RESEARCH

Use of blood pressure lowering drugs in the prevention of cardiovascular disease: meta-analysis of 147 randomised trials in the context of expectations from prospective epidemiological studies

M R Law, J K Morris, N J Wald

Centre for Environmental and Preventive Medicine, Wolfson Institute of Preventive Medicine, Barts and The London School of Medicine, Queen Mary University of London, London EC1M 6BQ
Correspondence to: M R Law
m.r.law@qmul.ac.uk

ABSTRACT

Objectives To determine the quantitative efficacy of different classes of blood pressure lowering drugs in preventing coronary heart disease (CHD) and stroke, and who should receive treatment.

Design Meta-analysis.

Data source Medline (1966-2007).

Study selection Randomised trials of blood pressure lowering drugs recording CHD events and strokes. 108 trials studied differences in blood pressure between study drug and placebo (or control group not receiving the study drug) (blood pressure difference trials), and 46 trials compared drugs (drug comparison trials). Seven trials with three randomised groups fell into both categories. The results were interpreted in the context of those expected from the largest published meta-analysis of cohort studies, totalling 958 000 people.

Participants 464 000 people defined into three mutually exclusive categories: participants with no history of vascular disease, a history of CHD, or a history of stroke.

Results In the blood pressure difference trials β blockers had a special effect over and above that due to blood pressure reduction in preventing recurrent CHD events in people with a history of CHD: risk reduction 29% (95% confidence interval 22% to 34%) compared with 15% (11% to 19%) in trials of other drugs. The extra effect was limited to a few years after myocardial infarction, with a risk reduction of 31% compared with 13% in people with CHD with no recent infarct (P=0.04). In the other blood pressure difference trials (excluding CHD events in trials of β blockers in people with CHD), there was a 22% reduction in CHD events (17% to 27%) and a 41% (33% to 48%) reduction in stroke for a blood pressure reduction of 10 mm Hg systolic or 5 mm Hg diastolic, similar to the reductions of 25% (CHD) and 36% (stroke) expected for the same difference in blood pressure from the cohort study meta-analysis, indicating that the benefit is explained by blood pressure reduction in itself. The five main classes of blood pressure lowering drugs (thiazides, β blockers, angiotensin converting enzyme inhibitors, angiotensin receptor blockers, and calcium channel blockers) were similarly effective (within a few percentage points) in preventing CHD events and strokes, with the exception that calcium channel blockers had a greater preventive effect on stroke (relative risk 0.92, 95% confidence interval 0.85 to 0.98). The percentage reductions in CHD and stroke were similar in people with and without cardiovascular disease and regardless of blood pressure before treatment (down to 110 mm Hg systolic and 70 mm Hg diastolic). Combining our results with those from two other studies (the meta-analyses of blood pressure cohort studies and of trials determining the blood pressure lowering effects of drugs according to dose) showed that in people aged 60-69 with a diastolic blood pressure before treatment of 90 mm Hg, three drugs at half standard dose in combination reduced the risk of CHD by an estimated 46% and of stroke by 62%; one drug at standard dose had about half this effect. The present meta-analysis also showed that drugs other than calcium channel blockers (with the exception of non-cardioselective β blockers) reduced the incidence of heart failure, by 24% (19% to 28%) and calcium channel blockers by 19% (6% to 31%).

Conclusions With the exception of the extra protective effect of β blockers given shortly after a myocardial infarction and the minor additional effect of calcium channel blockers in preventing stroke, all the classes of blood pressure lowering drugs have a similar effect in reducing CHD events and stroke for a given reduction in blood pressure so excluding material pleiotropic effects. The proportional reduction in cardiovascular disease events was the same or similar regardless of pretreatment blood pressure and the presence or absence of existing cardiovascular disease. Guidelines on the use of blood pressure lowering drugs can be simplified so that drugs are offered to people with all levels of blood pressure. Our results indicate the importance of lowering blood pressure in everyone over a certain age, rather than measuring it in everyone and treating it in some.

INTRODUCTION

Despite the widespread use of blood pressure lowering drugs and the results of many randomised trials,[1-20 w1-w162] questions remain about which drugs to use and who to treat. Firstly, do β blockers have a special effect over lowering blood pressure in preventing coronary heart

