Open Access: free full text available online
For the full versions of these articles see bmj.com

Changing perceptions of weight in Great Britain: comparison of two population surveys

F Johnson, L Cooke, H Croker, Jane Wardle

EDITORIAL by Bleich

Cancer Research UK Health Behaviour Research Centre, Department of Epidemiology and Public Health, University College London, London WC1E 6BT

Correspondence to: J Wardle
j.wardle@ucl.ac.uk

Cite this as: *BMJ* 2008;337:a494
doi:10.1136/bmj.a494

ABSTRACT

Objectives To examine changes in public perceptions of overweight in Great Britain over an eight year period.

Design Comparison of data on self perceived weight from population surveys in 1999 and 2007.

Setting Household surveys of two representative samples in Great Britain.

Participants 853 men and 944 women in 1999, and 847 men and 989 women in 2007.

Main outcome measures Participants were asked to report their weight and height and classify their body size on a scale from "very underweight" to "obese."

Results Self reported weights increased dramatically over time, but the weight at which individuals perceived themselves to be overweight also rose significantly. In 1999, 81% of overweight participants correctly identified themselves as overweight compared with 75% in 2007, demonstrating a decrease in sensitivity in the self diagnosis of overweight.

Conclusions Despite media and health campaigns aiming to raise awareness of healthy weight, increasing numbers of overweight people fail to recognise that their weight is a cause for concern. This makes it less likely that they will see calls for weight control as personally relevant.

INTRODUCTION

A considerable proportion of overweight adults—men in particular—do not recognise that their body weight is too high,[1-3] and many parents fail to recognise that their children are overweight.[4-6]

Changes in the social environment over recent years could have affected weight perceptions in several ways. Increased attention to the "obesity epidemic" and publicity channelled through the media and health professionals to encourage appropriate action for weight control[7 8] might be expected to promote recognition of overweight. There has also been an emphasis on positive body images for young women, which should have reduced inaccurate perceptions of overweight among normal weight women. On this basis, weight recognition should have become more accurate.

Accuracy in self diagnosing overweight can be approached with the diagnostic concepts of sensitivity and specificity. Sensitivity is the proportion of truly overweight people who identify themselves as such, while specificity is the proportion of people who are not overweight who identify themselves correctly as not overweight. If the combined emphasis on public awareness of the risks of obesity and promotion of a healthy body image in young women has been successful, then both sensitivity and specificity of self diagnosed overweight should have increased.

We investigated changes in public perception of overweight over an eight year period, and assessed effects on the self diagnostic abilities of overweight and normal weight British adults.

METHODS

Study design and participants

We compared self reported weights and perceptions of weight from two population based surveys carried out eight years apart. The first survey was carried out through the Office for National Statistics (ONS) omnibus survey of March 1999. Further details can be obtained at www.ons.gov.uk/about/who-we-are/our-services/omnibus-survey.

The second survey was a face to face omnibus survey conducted in May 2007 by the British Market Research Bureau (BMRB). Further information is available from www.bmrb.co.uk/?id=755. In both surveys the interviews were undertaken in the home with only one interview per household. Both produced samples that closely resembled the demography of the population of Great Britain.

Measures

Anthropometric data—Weight and height were self reported in whichever metric the individual preferred. Use of self reported anthropometric data means that height is likely to be overestimated and weight underestimated,[9 10] and therefore average BMI and the proportion of the population who are overweight or obese will be underestimated in both samples. Participants were divided into weight groups using BMI cut offs of <18.5 (underweight), >25 (overweight), and >30 (obese).

Perceived weight—Participants were asked to select a descriptor for their own body weight from the following list: very underweight, underweight, about right, overweight, and very overweight. The 2007 survey also included the category "obese." For most of

This article is an abridged version of a paper that was published on bmj.com. Cite this article as: *BMJ* 2008;337:a494

Prevalence of overweight and sensitivity and specificity of recognition of overweight in men and women, 1999 and 2007

	Prevalence/sensitivity/specificity (95% CI), No in group		
	1999	2007	χ^2, P value
Prevalence			
Men	0.50 (0.46 to 0.53), 903	0.57 (0.54 to 0.60), 916	7.72, P=0.005
Women	0.36 (0.33 to 0.40), 895	0.49 (0.46 to 0.53), 921	25.65, P<0.001
All	0.43 (0.40 to 0.45), 1798	0.53 (0.51 to 0.55), 1837	30.15, P<0.001
Sensitivity			
Men	0.75 (0.70 to 0.79), 446	0.67 (0.62 to 0.71), 512	3.09, P=0.079
Women	0.90 (0.86 to 0.93), 331	0.83 (0.80 to 0.87), 449	10.40, P=0.001
All	0.81 (0.78 to 0.84), 777	0.75 (0.72 to 0.78), 961	8.02, P=0.005
Specificity			
Men	0.86 (0.83 to 0.89), 457	0.90 (0.87 to 0.93), 404	3.51, P=0.061
Women	0.74 (0.70 to 0.77), 564	0.78 (0.75 to 0.82), 472	6.47, P=0.011
All	0.79 (0.76 to 0.81), 1012	0.83 (0.81 to 0.86), 876	9.87, P=0.002

the analyses reported here, we dichotomised the data into a "perceived overweight" group, comprising the top two (three in 2007) categories, versus the rest.

Data analysis

We used *t* tests and χ^2 analyses for comparisons between the two surveys and further examined perceptions of overweight using log binomial regression, with dichotomised perceived overweight as the dependent variable. Independent variables were age, age on leaving education, sex, survey year, and weight group. We calculated specificity and sensitivity of perceptions of overweight and 95% confidence intervals.

RESULTS

Sample characteristics, 1999 and 2007

In 1999, 1894 interviews were carried out, comprising 882 men and 1012 women. Adequate weight and height data were collected from 853 men and 944 women. The 2007 sample of 1998 participants comprised 895 men and 1103 women, of whom 847 men and 989 women provided adequate data on height and weight.

There was no significant difference in sex balance between the two time points: 53% women in 1999 and 54% women in 2007. There was no overall difference in age, but women in the 1999 sample were slightly older than those in the 2007 sample (t=3.05, P<0.01). See bmj.com. There was a significant difference in age of completing education (t=2.61, P<0.01) with participants in the 2007 sample being slightly older. This was probably because of increases in the legal school leaving age.

Weight and perceptions of overweight

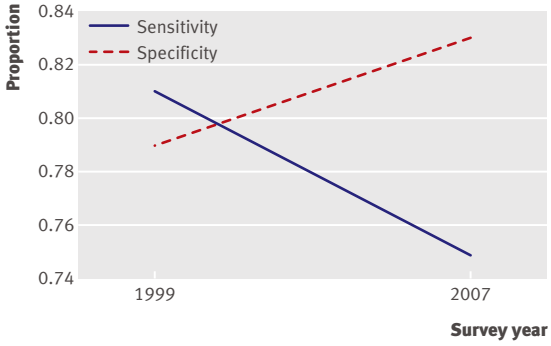
Height did not differ significantly between the samples, but both weight and BMI were higher in 2007 (t=6.09, P<0.001, and t=7.77, P<0.001, respectively). The proportion of respondents whose BMI placed them in the obese category had nearly doubled, from 11% to 19%.

In contrast with the upward trends in overweight and obesity, trends in perceived overweight were downward. In 1999, 43% of the population had a BMI that put them in the overweight or obese range, of whom 81% perceived themselves to be overweight or very overweight. In 2007, 53% of the population had a BMI in the overweight or obese range, of whom only 75% reported themselves to be overweight, very overweight, or obese.

We used log binomial regression to establish the significance of differences in weight perceptions between 1999 and 2007, controlling for differences in demographic composition of the samples. Age on leaving education did not achieve significance, but all other independent variables (age, weight group, sex, and survey year) were significant predictors of perceived overweight. See bmj.com.

Sensitivity and specificity of self perception of overweight

We examined sensitivity and specificity of weight perceptions, together with 95% confidence intervals, for the two samples (table). In the sample as a whole, sensitivity of recognition of overweight decreased



Sensitivity and specificity of perception of overweight. Sensitivity denotes proportion of overweight participants who correctly identify themselves as overweight. Specificity denotes proportion of normal and underweight participants who correctly identify that they are not overweight

WHAT IS ALREADY KNOWN ON THIS TOPIC

Perceptions of overweight in the population do not correspond well to the definitions used by health professionals

Many overweight and obese individuals fail to recognise that their weight is too high

WHAT THIS STUDY ADDS

As the proportion of overweight people in Great Britain has increased, the ability of overweight individuals to "self diagnose" their weight problem has declined

between 1999 and 2007, alongside an increase in specificity. When we analysed data for men and women separately, we found a similar pattern of results for both groups, but results for the men did not reach significance. The figure shows changes in sensitivity and specificity.

