm.unibe.ch

Cite this as: *BMJ* 2009;339:b3244
doi: 10.1136/bmj.b3244

were more pronounced in meta-analyses with high between trial heterogeneity, in meta-analyses with large estimated treatment benefits, and in meta-analyses of complementary medicine. Restriction of meta-analyses to trials without exclusions resulted in smaller estimated treatment benefits, larger P values, and considerable decreases in between trial heterogeneity.

Conclusion Excluding patients from the analysis in randomised trials often results in biased estimates of treatment effects, but the extent and direction of bias is unpredictable. Results from intention to treat analyses should always be described in reports of randomised trials. In systematic reviews, the influence of exclusions from the analysis on estimated treatment effects should routinely be assessed.

INTRODUCTION

In clinical trials, deviations from protocol and losses to follow-up often lead to the exclusion of patients from the analysis.[1 2] Patients excluded after randomisation are unlikely to be representative of those remaining in the trial. The selective occurrence and biased handling of protocol deviations and losses to follow-up may lead to results that differ from the true values, a systematic error called attrition bias.[2] To ensure that intervention and control groups are comparable and to prevent attrition bias the analysis should be done according to the intention to treat principle—that is, all randomised patients should be included in the analysis and kept in their original groups, regardless of their adherence to the protocol.[2 3]

Four meta-epidemiological studies explored the association of withdrawals, dropouts, and exclusions after randomisation with estimated treatment effects.[1 4 5-7] The direction and magnitude of attrition bias varied between the studies according to methods and definitions used and clinical areas.[1 6 8 9] In general, randomised controlled trials using subjective outcomes are more susceptible to bias than trials using objective outcomes such as overall mortality. In trials of osteoarthritis, treatment effects are often evaluated using subjective outcomes, such as intensity of pain. Meta-analyses of osteoarthritis trials may therefore be particularly prone to attrition bias.[2]

We examined whether excluding patients from the analysis were associated with biased estimates of treatment effects and increased heterogeneity between trials in meta-analyses of interventions used for the treatment of pain in osteoarthritis.

WHAT IS ALREADY KNOWN ON THIS TOPIC

Excluding randomised patients from the analysis in randomised controlled trials may result in attrition bias

The direction and magnitude of attrition bias varies between different studies according to different methods and definitions used and different clinical areas addressed

WHAT THIS STUDY ADDS

Excluding patients from the analysis of randomised trials often resulted in biased estimates of treatment effects, but the extent and direction of bias remained unpredictable in a specific situation

Overestimation of treatment benefits seemed particularly pronounced in meta-analyses with high between trial heterogeneity, in meta-analyses with large estimated treatment benefits, and in meta-analyses of complementary medicine

Restriction of meta-analyses to trials without exclusions of patients resulted in smaller estimated treatment benefits, larger P values, and considerable decreases in between trial heterogeneity

METHODS

We carried out an electronic search for meta-analyses of randomised trials (those using an unpredictable allocation sequence) or quasi-randomised trials (those using potentially predictable allocation mechanisms) in patients with osteoarthritis of the knee or hip (see bmj.com). Meta-analyses were eligible if they assessed patient reported pain comparing any intervention with placebo, sham, or a non-intervention control.