disease (CHD) events in people with a history of CHD? Secondly, does the effect of blood pressure lowering drugs in preventing CHD and stroke differ in people with and without a history of cardiovascular disease? Thirdly, does blood pressure reduction alone explain the effect of blood pressure lowering drugs in preventing CHD and stroke? There are claims of additional non-blood pressure lowering (so called pleiotropic) effects of drugs.[7 8 13 w135 w136 w139] Selected trial data have been used to suggest that each of the five main classes of blood pressure lowering drugs (thiazides, β blockers, angiotensin converting enzyme inhibitors, angiotensin receptor blockers, and calcium channel blockers) has a greater preventive effect,[1-13 w126 w129] and each a lesser preventive effect,[9-20 w126 w135] than other drugs. Fourthly, should the use of blood pressure lowering drugs be limited to people with "high" blood pressure and not given to those at high risk of cardiovascular disease who have a lower blood pressure? A corollary is whether blood pressure should be reduced to a limited extent only—a treat to target approach.[9-11 21 22 23 24] Finally, what is the quantitative effect of taking one or more blood pressure lowering drugs in lowering blood pressure and preventing CHD events and stroke according to dose, pretreatment blood pressure, and age? We answered these questions using the results from 147 randomised trials of blood pressure lowering drugs and CHD events (n=22 000) and stroke (n=12 000), examined in the context of the results from the largest meta-analysis of epidemiological cohort studies of blood pressure and CHD and stroke.[25]

METHODS

The database search (by MRL) used Medline (1966 to December 2007) to identify randomised trials of blood pressure lowering drugs in which CHD events or strokes were recorded. We also searched the Cochrane Collaboration and Web of Science databases and the citations in trials and meta-analysis and review articles.

We recorded the numbers of participants having one or more CHD events (fatal or non-fatal myocardial infarction or sudden cardiac death) and one or more

strokes (haemorrhagic and ischaemic). We also recorded the numbers of participants with a new diagnosis of heart failure or an exacerbation of an existing heart failure based on new hospital admissions or death from the disorder. Outcomes were recorded regardless of whether participants took their allocated tablets. Change in blood pressure (value on entry minus average value during trial in treated group, minus same change in control group) was recorded on an intention to treat basis by determining the numbers of participants in the treated and control groups who stopped attending clinics and taking the difference in blood pressure between them to be zero after they left the trial.

Categories of trial

The trials were divided into three categories according to whether the recruitment of participant was based on having no history of cardiovascular disease, a history of CHD (acute myocardial infarction, coronary artery disease without recent infarction, or heart failure), or a history of stroke (or other cerebrovascular disease). We also categorised the trials into "blood pressure difference trials" and "drug comparison trials." The blood pressure difference trials were those designed to achieve a difference in blood pressure between randomised groups who were given and not given the study drugs to show the effect of this difference on the incidence of CHD events and stroke: 92 of the 108 such trials were placebo controlled. Additional blood pressure lowering drugs were commonly used in the different groups in each trial. Trials were regarded as single drug trials if the difference between the groups in the mean number of drugs prescribed per participant was less than 1.5, and combination drug trials if 1.5 or greater.

The drug comparison trials were those that compared two blood pressure lowering drugs with each other. Although additional drugs could be used there was no intention to achieve a blood pressure reduction in one group compared with another. These trials
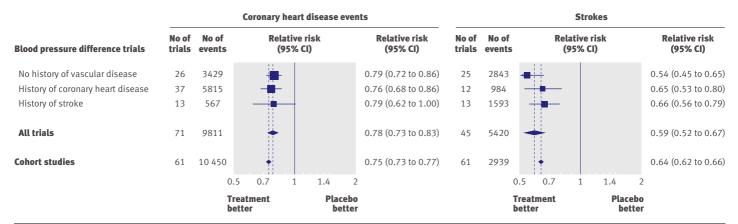

	Coronary heart disease events				**Strokes**			
Blood pressure difference trials	No of trials	No of events	Relative risk (95% CI)	Relative risk (95% CI)	No of trials	No of events	Relative risk (95% CI)	Relative risk (95% CI)
No history of vascular disease	26	3429		0.79 (0.72 to 0.86)	25	2843		0.54 (0.45 to 0.65)
History of coronary heart disease	37	5815		0.76 (0.68 to 0.86)	12	984		0.65 (0.53 to 0.80)
History of stroke	13	567		0.79 (0.62 to 1.00)	13	1593		0.66 (0.56 to 0.79)
All trials	71	9811		0.78 (0.73 to 0.83)	45	5420		0.59 (0.52 to 0.67)
Cohort studies	61	10 450		0.75 (0.73 to 0.77)	61	2939		0.64 (0.62 to 0.66)

Fig 1 | Relative risk estimates of coronary heart disease events and stroke for a blood pressure reduction of 10 mm Hg systolic or 5 mm Hg diastolic in the blood pressure difference trials and in epidemiological cohort studies. (Total number of trials is fewer than the sum of three categories as five included participants with and without vascular disease)

therefore tested for effects of a drug that were unrelated to lowering blood pressure.