DISCUSSION

Despite the topic of weight scarcely being out of the news, these data from two population surveys show that fewer overweight and obese people defined themselves as overweight in 2007 than in 1999. The changes indicate a marked decline in sensitivity with respect to individuals' detection of their own overweight. There was a concurrent improvement in specificity, with fewer people of normal or low weight believing themselves to be overweight. These effects were strongest in women, marginally failing to reach significance in men.

Interpretation

A decline in sensitivity of recognition of overweight has important implications for the targeting of public health messages, which are unlikely to reach marginally overweight individuals if they fail to identify themselves as targets.

Increased attention to the health risks of excess weight might have left individuals more reluctant to identify themselves with labels such as "overweight" or "obese." Health professionals must discuss body weight in a way that is precise enough and neither minimises the risks associated with excess weight nor provokes disengagement on the part of the patient.

Slightly fewer normal weight women think they are overweight. This has implications for practitioners and policy makers working in the field of eating disorders as well as obesity prevention.

Explaining changing weight perceptions

Various factors might have contributed to the declining ability of overweight individuals to recognise that their weight is too high. Social comparison is likely to play an important role, resulting in the threshold for perceived overweight rising in line with increasing weight in the population. International data have suggested that perceptions of overweight are related to levels of overweight in the local population, supporting the social norm hypothesis.[11]

Also images that accompany media and health information often depict severely obese people, untypical of the overweight population. This might act as false reassurance for those who are "merely" overweight, implicitly reinforcing a perception that messages about healthy eating and exercise are not aimed at them.

Strengths and weaknesses of the study

Women in the 2007 survey were slightly older, and both men and women report more years of schooling. Inclusion of these as covariates in the analyses, however, did not change the findings. Data collection methods were not identical between the two surveys, but both samples produced were representative of the population and in both cases a computer assisted face to face survey was used.

The use of self reported heights and weights facilitates large scale data collection but is always a source of inaccuracy, resulting in underestimates of weight (particularly in women) and overestimates of height (particularly in men).[9 10] It is therefore likely that BMI and the prevalence of overweight are underestimated in both samples.

Data collection was carried out by the Office for National Statistics (ONS) and the British Market Research Bureau (BMRB).
Contributors: See bmj.com.
Funding: Cancer Research UK and the Economic and Social Research Council.
Competing interests: None declared.
Ethical approval: Not required.
Provenance and peer review: Not commissioned; externally peer reviewed.

1 Jeffery AN, Voss LD, Metcalf BS, Alba S, Wilkin TJ. Parents' awareness of overweight in themselves and their children: cross sectional study within a cohort (EarlyBird 21). *BMJ* 2005;330:23-4.
2 Wardle J, Johnson F. Weight and dieting: examining levels of weight concern in British adults. *Int J Obes Relat Metab Disord* 2002;26:1144-9.
3 Crawford D, Campbell K. Lay definitions of ideal weight and overweight. *Int J Obes Relat Metab Disord* 1999;23:738-45.
4 Carnell S, Edwards C, Croker H, Boniface D, Wardle J. Parental perceptions of overweight in 3-5 y olds. *Int J Obes (Lond)* 2005;29:353-5.
5 Eckstein KC, Mikhail LM, Ariza AJ, Thomson JS, Millard SC, Binns HJ. Parents' perceptions of their child's weight and health. *Pediatrics* 2006;117:681-90.
6 Miller JC, Grant AM, Drummond BF, Williams SM, Taylor RW, Goulding A. DXA measurements confirm that parental perceptions of elevated adiposity in young children are poor. *Obesity (Silver Spring)* 2007;15:165-71.
7 International Food Information Council (IFIC) Foundation and the Center for Media and Public Affairs. *Food for thought VI. Reporting of diet, nutrition and food safety*. Washington, DC: IFIC Foundation, 2005.
8 Evans WD, Renaud JM, Kamerow DB. News media coverage, body mass index and public attitudes about obesity. *Soc Mark Q* 2006;12:19-33.
9 Cameron R, Evers SE. Self-report issues in obesity and weight management: state of the art and future directions. *Behav Assess* 1990;12:91-106.
10 Roberts RJ. Can self-reported data accurately describe the prevalence of overweight? *Public Health* 1995;109:275-84.
11 Wardle J, Haase AM, Steptoe A. Body image and weight control in young adults: international comparisons in university students from 22 countries. *Int J Obes (Lond)* 2006;30:644-51.

Accepted: 13 May 2008

Evaluation of *Talking Parents, Healthy Teens*, a new worksite based parenting programme to promote parent-adolescent communication about sexual health: randomised controlled trial

Mark A Schuster,[1,2,4] Rosalie Corona,[3,4] Marc N Elliott,[2,4] David E Kanouse,[2,4] Karen L Eastman,[4] Annie J Zhou,[2,4] David J Klein[2,4]

EDITORIAL by Kirby

[1]Children's Hospital Boston and Harvard Medical School, Boston, MA 02115, USA
[2]RAND, Santa Monica, CA 90407
[3]Department of Psychology, Virginia Commonwealth University, Richmond, VA 23284
[4]UCLA/RAND Center for Adolescent Health Promotion, University of California, Los Angeles, CA 90095
Correspondence to: M A Schuster, Children's Hospital Boston, 300 Longwood Ave, Boston, MA 02115 mark.schuster@childrens.harvard.edu

Cite this as: *BMJ* 2008;337:a308
doi:10.1136/bmj.39609.657581.25

This article is an abridged version of a paper that was published on bmj.com. Cite this article as: *BMJ* 2008;337:a308

ABSTRACT

Objective To evaluate a worksite based parenting programme—*Talking Parents, Healthy Teens*—designed to help parents learn to address sexual health with their adolescent children.

Design Randomised controlled trial (April 2002-December 2005).

Setting 13 worksites in southern California.

Participants 569 parents completed baseline surveys at work, gave permission for confidential surveys to be posted to their adolescent children, and were randomised to intervention or control groups. Parents and adolescents completed follow-up surveys at one week, three months, and nine months after the programme.

Intervention *Talking Parents, Healthy Teens* consists of eight weekly one hour sessions at worksites for parents of adolescent children aged in 6th-10th grade (about ages 11-16 years).

Main outcome measures Parent-adolescent communication about a list of sexual topics; whether parent taught adolescent how to use a condom; ability to communicate with parent/adolescent about sex; openness of parent-adolescent communication about sex.

Results Differences between intervention and control groups were significant for the mean number of new sexual topics that parents and adolescents reported discussing between baseline and each follow-up (P<0.001 for each); intervention parents were less likely than controls to discuss no new topics (8% *v* 29%, 95% confidence interval for difference 16% to 24%) and more likely to discuss seven or more new topics (38% *v* 8%, 19% to 41%) at nine months. Some differences increased after completion of the programme: at one week after the programme, 18% of adolescents in the intervention group and 3% in the control group (6% to 30%) said that their parents had reviewed how to use a condom since baseline (P<0.001); this grew to 29% *v* 5% (13% to 36%) at nine months (P<0.001). Compared with controls at nine months, parents and adolescents in the intervention group reported greater ability to communicate with each other about sex (P<0.001) and more openness in communication about sex (P<0.001).

Conclusions A worksite based programme can have substantial effects on communication between parents and adolescents about sexual health.

Trial registration Clinical Trials NCT00465010.

INTRODUCTION

Parents generally have a distinct emotional bond with[1] and influence over their children[1 2] and can tailor conversations to children's cognitive, social, emotional, and physical development and needs.[3] Thus, they might be in a unique position to promote healthy adolescent sexual development. Yet, many parents do not talk to adolescents, particularly younger adolescents, about sexual topics.[4] Parents report feeling embarrassed and inadequately informed[5] and unsure of what to say or how to begin.[6]

Programmes to teach parents to communicate about sexual health in community settings (such as schools) can be impractical, especially for employed parents, because of scheduling and location issues. A promising but untested approach is to bring the intervention to parents where they work.[7] We developed *Talking Parents, Healthy Teens*, a theory based worksite parenting programme to help parents become more comfortable and skilled at communicating with adolescents about sexual health.[8]

METHODS

Intervention

Talking Parents, Healthy Teens provides eight weekly one hour sessions to groups of about 15 parents of children in 6th-10th grade (about ages 11-16) at their worksite during the lunch hour (with free lunch). The programme uses roleplay, videotaped interactions, games, and discussions to help parents communicate with children about sex related topics; teach their children communication, assertiveness, and decision making skills; and better supervise and interact with their children. Home assignments strengthen parent-child relationships, and handouts review programme content and sexual health topics. Formative qualitative research[6 7] and three pilot tests informed programme development.

Study design

We conducted a randomised controlled trial, with randomisation at the individual parent level. Parents were recruited from 13 large public and private (for profit and non-profit) worksites in southern California through worksite email, newsletters, etc. Parents completed baseline surveys a month before the programme; adolescents also completed their surveys

before the programme; and both completed follow-up surveys at one week, three months, and nine months after the programme.

Outcome variables

Dependent variables were measured at each follow-up. For the first two variables below (discussions and condom instruction), retrospective time frames were "ever" at baseline and the time period between each follow-up.