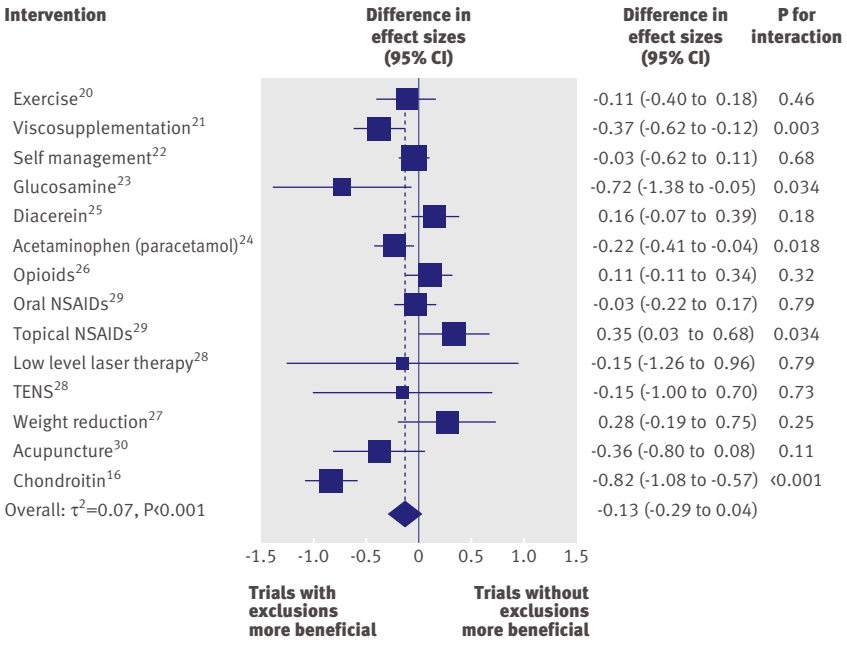
Component trials were classified to have had no exclusions of patients from the analysis if there was an explicit statement that all randomised patients were included in the analysis of the outcome we extracted or if the reported numbers of patients randomised and analysed on this outcome were identical. We classified trials to have had exclusions if they explicitly reported this, if the number of patients analysed was lower than the number randomised, or if it was unclear whether exclusions had occurred. Concealment of allocation was considered adequate if those responsible for patient inclusion were unable to suspect the next treatment allocation. Blinding of patients was considered adequate if experimental and control interventions were described as indistinguishable or if a double dummy technique was used.[2]

The primary outcome was patient reported pain. If different pain outcomes were reported we extracted one outcome per study according to a hierarchy.[10 11]

Statistical analysis

We expressed treatment effects as effect sizes (difference in mean values at end of follow-up divided by pooled standard deviation). Negative effect sizes indicate a beneficial effect of the experimental intervention. We used approximations for unavailable data.[11] To combine effect sizes across trials we used standard inverse variance random effects meta-analyses and calculated the variance estimate τ^2 as a measure of heterogeneity.[12]

We used random effects meta-analysis to estimate effect sizes separately for trials with and trials without exclusions of patients from the analysis. Then we derived differences between estimates from trials with exclusions and trials without exclusions for each meta-analysis and combined these differences using random effects meta-analysis.[13] A negative difference in effect sizes indicates that trials with exclusions show a more beneficial treatment effect. Formal tests of interaction between exclusions from the analysis and estimated treatment benefits were done separately for each meta-analysis (see bmj.com) using z



Difference in effect sizes between 128 trials with and 39 trials without exclusions of patients from analysis. A negative difference in effect sizes indicates that trials with exclusions of patients from analysis show more beneficial treatment effects. P values are for interaction between exclusions from analysis and effect sizes. NSAIDs=non-steroidal anti-inflammatory drugs; TENS=transcutaneous electrical nerve stimulation

scores. We carried out stratified analyses accompanied by interaction tests for the following characteristics: between trial heterogeneity in the overall meta-analysis (low, $\tau^2<0.06$, *v* high, $\tau^2\geq0.06$), treatment benefit in the overall meta-analysis (small, effect sizes >-0.5, *v* large, effect sizes ≤-0.5),[10 14] and type of intervention assessed in the meta-analysis (drug *v* other interventions, conventional *v* complementary medicine). To control confounding by concealment of allocation and by patient blinding, we used stratification by these factors to derive differences between trials with and trials without exclusions adjusted for concealment of allocation and adjusted for patient blinding.

Finally, we compared pooled effect sizes, between trial heterogeneity, precision (inverse of the standard error), and P values for pooled effect sizes between overall random effects meta-analyses including all trials and restricted meta-analyses including trials without exclusions only. Measures were compared using scatter plots and Wilcoxon rank tests for paired observations. P values were two sided.

RESULTS

Overall, 14 meta-analyses (167 trials in 41 170 patients) included at least one trial with and one without exclusion of patients from the analysis and contributed to the study.[11 15-25] Eight assessed drug interventions and five assessed complementary medicine. The number of trials per meta-analysis ranged from three to 24 (median 11) and the number of patients per meta-analysis from 278 to 13 659 (median 1731). The pooled effect sizes derived from random effects meta-analyses including all trials ranged from −0.07, indicating essentially no benefit, to −0.88, representing a large benefit. All meta-analyses favoured the experimental intervention and 11 showed significant differences between experimental and control intervention at P=0.05. The variance τ^2 varied between 0.00 and 0.52 (median 0.04, see bmj.com).