Discussions—For each adolescent, parents reported whether they had discussed 24 sex related topics (for example, "how to choose a method of birth control") (box).

Condom instruction—Adolescents responded yes/no to "My [mother/father] has reviewed the steps of how to use a condom with me." This item appeared in a battery of items about condoms, which were not asked of parents.

Communication ability—Parents and adolescents were asked, "How would you rate your ability to communicate with your [child/mother/father] about sexual topics?" and responded on a seven point response scale from excellent (7) to terrible (1).

Sex related discussion topics from parent and adolescent surveys

- How girls' bodies change physically as they grow up
- How boys' bodies change physically as they grow up
- Menstruation (having menstrual periods)
- Wet dreams (parents and boys only)
- How women get pregnant and have babies
- Masturbation (parents only)
- What qualities are important in choosing close friends
- How to ask someone out on a date
- How [you/your child] will make decisions about whether or not to have sex
- What it feels like to have sex
- Homosexuality/people being gay
- Consequences of getting pregnant/getting someone pregnant
- How well birth control can prevent pregnancy
- How well condoms can prevent sexually transmitted diseases (STDs)
- How to choose a method of birth control
- How to use a condom
- How people can prevent getting sexually transmitted diseases (STDs)
- Symptoms of sexually transmitted diseases (STDs)
- What to do if a partner doesn't want to use a condom
- The importance of not pressurising other people to have sex
- Reasons why people like to have sex
- Reasons why [you/your child] should not have sex
- How [you/your child] will know if [you/he/she] is in love
- How to say no if someone wants to have sex and [you/your child don't/doesn't] want to

Some items were omitted for adolescents based on feedback during the formative phase of the study: 24 items were asked of parents, 23 of boys, and 22 of girls. Results change little if analyses of parents' responses were limited to the 22 topics asked of both boys and girls or if all 23 topics are included for boys.

Communication openness—We used a 12 item parent scale to assess openness of parent-adolescent communication about sexual topics (such as, "My child [mother/father] and I talk openly and freely about sexual topics") using a four point strongly agree/strongly disagree scale (α=0.86). We used a similar seven item measure (α=0.85) for adolescents, with higher scores indicating more openness.

Independent variable

Our independent variable was intervention status (whether the person was in the intervention or control group).

Analysis

Count and ordinal outcomes were modelled with linear regression because results were insensitive to the normal residual assumption.[9] The dichotomous outcome was modelled with logistic regression.

We tested the intervention's effectiveness using parent and adolescent reports (separately) with models that predicted an outcome at each follow-up from both the intervention status and that outcome's baseline value. Tests of intervention effects analysed participants in their original randomised group, regardless of attendance. For more details on the methods see bmj.com.

RESULTS

Participants

We randomised 569 parents (with 683 participating adolescents), 288 to the intervention group and 281 to the control group. Parents attended a median of seven sessions. One week after the programme, 20% of control parents reported talking to intervention parents about the programme or seeing or reading the programme's educational materials.

New topics

At baseline, parents on average reported having ever discussed 8.9 of 24 topics (SD 5.5) with their adolescents; adolescents reported 7.2 of 22 topics (SD 5.3). Between baseline and one week after the programme, intervention parents reported discussing more new topics than control parents (mean 4.0 *v* 0.8, 95% confidence interval of the difference 2.7 to 3.6; P<0.001). This difference persisted at a similar magnitude and significance for newly incident discussions at three and nine months after the programme (fig 1).

There was a large difference between the intervention and control groups for parents in the mean number of new topics discussed since baseline (P<0.001) (table): intervention parents were much less likely than control parents to discuss no new topics and much more likely to discuss seven or more new topics. Similarly, 33% of adolescents in the intervention group and 13% in the control group (12% to 30%) reported seven or more new topics (P<0.001).

New sexual topics discussed for first time from baseline to nine months after intervention, reported by parents and adolescents. Figures are percentages (95% confidence intervals) of participants who reported each number of new topics

No of new topics	Control	Intervention*
Reported by parents:		
0	29 (24 to 34)	8 (5 to 11)
1	20 (15 to 24)	11 (7 to 14)
2	16 (12 to 21)	10 (7 to 13)
3-6	27 (21 to 32)	33 (28 to 38)
≥7	8 (5 to 12)	38 (32 to 44)
Reported by adolescents:		
0	31 (26 to 36)	18 (14 to 22)
1	19 (15 to 24)	11 (8 to 15)
2	9 (6 to 12)	8 (5 to 10)
3-6	28 (23 to 33)	30 (24 to 35)
≥7	13 (9 to 17)	33 (28 to 39)

*Intervention group discussed significantly more topics than control group (P<0.001) in linear regressions, predicting uncollapsed count of new topics discussed after baseline and adjusted for number of topics discussed at baseline. Separate models run for parents and adolescents.

Repeated topics

On average, intervention parents repeated 2.7 more topics than controls (P<0.001) one week after the programme. The difference grew through the three (P<0.001) and nine month (P=0.003) follow-up surveys (fig 1).

Condom teaching

At baseline, 4% of adolescents reported that their parent had reviewed how to use a condom. One week after the programme, more adolescents in the intervention than in the control group reported receiving this instruction since baseline (18% *v* 3%, 6% to 30%; P<0.001). The difference between the groups grew by nine months (29% *v* 5%, 13% to 36%; P<0.001).

Communication ability

At baseline parents rated their ability to discuss sexual topics with their adolescent as between "fair" and "good" (mean 4.6, SD 1.3) on average. Although control parents showed no significant change, intervention parents showed an increase from baseline relative to controls that remained significant throughout (P<0.001 at each follow-up) (fig 2).

Adolescents rated their baseline ability to discuss sexual topics with their parents as "fair" (mean 4.2, SD 1.8). Reported ability in adolescents in the control group declined over the follow-up surveys (P<0.001, linear trend test). Adolescents in the intervention group significantly differed from those in the control group in their ability to communicate about sexual topics at three months (mean 4.3 *v* 4.0, 0.1 to 0.5; P=0.02). At nine months, the decline in the control group was not evident in the intervention group, with a significant difference between the two (P<0.001).

Communication openness

After the programme parents in the intervention group reported significantly higher scores on a scale measuring openness of communication about sexual matters, compared with their scores at baseline (P<0.001 at each follow-up) and with scores in the control group (P<0.001 at each follow-up). Scores on the openness of communication scale in adolescents in the control group declined from baseline (P=0.006 at nine months), whereas scores for the intervention group improved compared with baseline (P=0.005), with higher scores at each follow-up (P<0.001 for each) (fig 2).

DISCUSSION

Reports from both parents and adolescents indicate that the *Talking Parents, Healthy Teens* programme had significant immediate effects on parent-adolescent communication. Compared with control parents, intervention parents reported more conversations about new sexual topics and more repeated conversations about topics that they had previously discussed. Reported improvement in adolescents paralleled parents' reports. The differences between the groups grew over time for parents' reports on repeated talks and adolescents' reports on new and repeated talks.

A particularly dramatic illustration of the programme's impact was with adolescents' reports of whether parents taught them how to use condoms. At baseline, few parents had done this, but at the end of the programme, a large and significant difference between the groups emerged. At three and nine months, the

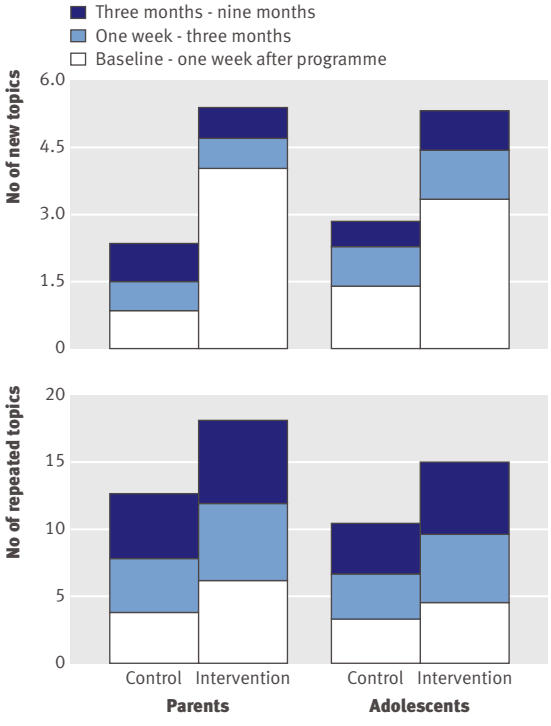


Fig 1 | New (top) and repeated (bottom) sexual topics discussed, reported by parents and adolescents. Cumulative difference between intervention and control from baseline to each subsequent survey was significant at P<0.001. For new topics, interval differences were not significant. For repeated topics, interval difference for parents was P<0.001 at three months and P=0.003 at nine months; for adolescents, difference was P<0.001 for each interval