Characteristics of component trials

Thirty nine of the 167 trials (23%) included all randomised patients in the analysis (see bmj.com). Exclusions occurred in 114 trials (69%) and in 14 trials (8%) it was unclear whether exclusions had occurred. Exclusions ranged from 0.1% to 40% (median 7.2%). Trials with exclusions were less likely to provide information on losses to follow-up (P=0.002). Data imputations using the last observation carried forward were reported by 27% of trials with exclusions and 49% of trials without exclusions and multiple imputation by 4% and 15%. For 68% and 15% it was unclear how the trialists dealt with missing data. Trials with exclusions were published earlier than those without exclusions (mean 1998 (SD 6) *v* 2001 (SD 4); P=0.002) and tended to report adequate concealment of allocation and sample size calculations less often.

Effect of exclusions on estimates of treatment effects

On average, treatment effects were more beneficial in trials with than in trials without exclusions (difference in effect sizes −0.13, 95% confidence interval −0.29 to 0.04, P=0.13), but the variability in bias between meta-analyses was considerable ($\tau^2=0.07$, P<0.001; figure). Differences in effect sizes ranged from −0.82 to 0.35. Tests of interaction between exclusions from the analysis and estimates of treatment effects were positive at P=0.05 in five meta-analyses: in four meta-analyses estimated effects were more beneficial in trials with exclusions from the analysis and in one meta-analysis estimated effects were more beneficial in trials without exclusions (figure).

Differences between trials with and trials without exclusions were evident in meta-analyses with a high degree of between trial heterogeneity, but not in meta-analyses with low between trial heterogeneity (P for interaction <0.001; see bmj.com). Similarly, differences were more pronounced in meta-analyses with large estimated treatment benefits in the overall meta-analysis than in meta-analyses with small estimated benefits and in meta-analyses of complementary interventions compared with conventional medicine (P for interaction <0.001 for both). When stratifying for these characteristics, the variability in bias decreased considerably. When adjusting for concealment of allocation (−0.11, 95% confidence interval −0.28 to 0.05, P=0.18) or patient blinding (−0.15, −0.30 to 0.00, P=0.047), average differences between trials with and trials without exclusions were robust. In both adjusted analyses the variability in bias between meta-analyses was much the same as in the primary analysis, with variance estimates τ^2 of 0.08 and 0.06 (both P<0.001), respectively.

Impact of restricting meta-analyses to trials without exclusions

When restricting meta-analyses to trials without exclusions only, the median number of trials included in a single meta-analysis decreased from 11 to 2 and the median number of patients from 1731 to 401. Estimates of treatment benefits decreased in 10 meta-analyses and increased in four (P=0.10). Between trial heterogeneity decreased in 12 meta-analyses and increased in one (P=0.006). For one meta-analysis only one trial had no exclusions from the analysis, and no between trial heterogeneity could be estimated after the restriction.[25] Precisions of pooled effect size estimates decreased in nine meta-analyses and increased in five (P=0.25). P values became larger in 10 meta-analyses and smaller in four (P=0.016). After the restriction to trials without exclusions only, six meta-analyses lost significance at the P=0.05 level.

DISCUSSION

In this meta-epidemiological study of 14 meta-analyses and 167 trials we found that excluding randomised patients from the analysis often resulted in biased estimates of treatment effects. The average estimate of bias of a difference in effect size of −0.13 may seem small (figure), but it corresponds to one quarter to one half of a typical treatment effect found for interventions in osteoarthritis.[10] The impact of exclusions on estimates of treatment effects seemed most pronounced in meta-analyses with large treatment benefits, complementary interventions, and a high degree of heterogeneity between trials, but the extent and direction of bias may be unpredictable in a specific situation. Tests of interaction between exclusions from the analysis and estimates of treatment effects were statistically significant in five meta-analyses; in four of these meta-analyses, estimated treatment effects were less beneficial in trials without exclusions.

When restricting meta-analyses to trials without exclusions P values increased in most cases and six meta-analyses lost significance at P=0.05 (see bmj.com). This increase in P values was not only due to a loss of statistical power.[26] As a result of the restriction the between trial heterogeneity τ^2 decreased considerably. Therefore the average loss of statistical precision of random effects meta-analyses was smaller than what could be expected after the exclusion of over half the trials. Only in five meta-analyses was there a relevant loss of precision after the restriction, in six meta-analyses the statistical precision remained much the same, and in three the precision increased.