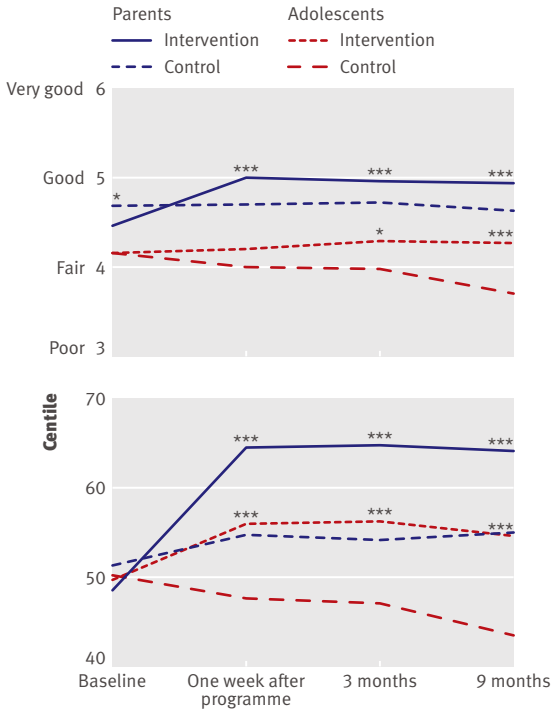


Fig 2 | **Top: ability to communicate with adolescent/parent about sexual topics (How would you rate your ability to communicate with your child [mother/father] about sexual topics? excellent/very good/good/fair/poor/very poor/terrible). Bottom: openness of parent-adolescent communication scale (12 item scale for parents; 7 item scale for adolescents). Because units of scale have no absolute meaning, we have expressed all scores on this scale as centile of baseline distribution (control and intervention combined) to which they correspond. In tests for linear time trend across survey waves, the parent control group does not change significantly over four surveys (P=0.14), while parent intervention group increases significantly (P=0.001); adolescent control group declines significantly over time (P=0.006), while adolescent intervention group increases (P=0.005). *P<0.05; ***P<0.001 for intervention-control comparisons, tested with a linear regression predicting communication at each survey and controlling for communication at baseline**

difference widened, indicating an ongoing influence of the programme on parents' behaviour.

Parents and adolescents in the intervention group reported not only more discussions than those in the control group but also a greater ability to talk about sexual matters and more openness discussing sex related topics with one another. These findings were sustained and continued to accumulate for at least nine months after the programme ended. By contrast, adolescents in the control group reported a decline over time for both measures. For communication ability, the intervention stemmed this natural decline, and for openness, the intervention caused absolute improvement. This ongoing expansion suggests that parents are not merely conveying static knowledge to adolescents but rather are altering sexual communication dynamics within the family in a way that might increase their effectiveness in promoting adolescent sexual health.

WHAT IS ALREADY KNOWN ON THIS TOPIC

Research shows that parents can significantly influence adolescents' sexual health and risk reduction through parent-adolescent relationships, parenting practices, and communication about sexual matters, but few interventions help parents raise sexually healthy youth

Worksite health promotion has been effective for adult health issues but has not been well studied as a way to influence adolescents

WHAT THIS STUDY ADDS

A worksite based programme for parents of adolescents can have significant immediate and longer term effects on parent-adolescent communication related to sexual health

Findings in context

This study's implications point to the potential to expand the scope of worksite health promotion beyond employees' own health. Our formative work found that employers showed enthusiasm for programmes that help parents of adolescents deal with adolescents.[6 7]

Because some control parents reported exposure, our study's substantial findings probably underestimate the true magnitude of the effects of intervention. It is also possible that those in the control group (as well as the intervention group) were prompted by the survey's list of discussion topics to engage in new or repeat conversations about particular topics, particularly given that these respondents were interested enough in communication to volunteer for the study.

Limitations

As with all self reported data, respondents might misreport answers and thereby create bias towards a particular result. Intervention parents might be inclined to report improved communication, but adolescents, who were not in the programme, would have little apparent reason to over-report communication. Our results might not be generalisable to all employed parents. Although participating worksites covered a broad cross section of large employers, smaller worksites might have different experiences implementing the programme. While our sample was diverse in terms of education, race/ethnicity, and manager/staff rank, incomes tended to be higher than average, consistent in the US with employment at large worksites, which might reduce generalisability to lower income parents at small worksites. Likewise, all the worksites were located in a large metropolitan area in the US; it will be important to test the intervention in various areas to assess its effectiveness in different populations. In addition, the offer of incentives for participation and the programme's implementation as part of a study could have influenced parents' willingness to participate. Some participants might have been attracted by the opportunity to participate in a study and receive monetary incentives, while others might not have volunteered because they were not interested in participating in a study and filling out surveys. Because most of the adolescents were middle school aged and because children of employed parents (who, on average, have higher household incomes) tend to have a later expected age of sexual initiation,[10] a

much longer follow-up period would be necessary to determine whether communication influences sexual behaviours. None the less, we have shown that *Talking Parents, Healthy Teens* significantly increases and improves parent-adolescent communication, which is not only linked with adolescent sexual health and risk reduction[11-13] but is also a worthy goal in itself.

We thank Rosa-Elena Garcia and the RAND Survey Research Group for data collection; Elaine Quiter, Diane Schoeff, and Katherine Vestal for project coordination; Susan Park and Gillian Smith for research assistance; Kate Sommers-Dawes, Scott Stephenson, and Katherine Woolsey for manuscript assistance; Laura Bogart and Steven Martino of RAND for comments on a draft of the manuscript; the programme's facilitators and assistant facilitators; and other UCLA/RAND Center for Adolescent Health Promotion staff who helped in the development, implementation, and evaluation of the programme. We also thank the centre's community advisory board and scientific advisory board for their guidance, the worksites for their assistance, and the study participants for sharing their time and experiences.

Contributors: See bmj.com.

Funding: Grant R01 MH61202 from the National Institute of Mental Health and cooperative agreements U48/CCU915773 and U48/DP000056 from the Centers for Disease Control and Prevention.

Competing interests: None declared.

Ethical approval: The institutional review boards of RAND and the University of California, Los Angeles approved the study protocol.

Provenance and peer review: Not commissioned; externally peer reviewed.

1 Lezin N, Rolleri L, Bean S, Taylor J. *Parent-child connectedness: implications for research, interventions, and positive impacts on adolescent health care*. Santa Cruz, CA: ETR Associates, 2004.
2 Meschke LL, Bartholomae S, Zentall SR. Adolescent sexuality and parent-adolescent processes: promoting healthy teen choices. *J Adolesc Health* 2002;31:264-79.
3 Kotchick BA, Shaffer A, Forehand R, Miller KS. Adolescent sexual risk behavior: a multi-system perspective. *Clin Psychol Rev* 2001;21:493-519.
4 Raffaelli M, Bogenschneider K, Flood MF. Parent-teen communication about sexual topics. *J Fam Issues* 1998;19:315-33.
5 Jaccard J, Dittus PJ, Gordon VV. Parent-teen communication about premarital sex: factors associated with the extent of communication. *J Adolesc Res* 2000;15:187-208.
6 Eastman KL, Corona R, Ryan GW, Warsofsky AL, Schuster MA. Worksite-based parenting programs to promote healthy adolescent sexual development: a qualitative study of feasibility and potential content. *Perspect Sex Reprod Health* 2005;37:62-9.
7 Schuster MA, Eastman KA, Fielding JE, Rotheram-Borus MJ, Breslow L, Franzoi LL, et al. Promoting adolescent health: worksite-based interventions with parents of adolescents. *J Public Health Man* 2001;7:41-52.
8 Eastman K, Corona R, Schuster M. Talking parents, healthy teens: a worksite-based program for parents to promote adolescent sexual health. *Prev Chron Dis* 2006;3:1-10.
9 Jarque CM, Bera AK. Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters* 1980;6:255-9.
10 Lammers C, Ireland M, Resnick M, Blum R. Influences on adolescents' decision to postpone onset of sexual intercourse: a survival analysis of virginity among youths aged 13 to 18 years. *J Adolesc Health* 2000;26:42-8.
11 DiClemente RJ, Wingood GM, Crosby R, Cobb BK, Harrington K, Davies SL. Parent-adolescent communication and sexual risk behaviors among African American adolescent females. *J Pediatr* 2001;139:407-12.
12 Hutchinson MK, Jemmott JB, Jemmott LS, Braverman P, Fong GT. The role of mother-daughter sexual risk communication in reducing sexual risk behaviors among urban adolescent females: a prospective study. *J Adolesc Health* 2003;33:98-107.
13 Dilorio C, Pluhar E, Belcher L. Parent-child communication about sexuality: a review of the literature from 1980-2002. *J HIV AIDS Prev Educ Adolesc Child* 2003;5:7-32.