Strengths and limitations of the study

We did not rely on statements in the reports on whether an intention to treat analysis was done or not. Rather we required explicit information on the flow of patients [27 28] or statements that all randomised patients were included in the analysis. Some might argue that our distinction between trials with and trials without exclusions from the analysis was overly stringent. We would expect that any bias associated with exclusions will increase with the number of exclusions. Others may argue that our classification was not stringent enough and that we should have required an affirmative statement that no crossovers had occurred and that all randomised patients were included in the analysis in the group to which they were originally allocated. Only seven of the 167 trials (4%) provided this information.

Our study is based on published information and depends on the quality of the reports. We were able to determine from trial reports in all but 14 trials whether exclusions had occurred, however, and bias introduced by misclassification of trials due to inadequate reporting is likely to be minimal. At least two thirds of the trials included in our study had incomplete outcome data. We were unable to examine whether the approach used for data imputation influences estimates of treatment effects because of the low quality of reporting.[28-30] In our study the observed association between exclusions of patients from the analysis and estimates of treatment effects could be confounded by concealment of allocation, which may have resulted in spurious associations. When accounting for concealment of allocation in a sensitivity analysis, however, we found our results to be robust.

Implications

The most stringent interpretation of intention to treat includes the analysis of all patients, regardless of whether they were eligible, received treatment, and adhered to the protocol.[3] Many trialists exclude randomised patients who did not receive at least one dose of the allocated intervention, whereas others exclude patients found retrospectively to be ineligible.[3 31] Both approaches may produce unbiased estimates if patients and treating doctors are unaware of the allocated intervention and if the decision to exclude patients is based solely on information collected before randomisation and unrelated to group assignment and clinical outcome.[31] In addition, exclusions from the analysis owing to randomly missing outcome data may be less problematic than the selective exclusion of patients owing to protocol violations. These assumptions are hardly ever verifiable: details on the flow of participants through a trial and descriptions of procedures used to determine whether patients should be excluded from the analysis are often omitted from published reports of randomised trials.[3 28] Therefore it is difficult to determine from published information whether reported exclusions from the analyses resulted in bias,[2] and strict adherence to the intention to treat principle should be advocated.[27 32]

Per protocol analyses include only those patients who received treatment as defined in the protocol and provided outcome data. However, patients excluded from per protocol analyses are likely to be different from those analysed.[33] Trials without exclusions more often reported imputations of missing data than those with exclusions. The last observation carried forward approach was used most often. This method is popular in musculoskeletal research[34 35] but leads

to overly precise estimates and potential bias.[36 37] Multiple imputation is more difficult to carry out but avoids those problems.[38] The CONSORT statement urges transparent reporting of the flow of participants through a trial, including a description of withdrawals and losses to follow-up and reasons for exclusions.[27 32] In our view a description of strategies used to handle missing outcome data is also essential.

Conclusions

To avoid potential attrition bias, trialists should ensure low dropout rates and high compliance rates and minimise missing outcome data. Results of intention to treat analyses should always be reported. Sensitivity analyses, restricted to patients adhering to the protocol, may be described in addition. In systematic reviews and meta-analyses, data extraction should be based on results from analyses of all randomised patients, whenever possible. The influence of exclusions from randomisation on estimated treatment benefits should be routinely assessed in stratified analyses.

We thank Sacha Blank, Liz King, Katharina Liewald, Linda Nartey, Rebekka Sterchi, and Beatrice Tschannen for help with data extraction; the authors who provided study data; Malcolm Sturdy for the development and maintenance of the database; and Fitore Sallahaj for data entry.

Contributors: See bmj.com.

Funding: Swiss National Science Foundation (grant Nos 4053-40-104762/3 and 3200-066378) to PJ and SR, and the Swiss Society of Internal Medicine to PJ. The study was part of the Swiss National Science Foundation's National Research Programme 53 on musculoskeletal health. SR was a recipient of a research fellowship funded by the Swiss National Science Foundation (grant No PBBEB-115067). MS was supported by a young investigators' award of the German Ministry of Education and Research (grant No 01 GK 0516). PJ was a PROSPER (programme for social medicine, preventive and epidemiological research) fellow funded by the Swiss National Science Foundation (grant No 3233-066377). The funders had no role in the study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all data of the study and had final responsibility for the decision to submit for publication. None of the authors is affiliated with or funded by any manufacturer of any intervention used for osteoarthritis.

Competing interests: None declared.

Ethical approval: Not required.

1 Tierney JF, Stewart LA. Investigating patient exclusion bias in meta-analysis. *Int J Epidemiol* 2005;34:79-87.
2 Juni P, Altman DG, Egger M. Systematic reviews in health care: assessing the quality of controlled clinical trials. *BMJ* 2001;323:42-6.
3 Hollis S, Campbell F. What is meant by intention to treat analysis? Survey of published randomised controlled trials. *BMJ* 1999;319:670-4.
4 Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;273:408-12.
5 Kjaergard LL, Villumsen J, Gluud C. Reported methodologic quality and discrepancies between large and small randomized trials in meta-analyses. *Ann Intern Med* 2001;135:982-9.
6 Balk EM, Bonis PA, Moskowitz H, Schmid CH, Ioannidis JP, Wang C, et al. Correlation of quality measures with estimates of treatment effect in meta-analyses of randomized controlled trials. *JAMA* 2002;287:2973-82.
7 Gluud LL, Thorlund K, Gluud C, Woods L, Harris R, Sterne JA. Correction: reported methodologic quality and discrepancies between large and small randomized trials in meta-analyses. *Ann Intern Med* 2008;149:219.
8 Juni P, Egger M. Commentary: empirical evidence of attrition bias in clinical trials. *Int J Epidemiol* 2005;34:87-8.
9 Wood L, Egger M, Gluud LL, Schulz KF, Juni P, Altman DG, et al. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. *BMJ* 2008;336:601-5.
10 Juni P, Reichenbach S, Dieppe P. Osteoarthritis: rational approach to treating the individual. *Best Pract Res Clin Rheumatol* 2006;20:721-40.
11 Reichenbach S, Sterchi R, Scherer M, Trelle S, Burgi E, Burgi U, et al. Meta-analysis: chondroitin for osteoarthritis of the knee or hip. *Ann Intern Med* 2007;146:580-90.
12 DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986;7:177-88.
13 Sterne JA, Juni P, Schulz KF, Altman DG, Bartlett C, Egger M. Statistical methods for assessing the influence of study characteristics on treatment effects in 'meta-epidemiological' research. *Stat Med* 2002;21:1513-24.
14 Cohen J. *Statistical power analysis for the behavioral sciences*. 2nd ed. Hillsdale, NJ: Lawrence Earlbaum, 1988.
15 Fransen M, McConnell S, Bell M. Exercise for osteoarthritis of the hip or knee. *Cochrane Database Syst Rev* 2003(3):CD004286.
16 Lo GH, LaValley M, McAlindon T, Felson DT. Intra-articular hyaluronic acid in treatment of knee osteoarthritis: a meta-analysis. *JAMA* 2003;290:3115-21.
17 Chodosh J, Morton SC, Mojica W, Maglione M, Suttorp MJ, Hilton L, et al. Meta-analysis: chronic disease self-management programs for older adults. *Ann Intern Med* 2005;143:427-38.
18 Towheed TE, Maxwell L, Anastassiades TP, Shea B, Houpt J, Robinson V, et al. Glucosamine therapy for treating osteoarthritis. *Cochrane Database Syst Rev* 2005(2):CD002946.
19 Towheed TE, Maxwell L, Judd MG, Catton M, Hochberg MC, Wells G. Acetaminophen for osteoarthritis. *Cochrane Database Syst Rev* 2006(1):CD004257.
20 Rintelen B, Neumann K, Leeb BF. A meta-analysis of controlled clinical studies with diacerein in the treatment of osteoarthritis. *Arch Intern Med* 2006;166:1899-906.
21 Avouac J, Gossec L, Dougados M. Efficacy and safety of opioids for osteoarthritis: a meta-analysis of randomized controlled trials. *Osteoarthritis Cartilage* 2007;15:957-65.
22 Christensen R, Bartels EM, Astrup A, Bliddal H. Effect of weight reduction in obese patients diagnosed with knee osteoarthritis: a systematic review and meta-analysis. *Ann Rheum Dis* 2007;66:433-9.
23 Bjordal JM, Johnson MI, Lopes-Martins RA, Bogen B, Chow R, Ljunggren AE. Short-term efficacy of physical interventions in osteoarthritic knee pain. A systematic review and meta-analysis of randomised placebo-controlled trials. *BMC Musculoskelet Disord* 2007;8:51.
24 Bjordal JM, Klovning A, Ljunggren AE, Slordal L. Short-term efficacy of pharmacotherapeutic interventions in osteoarthritic knee pain: a meta-analysis of randomised placebo-controlled trials. *Eur J Pain* 2007;11:125-38.
25 Manheimer E, Linde K, Lao L, Bouter LM, Berman BM. Meta-analysis: acupuncture for osteoarthritis of the knee. *Ann Intern Med* 2007;146:868-77.
26 Pildal J, Hrobjartsson A, Jorgensen K, Hilden J, Altman D, Gotzsche P. Impact of allocation concealment on conclusions drawn from meta-analyses of randomized trials. *Int J Epidemiol* 2007;36:847-57.
27 Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med* 2001;134:663-94.
28 Egger M, Juni P, Bartlett C. Value of flow diagrams in reports of randomized controlled trials. *JAMA* 2001;285:1996-9.
29 Gravel J, Opatrny L, Shapiro S. The intention-to-treat approach in randomized controlled trials: are authors saying what they do and doing what they say? *Clin Trials* 2007;4:350-6.
30 Huwiler-Muntener K, Juni P, Junker C, Egger M. Quality of reporting of randomized trials as a measure of methodologic quality. *JAMA* 2002;287:2801-4.
31 Fergusson D, Aaron SD, Guyatt G, Hébert P. Postrandomisation exclusions: the intention to treat principle and excluding patients from analysis. *BMJ* 2002;325:652-4.
32 Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *Ann Intern Med* 2001;134:657-62.
33 Sackett DL, Gent M. Controversy in counting and attributing events in clinical trials. *N Engl J Med* 1979;301:1410-2.
34 Baron G, Boutron I, Giraudeau B, Ravaud P. Violation of the intent-to-treat principle and rate of missing data in superiority trials assessing structural outcomes in rheumatic diseases. *Arthritis Rheum* 2005;52:1858-65.
35 Kim M. Statistical methods in Arthritis & Rheumatism: current trends. *Arthritis Rheum* 2006;54:3741-9.
36 Julious SA, Mullee MA. Issues with using baseline in last observation carried forward analysis. *Pharm Stat* 2008;7:142-6.
37 Streiner DL. The case of the missing data: methods of dealing with dropouts and other research vagaries. *Can J Psychiatry* 2002;47:68-75.
38 Baron G, Ravaud P, Samson A, Giraudeau B. Missing data in randomized controlled trials of rheumatoid arthritis with radiographic outcomes: a simulation study. *Arthritis Rheum* 2008;59:25-31.