Accepted: 2 May 2008

Academic achievement of twins and singletons in early adulthood: Taiwanese cohort study

Meng-Ting Tsou,[1,2] Meng-Wen Tsou,[3] Ming-Ping Wu,[4] Jin-Tan Liu[5]

EDITORIAL by Christensen and McGue

[1]Department of Family Medicine, Mackay Memorial Hospital, Taiwan
[2]Nursing and Management College, Mackay Medicine, Taiwan
[3]Department of International Trade, Tamkang University, Taiwan
[4]Division of Urogynecology and Pelvic Floor Reconstruction, Department of Obstetrics and Gynecology, Chi Mei Foundation Hospital, Taiwan
[5]Department of Economics, National Taiwan University, Taiwan

Correspondence to: J-T Liu
liujt@ntu.edu.tw

Cite this as: *BMJ* 2008;337:a438
doi:10.1136/bmj.a438

ABSTRACT

Objectives To examine the long term effects of low birth weight on academic achievements in twins and singletons and to determine whether the academic achievement of twins in early adulthood is inferior to that of singletons.

Design Cohort study.

Setting Taiwanese nationwide register of academic outcome.

Participants A cohort of 218 972 singletons and 1687 twins born in Taiwan, 1983-5.

Main outcome measure College attendance and test scores in the college joint entrance examinations.

Results After adjustment for birth weight, gestational age, birth order, and sex and the sociodemographic characteristics of the parents, twins were found to have significantly lower mean test scores than singletons in Chinese, mathematics, and natural science, as well as a 2.2% lower probability of attending college. Low birthweight twins had an 8.5% lower probability of college attendance than normal weight twins, while low birthweight singletons had only a 3.2% lower probability. The negative effects of low birth weight on the test scores in English and mathematics were substantially greater for twins than for singletons. The twin pair analysis showed that the association between birth weight and academic achievement scores, which existed for opposite sex twin pairs, was not discernible for same sex twin pairs, indicating that birth weight might partly reflect other underlying genetic variations.

Conclusions These data support the proposition that twins perform less well academically than singletons. Low birth weight has a negative association with subsequent academic achievement in early adulthood, with the effect being stronger for twins than for singletons. The association between birth weight and academic performance might be partly attributable to genetic factors.

INTRODUCTION

The cognitive disadvantage for twins during childhood has been well documented,[1-3] although a study from the

Netherlands reported little difference in intelligence between adult twins and singleton siblings,[4] and a more recent study from Denmark found the academic performance of twins and singletons aged 15-16 years was quite similar.[5]

We used the Taiwanese nationwide registers of academic outcomes, which is based on the results of the college joint entrance examination, to compare performance levels between twins and singletons. We undertook separate analyses of the effects of low birth weight on long term academic outcomes for twins and singletons and the effects on twin pairs.

METHODS

Data

We linked the Taiwanese nationwide academic outcomes in the college joint entrance examinations for 2002-3 to national birth certificate record data for 1983-5. The examination is the main way for Taiwanese students to enter universities and colleges.

Comprehensive details on sex, gestational age, birth weight, birth order, birth place, and multiple birth status are contained within the birth certificate files, along with details on the age, years of education, and working status of the parents at the time of the birth. The college joint entrance examination files contain information on whether the student was enrolled in a college and the test scores obtained in the five major subjects (Chinese, English, mathematics, natural science, and social science). The overall test score was measured as the sum of the five subjects.

We used the unique personal identifier to merge the college joint entrance examination files with the birth certificate files and restricted our sample to those born from September 1983 to August 1985. After excluding those with missing data on parental characteristics, we were left with 220 659 observations for our analysis, comprising 1687 twins and 218 972 singletons.

Statistical methods

We used the univariate comparison method to compare academic performance of twins and singletons. To account for potential confounding factors, we used logistic regression analysis to determine the probability of college attendance (dichotomous outcome) and multiple linear regression analysis for the test scores (continuous outcomes), with adjustment for sex, gestational age, birth weight, and birth order and socioeconomic characteristics of the parents. We examined the impact of low birth weight on long term academic outcomes by estimating the regressions separately for twins and singletons. Finally, focusing on the twin pair sample, we used the twin fixed effect approach to examine the effect of birth weight on academic achievement to provide further control of unobserved parental and environmental factors across families.

RESULTS

The average birth weight of twins (2570 g) was substantially lower than that of the singletons (3287 g). Over 40% of the twin births were classified as low birth weight (<2500 g) compared with only 3% of singletons. The gestational period tended to be shorter for twins (38.1 weeks) than singletons (39.7 weeks), while birth order was higher for twins (1.9) than singletons (1.8). On average, parents of twins were older at the time of birth than parents of singletons (mothers 27.6 *v* 26.8 years; fathers 30.7 *v* 29.8 years) and had more years of education (mothers 10.9 *v* 10.4 years; fathers 11.9 *v* 11.4 years).

After adjustment for sex, gestational age, birth weight, and birth order and the age, education, and working status of the parents at the time of the birth, logit estimates indicated that twins had a 2.2% lower probability of college attendance than singletons. Twins had significantly lower mean difference in test scores in Chinese, mathematics, and natural science than singletons, although there was little difference in the mean scores for English and social science (table). The overall test score was significantly lower for twins than for singletons.

When we carried out separate regression estimations for twins and singletons, we found that low birth weight was an important predictor of long term academic performance. Low birthweight twins had an 8.5% lower probability of college attendance than their normal birth weight counterparts, while low birthweight singletons had only a 3.2% lower probability of college attendance.

Within the sample of twins, those with low birth weight had significantly lower test scores in English and mathematics than twins with normal birth weight; the estimated coefficients indicating that low birth weight was responsible for a substantial reduction, of 0.48, in the scores for these two subjects. There was little

Descriptive statistics and adjusted mean differences (95% CI) in test scores between twins and singletons

	Singletons		Twins		
	Mean (SD)	Median (IQR)	Mean (SD)	Median (IQR)	Estimated coefficient* (95% CI)
Chinese	10.9 (2.0)	11 (10-12)	10.8 (2.0)	11 (10-12)	−0.108 (−0.202 to −0.014)
English	9.1 (3.2)	9 (6-12)	9.1 (3.3)	9 (6-12)	0.002 (−0.149 to 0.154)
Mathematics	6.9 (3.4)	7 (4-9)	6.7 (3.4)	6 (4-9)	−0.168 (−0.326 to −0.010)
Natural science	10.0 (2.2)	10 (8-12)	9.8 (2.2)	10 (8-11)	−0.174 (−0.277 to −0.070)
Social science	12.2 (1.7)	12 (11-14)	12.1 (1.8)	12 (11-13)	−0.071 (−0.156 to −0.013)
Overall test score	49.1 (9.7)	49 (42-56)	48.5 (9.8)	49 (41-56)	−0.528 (−0.928 to −0.075)

IQR=interquartile range.
*Adjusted for birth weight, gestation age, birth order, and sex and parents' age, education, and work status at time of birth.

This article is an abridged version of a paper that was published on bmj.com. Cite this article as: *BMJ* 2008;337:a438

WHAT IS ALREADY KNOWN ON THIS TOPIC

Some early studies reported that twins had lower IQs than singletons during childhood, while more recent evidence indicates that academic performance during adolescence is similar for twins and singletons

Birth weight had a positive association with IQ in older cohorts, whereas little effect has been reported on academic performance in more recent cohorts

WHAT THIS STUDY ADDS

In Taiwan, twins have lower academic achievement scores than singletons and are less likely to go to college

Low birth weight has a negative effect on long term academic performance for both twins and singletons, though the effect is stronger for twins

Comparisons of effects of birth weight across same sex and opposite sex twins suggest that birth weight might partly reflect other underlying genetic variations

association between low birth weight and the scores for Chinese, social science, and natural science.

Within the sample of singletons, there was a significant and negative association between low birth weight and the test scores for all subjects. The effects of low birth weight on the singletons tended to be smaller than those for the twins.

To identify whether there is any association between birth weight and levels of academic performance within twin pairs we used the twin fixed effect approach to regress the differences within twins in birth weight on the differences within twins in the test scores. This enabled us to account for unobserved heterogeneity across families.

The twin pair analysis included 377 pairs of twins, of which 316 were same sex pairs. Given that we were unable to distinguish between dizygotic twins and monozygotic twins, by separating the sample into opposite sex twins and same sex twins we were able to investigate whether the effect of birth weight might partly reflect genetic differences between the twins.

For the whole sample of twin pairs, twins with higher birth weight performed better in mathematics: with an increase of 100 g in birth weight, there was a corresponding increase of 0.09 in the score for mathematics. Nevertheless, after reclassifying the twin sample into same sex and opposite sex twin pairs, we found that although birth weight in opposite sex pairs did have a significant effect on the scores for English and mathematics, the same effect was not discernible for same sex pairs.

DISCUSSION

In this large study in Taiwan, we merged the college joint entrance examinations files and birth record files and found that twins had lower academic achievement than singletons. Even after we controlled for potential confounding factors, the test scores in Chinese, mathematics, and natural science and the likelihood of college attendance were all significantly lower for twins than for singletons.