Accepted: 23 March 2009

pico

Understanding why some ethnic minority patients evaluate medical care more negatively than white patients: a cross sectional analysis of a routine patient survey in English general practices

Nicola Mead,[1] Martin Roland[2]

EDITORIAL by Skeikh

[1]NIHR School for Primary Care, University of Manchester, Manchester M13 9PL
[2]General Practice and Primary Care Research Unit, University of Cambridge, Cambridge CB2 0SR
Correspondence to: Martin Roland
martin.roland@medschl.cam.ac.uk

Cite this as: *BMJ* 2009;339:b3450
doi: 10.1136/bmj.b3450

STUDY QUESTION Why do ethnic minority survey respondents give poor ratings of primary care compared with white patients?

SUMMARY ANSWER Relatively negative evaluations by ethnic minority patients of waiting times to see general practitioners and of continuity of care are partly explained by reported poor experiences of these aspects of care, but may also reflect communication problems and different expectations of waiting times for appointments.

WHAT IS KNOWN AND WHAT THIS PAPER ADDS Studies in the United States and the United Kingdom have shown that ethnic minority patients evaluate their care more negatively than do white patients. We found important differences in assessments of quality of care by different ethnic groups even after adjusting for factors that might account for observed differences.

Participants and setting

Between April 2005 and March 2006, 1098 English general practices undertook routine surveys of patients using the General Practice Assessment Questionnaire (GPAQ). The study analysed surveys from 188 572 respondents, of whom 95.8% classed themselves as "white", "black/black British," "Asian/Asian British," or "Chinese."

Design

We undertook a cross sectional analysis of patient survey data. Regression modelling was used to examine differences between white and ethnic minority patients' ratings of four aspects of primary care adjusting for mode of questionnaire administration, demographic factors, health need, and variation in reported standards of received care.

Primary outcome(s)

Patient evaluations of waiting times for GP appointments, time spent waiting in surgeries for consultations to start, and continuity of care were the primary outcomes.

Main results and the role of chance

Respondents from the three minority ethnic groups rated all aspects of care more than 2.5% lower than did white patients. Poorer evaluations of time spent waiting for consultations to begin (rated lowest by Asian patients) and continuity of care (rated lowest by Chinese patients) appeared to reflect worse reported experiences by ethnic minority groups. However, substantial differences between white and ethnic minority patients' ratings of appointment waiting times persisted even after adjusting for the actual time patients reported waiting. This effect disappeared for Chinese and black respondents after adjusting for evaluations of reception staff and doctors' communication skills, but Asian patients' ratings remained more than 2.5% lower than those of white respondents.

Bias, confounding and other reasons for caution

The simple measure of ethnicity used in this study may have masked important differences within categories (such as between patients of Indian, Pakistani, and Bangladeshi origin all categorised as "Asian"). The key variables in the study are patients' reported experiences of care and their evaluations of those experiences; it cannot be assumed these reported data reflect actual quality of care in practices.

Generalisability to other populations

Data were from routine surveys, thus no details of response rates are available. Surveys were completed in English mainly by consulting patients, who cannot be assumed representative of the wider population or of populations in other settings.

Study funding/potential competing interests

This study was conducted by the National Primary Care Research and Development Centre and supported by funding from the Department of Health and the National Institute for Health Research.

This is a summary of a paper that was published on bmj.com as *BMJ* 2009;339:b3450

SUMMARY OF ETHNIC MINORITY GROUPS THAT RATED EACH ASPECT OF CARE SUBSTANTIALLY LOWER THAN WHITE RESPONDENTS*

Evaluation of:	Unadjusted analysis	Adjusted additionally for:					
		Mode of administration	Demographic factors	Health need	Reported experience of aspect of care evaluated	Evaluation of receptionists	Evaluation of GP communication
Waiting time for appointments with a particular GP	Black	Black	–	–	Black	–	–
	Asian	Asian	Asian	Asian	Asian	Asian	Asian
	Chinese	Chinese	Chinese	Chinese	Chinese	–	–
Waiting time for appointments with any GP	Black	Black	Black	Black	Black	–	–
	Asian	Asian	Asian	Asian	Asian	Asian	Asian
	Chinese	Chinese	Chinese	Chinese	Chinese	–	–
Waiting time for consultations to begin	Black	Black	Black	Black	–	–	–
	Asian	Asian	Asian	Asian	–	–	–
	Chinese	Chinese	Chinese	Chinese	–	–	–
Continuity of care	Black	Black	–	–	–	–	–
	Asian	Asian	Asian	Asian	–	–	–
	Chinese	Chinese	Chinese	Chinese	–	–	–

*Cells show ethnic minority groups with a mean score at least 2.5 percentage points lower than that of white respondents on the evaluation item in question, adjusting for other independent variables named in that column heading and all those to the left

pico

The benefits of steroids versus steroids plus antivirals for treatment of Bell's palsy: a meta-analysis

Eudocia C Quant,[1 2] Shafali S Jeste,[3 2] Rajeev H Muni,[4] Alison V Cape,[5 2] Manveen K Bhussar,[6] Anton Y Peleg[2 6 7]

[1]Department of Neurology, Massachusetts General Hospital, Boston, MA 02114
[2]Harvard Medical School, Boston, MA 02115
[3]Division of Neurology, Children's Hospital Boston, Boston, MA 02115
[4]Department of Ophthalmology and Vision Sciences, University of Toronto, St Michael's Hospital, Toronto, Ontario M5B 1W8, Canada
[5]Department of Obstetrics and Gynecology, Brigham and Women's Hospital, Boston, MA 02115
[6]Division of Infectious Diseases, Beth Israel Deaconess Medical Center, Boston, MA 02215
[7]Division of Infectious Diseases, Massachusetts General Hospital, Boston, MA 02114

Correspondence to: A Y Peleg
apeleg@bidmc.harvard.edu

Cite this as: *BMJ* 2009;339:b3354
doi: 10.1136/bmj.b3354

STUDY QUESTION Does adding antivirals to steroids provide a better degree of facial muscle recovery in patients with Bell's palsy than treatment with steroids alone?