Our results agree with several previous studies in which the performance of young twins was found to be inferior to that of singletons,[1 2 6 7] and are in line with the findings of another study from the Netherlands.[8] The results differ from the Danish study in which both twins and singletons showed similar academic performance during adolescence.[5]

Our results indicate that the effect of low birth weight on academic performance persists into early adulthood. Low birth weight might be responsible for some impairment in brain development, and could result in lower intellectual performance. Given that the incidence of low birth weight is substantially higher among twins and that the negative effect of low birth weight on academic performance is greater for twins than for singletons, the low fetal growth might well result in a long term educational disadvantage for twins.[9 10]

We cannot exclude the possibility that the positive effect of birth weight, which was discernible in our opposite sex twin pairs, might be related to sex effects—for example, boys usually have higher birth weight than girls, and males quite often have higher scores than females in mathematics.

Strengths and limitations

The large sample size and high quality of our national datasets permit powerful comparisons between the academic performance of twins and singletons. Our sample of births from the 1980s, also provides initial evidence from a more recent cohort in an Asian country. Finally, we used both a multivariate regression approach to the whole cohort analysis, and a twin fixed effect approach to our twin pair analysis.

The limitations of our data are the possibility of selection bias as those who did not take the college joint entrance examinations would be excluded from our analysis; the absence of information on several potential confounding factors, such as admission to hospital or morbidity of a child during its early neonatal stage, whether the child was born after assisted conception, the quality of care during early childhood, the size of the family, and the peer group effect on academic achievement; and our inability to distinguish dizygotic twins from monozygotic twins in the same sex twin pairs. Also, although 2172 subjects had missing data on parental characteristics the descriptive statistics of the birth characteristics for twins and singletons among these subjects were generally similar to those for our whole cohort analysis. Thus, we think this is unlikely to cause serious bias to our basic estimates.

Our comparisons within twins suggest that the association between birth weight and academic performance might be partly explained by genetic factors, which leads us to one important caveat, that the estimates of birth weight based on the entire population might be biased upward as a result of genetic variations.

Contributors: See bmj.com.
Funding: National Health Research Institutes, Taiwan (NHRI-EX94-9226, NHRI-EX96-9622).
Competing interests: None declared.
Ethical approval: Ministry of Interior Affairs and Ministry of Education in Taiwan.

Provenance and peer review: Not commissioned; externally peer reviewed.

1 Record RG, McKeown T, Edwards JH. An investigation of the difference in measured intelligence between twins and single births. *Ann Hum Genet* 1970;34:11-20.
2 Allin AB, Fischbein S. Twins: are they at risk? A longitudinal study of twins and nontwins from birth to 18 years of age. *Acta Genet Med Gemellol (Roma)* 1991;40:29-40.
3 Posthuma D, De Geus EJC, Bleichrodt N, Boomsma DI. Twin-singleton differences in intelligence? *Twin Res* 2000;3:83-7.
4 Ronalds GA, De Stavola BL, Leon DA. The cognitive cost of being a twin: evidence from comparisons within families in the Aberdeen children of the 1950s cohort study. *BMJ* 2005;331:1306.
5 Christensen K, Petersen I, Skytthe A, Herskind AM, McGue M, Bingley P. Comparison of academic performance of twins and singletons in adolescence: follow-up study. *BMJ* 2006;333:1095.
6 Myrianthopoulos NC, Nichols PL, Broman SH. Intellectual development of twins—comparison with singletons. *Acta Genet Med Gemellol (Roma)* 1976;25:376-80.
7 Hay DA, O'Brien PJ, Johnston CJ, Prior M. The high incidence of reading disability in twin boys and its implications for genetic analyses. *Acta Genet Med Gemellol (Roma)* 1984;33:223-36.
8 Boomsma DI, van Beijsterveldt CE, Rietveld MJ, Bartels M, van Baal GC. Genetics mediate relation of birth weight to childhood IQ. *BMJ* 2001;323:1426-7.
9 Bouchard TJ Jr. Genetic and environmental influences on adult intelligence and special mental abilities. *Hum Biol* 1998;70:257-79.
10 Bergvall N, Iliadou A, Tuvemo T, Cnattingius S. Birth characteristics and risk of low intellectual performance in early adulthood: are the associations confounded by socioeconomic factors in adolescence or familial effects? *Pediatrics* 2006;117:714-21.

Accepted: 20 May 2008

Gallbladder disease and use of transdermal versus oral hormone replacement therapy in postmenopausal women: prospective cohort study

Bette Liu, Valerie Beral, Angela Balkwill, Jane Green, Siân Sweetland, Gillian Reeves, for the Million Women Study Collaborators

Epidemiology Unit, University of Oxford, Oxford OX3 7LF
Correspondence to: B Liu
Bette.Liu@ceu.ox.ac.uk

Cite this as: *BMJ* 2008;337:a386
doi:10.1136/bmj.a386

ABSTRACT

Objective To determine whether transdermal compared with oral use of hormone replacement therapy reduces the risk of gallbladder disease in postmenopausal women.
Design Prospective cohort study (Million Women Study).
Setting Women registered with the National Health Service (NHS) in England and Scotland.
Participants 1 001 391 postmenopausal women (mean age 56) recruited between 1996 and 2001 from NHS breast screening centres and followed by record linkage to routinely collected NHS hospital admission data for gallbladder disease.
Main outcome measures Adjusted relative risk and standardised incidence rates of hospital admission for gallbladder disease or cholecystectomy according to use of hormone replacement therapy.
Results During follow-up 19 889 women were admitted for gallbladder disease; 17 190 (86%) had a cholecystectomy. Compared with never users of hormone replacement therapy, current users were more likely to be admitted for gallbladder disease (relative risk 1.64, 95% confidence interval 1.58 to 1.69) but risks were substantially lower with transdermal therapy than with oral therapy (relative risk 1.17, 1.10 to 1.24 *v* 1.74, 1.68 to 1.80; heterogeneity P<0.001). Among women using oral therapy, equine oestrogens were associated with a slightly greater risk of gallbladder disease than estradiol (relative risk 1.79, 1.72 to 1.87 *v* 1.62, 1.54 to 1.70; heterogeneity P<0.001) and higher doses of oestrogen increased the risk more than lower doses: equine oestrogens >0.625 mg, 1.91 (1.78 to 2.04) *v* ≤0.625 mg, 1.76 (1.68 to 1.84); heterogeneity P=0.02; estradiol >1 mg, 1.68 (1.59 to 1.77) *v* ≤1 mg, 1.44 (1.31 to 1.59); heterogeneity P=0.003. The risk of gallbladder disease decreased with time since stopping therapy (trend P=0.004). Results were similar taking cholecystectomy as the outcome. Standardised hospital admission rates per 100 women over five years for cholecystectomy were 1.1 in never users, 1.3 with transdermal therapy, and 2.0 with oral therapy.
Conclusion Gallbladder disease is common in postmenopausal women and use of hormone replacement therapy increases the risk. Use of transdermal therapy rather than oral therapy over a five year period could avoid one cholecystectomy in every 140 users.

INTRODUCTION

Studies have shown an increased risk of gallbladder disease in postmenopausal women using hormone replacement therapy.[1-4] Oestrogen administered orally is metabolised by the liver before entering the systemic circulation ("first pass metabolism"). Oestrogen administered transdermally avoids this and it has been suggested that it might have a lesser effect on the risk of gallbladder disease than oral oestrogen.[1] We examined the relation between use of different types of hormone replacement therapy and the incidence of gallbladder disease in postmenopausal women.

METHODS

The Million Women Study is a population based prospective study of 1.3 million women, mean age 56, recruited from National Health Service (NHS) breast screening clinics in England and Scotland during 1996-2001. Information was obtained on hormone replacement therapy use, personal factors, and medical and reproductive history at recruitment and on resurvey about three years after recruitment.

All study participants were registered with the NHS at recruitment and were followed by record linkage for

This article is an abridged version of a paper that was published on bmj.com. Cite this article as: *BMJ* 2008;337:a386

deaths, cancer, emigration, and NHS hospital admissions.[5-7] The hospital data include the primary reason for admission and up to 12 procedures in each record.[8 9] For these analyses gallbladder disease is defined as the first hospital admission after recruitment with a diagnosis of cholelithiasis or cholecystitis, or a cholecystectomy. Analyses were also done using admission for cholecystectomy alone.

Statistical analysis

We excluded from analyses women who were premenopausal or perimenopausal, had a history of cancer, had an admission for gallbladder disease before recruitment, or whose use of hormone replacement therapy was unknown. Person years were calculated from the date of recruitment to the date of first admission for gallbladder disease, death, emigration, or end of follow-up, whichever came first.

We used Cox regression to estimate the relative risk of admission for gallbladder disease in relation to use of hormone replacement therapy. Analyses were adjusted for recruitment area, socioeconomic status, body mass index, and parity, and stratified for age and hysterectomy. We also examined the effect of adjusting for other potential confounders including smoking, alcohol use, physical activity, diet, oral contraceptive use, treatment for hypertension, high cholesterol levels, and previous stroke, heart disease, thrombosis, or bilateral oophorectomy. Sensitivity analyses were done by censoring follow-up at incident cardiovascular and breast cancer events as well at July 2002 (to coincide with publication of the first results from the women's health initiative trials[10]).

Participants were categorised according to their reported use of hormone replacement therapy at recruitment. Women were classified as users of oral therapy if they used any tablet oestrogen formulation and users of transdermal therapy if they used a patch or gel oestrogen formulation. Women using oestrogen implants were categorised separately. As hormone use may have changed during follow-up, we also examined the likely effect of this by using resurvey information collected from women on average 2.8 years after recruitment.

We calculated standardised incidence rates for admissions for cholecystectomy in users of hormone replacement therapy, taking incidence rates in never users as the standard and adjusting for age, hysterectomy, region, socioeconomic group, body mass index, and parity. Relative risks are reported with 95% confidence intervals. For further details of methods see bmj.com.

RESULTS

Overall, 1 001 391 women were included in the analyses; at recruitment 32% were current users and 18% were past users of hormone replacement therapy. Among the current users 77% used oral therapy and 18% transdermal therapy. Current users of oral and transdermal therapies were similar on all baseline characteristics except that hysterectomy and bilateral oophorectomy were more common in women using transdermal therapy (see bmj.com). From the resurvey data few women who were never or past users became current users subsequently (1% and 3% per year, respectively). Similar proportions of current users of oral and transdermal therapies at recruitment stopped use subsequently (10% and 9% per year, respectively). Among the continuing current users, 1% a year who used oral therapy changed to transdermal therapy, whereas 4% a year who used transdermal therapy changed to oral therapy.

Participants were followed for 6 102 811 person years. During follow-up 19 889 women had a first admission for gallbladder disease (mean 3.3 years after recruitment); 17 190 (86%) had a cholecystectomy.

Compared with never users of hormone therapy, the risk of gallbladder disease was significantly greater in current users (relative risk 1.64, 95% confidence interval 1.58 to 1.69) than in past users (1.27, 1.22 to 1.32). The risk among past users declined with increasing time since last use (trend P=0.004); for women who stopped hormone use more than 10 years previously, however, the risk remained significantly greater than in never users (1.19, 1.10 to 1.29). Results were similar for cholecystectomy as the outcome (see bmj.com).

Current users had higher risks of gallbladder disease than never users but the risks depended on the method of administration. Among current users transdermal therapy conferred a substantially lower risk of gallbladder disease than oral therapy (relative risk 1.17, 1.10 to 1.24 *v* 1.74, 1.68 to 1.80; heterogeneity P<0.001; fig 1). The lower risk with transdermal therapy was also seen for cholecystectomy alone: relative risk 1.18 (1.10 to 1.27) and 1.80 (1.74 to 1.87). The relative risks did not change substantially after adjustment for potential confounders (see bmj.com).

Among current users of oral hormone replacement therapy the risk of gallbladder disease varied by oestrogen type and dose (fig 2. Women using equine oestrogens had a significantly higher risk of gallbladder disease than those using estradiol: relative risk compared with never users 1.79 (1.72 to 1.87) *v* 1.62 (1.54 to 1.70); heterogeneity P<0.001. Higher doses

Method of administration	Duration of use (years)*	Cases/ population	Relative risk (95% CI)†	Relative risk (95% CI)†
All current users	6.7	8600/341 354	1.64 (1.58 to 1.69)	
Oral	6.6	6914/263 871	1.74 (1.68 to 1.80)	
Transdermal	7.2	1249/60 247	1.17 (1.10 to 1.24)	
Implanted	7.7	160/5892	1.47 (1.25 to 1.72)	
Other and unknown	6.4	277/11 344	1.55 (1.39 to 1.73)	

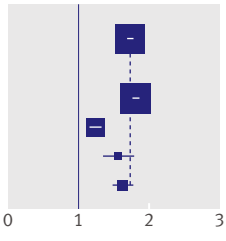


Fig 1 | Relative risk of hospital admission for gallbladder disease in current users of hormone replacement therapy compared with never users by method of administration. Dotted line represents overall relative risk for all current users compared with never users. *Mean years of use estimated for cases at time of admission for gallbladder disease, using resurvey information. †Relative risk compared with never users stratified by age and hysterectomy and adjusted for region, socioeconomic group, body mass index, and parity

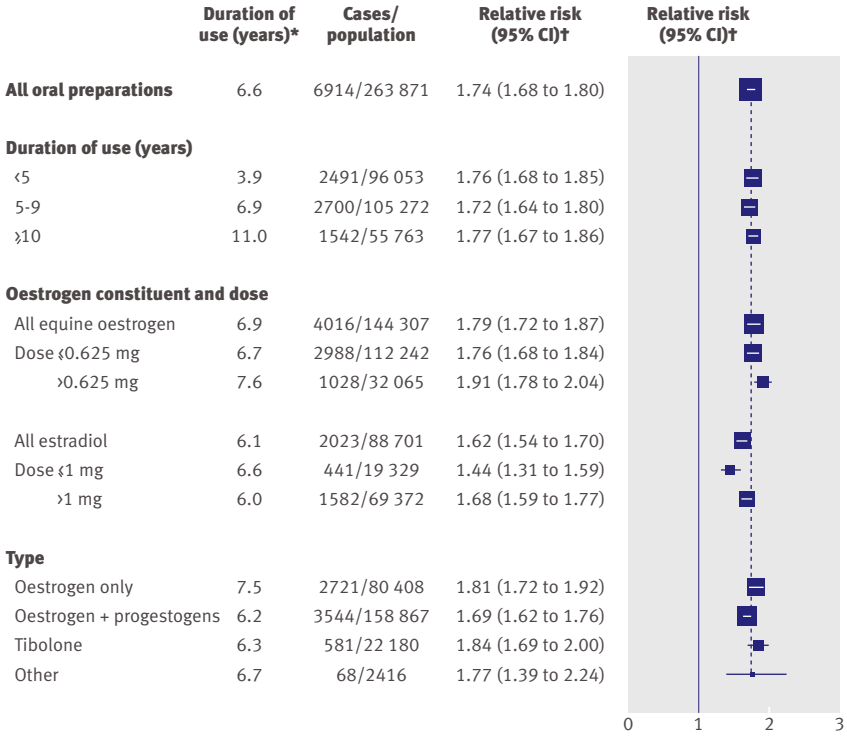


Fig 2 | Relative risk of hospital admission for gallbladder disease in current users of oral hormone replacement therapy compared with never users by duration of use and type of therapy reported at recruitment. Dotted line represents overall relative risk for all users of oral therapy compared with never users. *Mean years of use estimated for cases at time of admission for gallbladder disease, using resurvey information. †Relative risk compared with never users stratified by age and hysterectomy and adjusted for region, socioeconomic group, body mass index, and parity. Numbers in subcategories do not necessarily sum to totals owing to missing values

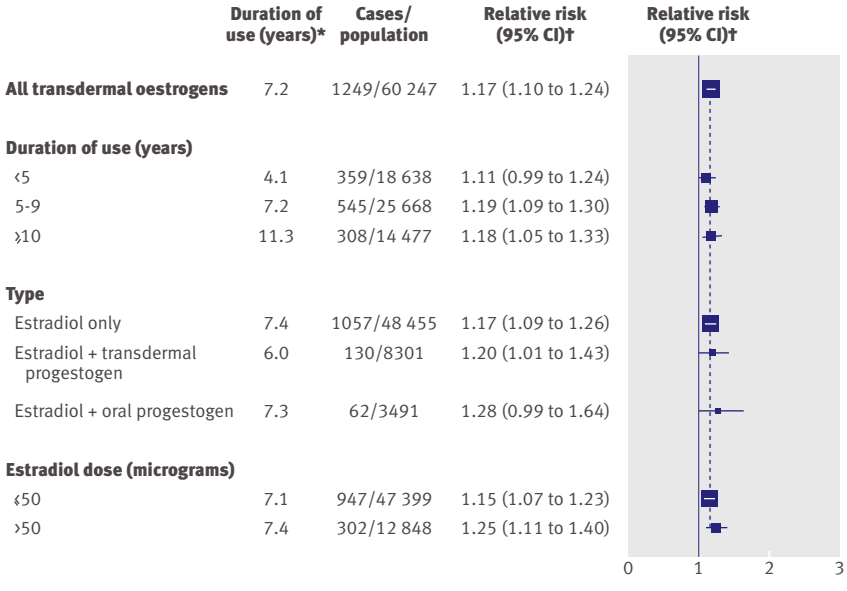


Fig 3 | Relative risk of hospital admission for gallbladder disease in current users of transdermal hormone replacement therapy compared with never users by duration of use and type of therapy reported at recruitment. Dotted line represents overall relative risk for all users of transdermal therapy compared with never users. *Mean years of use estimated for cases at time of admission for gallbladder disease, using resurvey information. †Relative risk compared with never users stratified by age and hysterectomy and adjusted for region, socioeconomic group, body mass index, and parity. Numbers in subcategories do not necessarily sum to totals owing to missing values

were also associated with greater risks than lower doses (equine oestrogen 1.76 (1.68 to 1.84) for ≤0.625 mg and 1.91 (1.78 to 2.04) for >0.625 mg; heterogeneity P=0.02; estradiol 1.44 (1.31 to 1.59) for ≤1 mg and 1.68 (1.59 to 1.77) for >1 mg; heterogeneity P=0.003). Among users of transdermal estradiol there were no significant differences found by duration of its use, whether it was used in combination with progestogens or whether the progestogen was given orally or transdermally (fig 3).