SUMMARY ANSWER In this meta-analysis, antivirals did not provide an added benefit in achieving at least partial facial muscle recovery compared with steroids alone in patients with Bell's palsy.

Selection criteria for studies

PubMed, Embase, Web of Science, and the Cochrane Central Register of Controlled Trials were searched for studies published in all languages between 1984 and January 2009. Additional studies were identified from cited references. We included all randomised controlled trials that compared steroids with the combination of steroids and antivirals in patients with Bell's palsy. Additional inclusion criteria were at least one month of follow-up and a primary end point of facial muscle recovery. Odds ratios and 95% confidence intervals were calculated and pooled using a random effects model.

Primary outcome

The primary outcome was the proportion of patients with at least partial facial muscle recovery from Bell's palsy at longest follow-up point and who attended a follow-up visit at least one month after the initiation of treatment. Partial facial muscle recovery was defined as a House-Brackmann grade of at least 2 or an equivalent score using an alternative scoring system.

Main results and role of chance

Six trials involving 1145 patients were included; 574 patients received steroids alone and 571 patients received steroids and antivirals. The pooled odds ratio for facial muscle recovery showed no benefit of steroids plus antivirals compared with steroids alone (odds ratio 1.50, 95% confidence interval 0.83 to 2.69; P=0.18, fig). A one study removed analysis showed that the highest quality studies had the greatest effect towards showing no difference between study arms. Subgroup analyses assessing causes of heterogeneity defined a priori (that is, time from symptom onset to treatment, length of follow-up, and type of antiviral studied) showed no benefit of antivirals in addition to steroids.

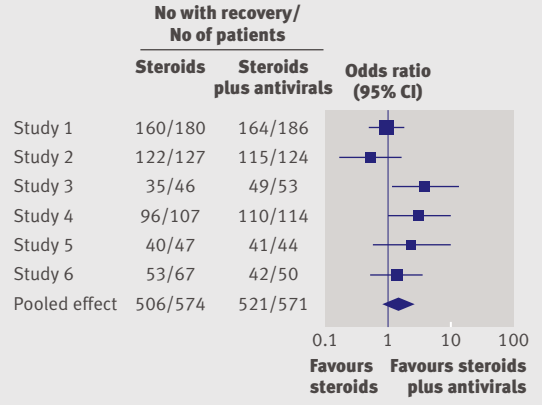
Bias, confounding, and other reasons for caution

It is unclear whether incremental benefits from the addition of antivirals to steroids depend on the severity of facial muscle paralysis at presentation, the extent of facial recovery, or the time to facial recovery. The inclusion of patients with Bell's palsy caused by *Varicella zoster* may dilute the potential benefit of antiviral therapy because this virus is less sensitive to antivirals than are other viruses, and the doses used in treatment trials are not high enough to treat Varicella zoster infection. Our subgroup analyses are limited by the small number of included studies and thus may lack statistical power. Finally, we were not able to perform intention to treat analysis because one study did not report such data; therefore, we used the number of patients at final follow-up as our denominator, as these data were available for all studies.

Study funding/potential competing interests

No specific funding was received for this study. AYP has acted as an adviser for Abbott Molecular. All other authors have no competing interests.

FOREST PLOT OF THE SIX INCLUDED STUDIES ORDERED ACCORDING TO JADAD SCORE (HIGHEST QUALITY TO LOWEST)



BMJ pico: advice to authors

The full text of all accepted *BMJ* research articles is published online in full, with open access and no word limit, on bmj.com as soon as it is ready. In the print *BMJ* each research article is abridged, as a one page BMJ pico, with the aim of making research more inviting and useful to readers. Starting in August 2009, authors are writing their own BMJ picos.

We have designed BMJ pico with evidence based medicine experts to succinctly present the key evidence from each study, to help minimise delay between online and print publication, and to enable us to publish more research in each week's print *BMJ*. For more details, see http://tinyurl.com/kp5c7o/.

There is no need for authors to prepare a BMJ pico to submit along with the full research article. Authors produce their own BMJ pico, using a template from us, only after the full article has been accepted.

Because publication of research on bmj.com is definitive, rather than interim "epublication ahead of print," authors who do not wish to abridge their articles using BMJ pico will be able to opt for online only publication.

This is a summary of a paper that was published on bmj.com as *BMJ* 2009;339:b3354