In the sensitivity analyses the relation between never users, current users, and past users did not change appreciably. Transdermal therapy consistently conferred a lower risk of gallbladder disease than oral therapy (see bmj.com).

The standardised incidence rate of cholecystectomy per 100 women over five years was 1.1 (95% confidence interval 1.1 to 1.1) in never users of hormone replacement therapy. The corresponding rate in current users of transdermal therapy was 1.3 (1.2 to 1.4) and in current users of oral therapy was 2.0 (1.9 to 2.0). Over five years the absolute difference in the risk of cholecystectomy between current users of oral and transdermal therapy was 0.7 per 100 women. Hence among women using transdermal therapy compared with oral therapy over five years, one fewer cholecystectomy would be expected in every 140 women.

DISCUSSION

In agreement with results from other studies,[1-4] we found that use of hormone replacement therapy by postmenopausal women increases the risk of gallbladder disease. We also found a substantially lower risk of gallbladder disease with use of transdermal therapy than with oral therapy. Oestrogens administered orally are metabolised in the liver before entering the systemic circulation, and the metabolites are excreted in the bile and urine.[11] Oestrogens administered transdermally avoid this first pass metabolism and this may well be contributing to the lower risk of gallbladder disease that we observed with transdermal therapy compared with oral therapy.

Our findings suggest that the method of administration, dose, and type of oestrogen all affect women's risk of developing gallbladder disease. Oestrogen implants, which can result in high blood concentrations of oestrogen[11 12] but avoid the first pass metabolism, increase the relative risk of gallbladder disease to a level somewhere between that of users of oral and transdermal therapies. Among users of oral therapy we found a dose-response effect with increased relative risks of gallbladder disease among users of higher doses compared with lower doses. Use of oral equine oestrogen was associated with a slightly higher risk of gallbladder disease than oral estradiol, possibly because equine preparations consist largely of conjugated oestrogens, which are metabolised somewhat differently from estradiol.[11]

Our results also suggest that the addition of progestogens to oestrogen therapy does not have a

WHAT IS ALREADY KNOWN ON THIS TOPIC

Hormone replacement therapy increases the risk of gallbladder disease in postmenopausal women

Oestrogens given transdermally rather than orally avoid first pass metabolism by the liver and may have a lesser effect on the risk of gallbladder disease than oral oestrogens

WHAT THIS STUDY ADDS

Use of transdermal oestrogens are associated with a substantially lower risk of gallbladder disease than use of oral oestrogens

Over five years one cholecystectomy could be avoided for every 140 postmenopausal women who use transdermal rather than oral hormone replacement therapy

large additional effect on the risk of gallbladder disease. This is in agreement with findings from the women's health initiative randomised controlled trial.[1] Furthermore, among users of transdermal oestrogens, the risks did not vary appreciably if the progestogens were administered orally or transdermally.

Strengths and limitations

Our results are based on hormone replacement therapy use reported at recruitment, and although some changes would have occurred in use during follow-up, the sensitivity analyses and data from the resurvey suggest this would not substantially alter our findings. In this cohort, the findings relating to the method of administration of hormone replacement therapy are specific for gallbladder disease, as we found no differences between transdermal and oral therapies for increased risk of breast cancer[13] and ovarian cancer,[14] or the reduced risk of fracture.[15] Prescribing of transdermal oestrogens rather than oral oestrogens differs according to some factors such as history of hysterectomy and bilateral oophorectomy (see bmj.com) but we were able to adjust for this as well as other potential confounders that could influence the choice of hormone replacement therapy.

Although some non-NHS funded admissions are not included in our analyses, we have shown that in this cohort such admissions are uncommon as most self reported cholecystectomies were included in our linked NHS hospital admission data.[16] Similarly, reporting of hormone replacement therapy use in this population compares well with prescription data.[17] In this study we had virtually complete follow-up, with objective recording of gallbladder disease. The large sample size allowed us to compare reliably a variety of types of hormone replacement therapy. The prospective design ensured that use of hormone replacement therapy was ascertained before outcomes thereby reducing biases resulting from recall or differential prescribing, and we were able to adjust for many potential confounders and found our results to be consistent.

Implications for practice

In the UK over five years an estimated 1.1% of middle aged women who have never used hormone replacement therapy are admitted to hospital for a cholecystectomy. Use of transdermal oestrogen increases the risk to 1.3% and use of oral oestrogens increases this to 2.0%. Transdermal hormone replacement therapy is generally more costly than oral and can cause local skin reactions.[11 12] However, for women who choose to use hormone replacement therapy, one cholecystectomy could be avoided for every 140 users of transdermal therapy rather than oral therapy over a five year period.

We thank the women who participated in the Million Women Study, the study steering committee, coordinating centre staff, and collaborating NHS breast screening centres. We also thank ISD Scotland, the Information Centre for Health and Social Care, and Northgate Solutions for the linkage of hospital records.

Contributors: See bmj.com.

Funding: Cancer Research UK [grant No C570/A5028], the NHS breast screening programme, and the Medical Research Council [grant No G0700474]. The funders had no involvement in the design and conduct of the study; the collection, management, analysis, and interpretation of the data; or the preparation, review, or approval of this manuscript.

Competing interests: None declared.

Ethical approval: The study has been approved by the Eastern multicentre research ethics committee.

Provenance and peer review: Not commissioned; externally peer reviewed.

1 Cirillo DJ, Wallace RB, Rodabough RJ, Greenland P, LaCroix AZ, Limacher MC, et al. Effect of estrogen therapy on gallbladder disease. *JAMA* 2005;293:330-9.
2 Simon JA, Hunninghake DB, Agarwal SK, Lin F, Cauley JA, Ireland CC, et al. Effect of estrogen plus progestin on risk for biliary tract surgery in postmenopausal women with coronary artery disease. *Ann Intern Med* 2001;135:493-501.
3 Grodstein F, Colditz GA, Stampfer MJ. Postmenopausal hormone use and cholecystectomy in a large prospective study. *Obstet Gynecol* 1994;83:5-11.
4 Mamdani MM, Tu K, van Walraven C, Austin PC, Naylor DC. Postmenopausal estrogen replacement therapy and increased rates of cholecystectomy and appendectomy. *CMAJ* 2000;162:1421-4.
5 Million Women Study Collaborators. The million women study: design and characteristics of the study population. *Breast Cancer Res* 1999;1:73-80.
6 Kendrick S, Clarke J. The Scottish record linkage system. *Health Bull (Edinb)* 1993;51:72-9.
7 Information Centre. Hospital Episode Statistics. 2007. www.hesonline.nhs.uk.
8 World Health Organization. *International statistical classification of diseases and related health problems*. 10th revision. Geneva: WHO, 1992.
9 Office of population censuses and surveys. *Classification of surgical operations and procedures*. Fourth revision. London: Crown copyright, 2000.
10 Writing group for the Women's Health Initiative Investigators. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results from the women's health initiative randomized controlled trial. *JAMA* 2002;288:321-33.
11 Sweetman S, ed. *Martindale: the complete drug reference*. 35th ed. London: Pharmaceutical Press, 2007.
12 Lumsden M. Hormone replacement therapy: regimen and routes of administration. In: Critchley H, Gebbie A, Beral V, eds. *Menopause and hormone replacement*. London: RCOG Press, 2004:187-203.
13 Million Women Study Collaborators. Breast cancer and hormone-replacement therapy in the million women study. *Lancet* 2003;362:419-27.
14 Million Women Study Collaborators. Ovarian cancer and hormone replacement therapy in the million women study. *Lancet* 2007;369:1703-10.
15 Banks E, Beral V, Reeves G, Balkwill A, Barnes I. Fracture incidence in relation to the pattern of use of hormone therapy in postmenopausal women. *JAMA* 2004;291:2212-20.
16 Liu B, Sweetland S, Beral V, Green J, Balkwill A, Casabonne D. Self-reported information on joint replacement and cholecystectomy agrees well with that in medical records. *J Clin Epidemiol* 2007;60:1190-4.
17 Banks E, Beral V, Cameron R, Hogg A, Langley N, Barnes I, et al. Agreement between general practice prescription data and self-reported use of hormone replacement therapy and treatment for various illnesses. *J Epidemiol Biostat* 2001;6:357-63.

Accepted: 9 May 2008