Statistical analysis

All statistical analyses were done using Stata software. We combined relative risk estimates of disease events from individual trials using a random effects model.[26] Summary relative risk estimates from blood pressure difference trials were standardised to a blood pressure reduction of 10 mm Hg systolic or 5 mm Hg diastolic, by raising the relative risk estimate in each trial to the appropriate power (10 divided by the observed reduction in systolic blood pressure or 5 divided by the observed reduction in diastolic pressure). If reductions in systolic and diastolic pressure were reported, we took the average of the two risk estimates. As the reduction in blood pressure was not reported in most trials in people with a history of CHD, we estimated the average reduction from the average pretreatment blood pressure and the average drug dose, using results from a meta-analysis.[27] The estimated blood pressure reduction was 5.9 mm Hg systolic and 3.1 mm Hg diastolic, close to the median reduction in the 27 trials in which blood pressure reduction was reported (6 mm Hg and 3 mm Hg, respectively).

Predicting the trial results on CHD and stroke from epidemiological studies and trials of drugs on blood pressure

Effect of blood pressure lowering drugs in lowering blood pressure according to dose

These estimates were taken from a meta-analysis of 354 short term randomised placebo controlled trials of blood pressure lowering drugs in fixed dose.[27] This showed that the five main classes of blood pressure lowering drugs all produced similar reductions in blood pressure when taken at standard dose or at the same multiple of standard dose, and that the effect of the drugs in lowering blood pressure increased with dose (by about 2 mm Hg systolic and 1 mm Hg diastolic for a doubling in dose) and with pretreatment blood pressure.

Expected reduction in disease events for a specified reduction in blood pressure

The associations between systolic and diastolic blood pressure and CHD events and stroke were taken from a meta-analysis of 61 cohort studies.[25] This showed that in every age group cardiovascular mortality plotted on a logarithmic scale against blood pressure on an arithmetic scale is well fitted by straight lines, indicating a constant proportional change in risk for a specified change in blood pressure from any level of pretreatment blood pressure. For a specified age specific regression slope S (indicating the relative risk for a blood pressure decrease of 20 mm Hg systolic or 10 mm Hg diastolic), and decrease in blood pressure, d, the relative risk is $S^{d/20}$ for systolic pressure and $S^{d/10}$ for diastolic pressure (see bmj.com).

The data were used to produce equations that predict blood pressure reductions given number of drugs, dose of drugs (as a multiple of standard), pretreatment blood pressure, and age (see bmj.com).

RESULTS

Overall, 147 trial reports were included (see bmj.com): 108 blood pressure difference trials and 46 drug comparisons trials (seven reports with two treatment groups and a placebo group fell into both categories, treatment versus placebo and one treatment versus the other). Forest plots of individual trial results and the summary relative risk estimates and results for heterogeneity testing are on bmj.com.

Do β blockers have a special effect in preventing CHD events in people with a history of CHD?

In the 37 blood pressure difference trials of β blockers in people with a history of CHD, that compared β blockers with placebo (32 trials) or with an untreated control group (5 trials), CHD events were, on average, reduced by 29% (relative risk 0.71, 95% confidence interval 0.66 to 0.78), significantly greater (P<0.001) than the 15% reduction in single drug trials of β blockers in people without a history of CHD and of other classes of drug in people with and without a history of CHD. The greater protective effect of β blockers in people with CHD was explained by a greater effect in the 27 trials that recruited participants at the time of an acute myocardial infarction. The risk reduction for recurrent CHD events over the 1-2 year follow-up in these 27 trials was 31% (relative risk 0.69, 0.62 to 0.76). In the 11 trials remaining (one recruited some participants with a recent infarct and some without[w62]) participants had a history of CHD but no recent infarct; in these the risk reduction was 13% (relative risk 0.87, 0.71 to 1.06; P=0.04 for the difference between the two groups of trials), similar to the 15% risk reductions in the other single drug trials. The 31% risk reduction after acute myocardial infarction was significantly greater (P<0.001). β blockers used for one or two years after an acute myocardial infarction were therefore about twice as effective as β blockers used in other circumstances and about twice as effective as other drugs used in any circumstances.

The four drug comparison trials of β blockers compared with other drugs in people with CHD but no recent infarct confirmed the absence of a special effect of β blockers in the absence of a recent infarct; the summary relative risk of CHD events was 0.99 (0.82 to 1.20), a relative risk of 1.0 indicating the same risk reduction from β blockers and other drugs.

In view of the special effect of β blockers, CHD events in all 37 blood pressure difference trials and all four drug comparison trials of β blockers in people with CHD were excluded from subsequent analyses according to the prior stipulation that we would do so if a special effect was observed, even though post hoc the special effect was limited to a subset (those with acute infarction).

Does the preventive effect of drugs differ in people with and without a history of cardiovascular disease?

The summary relative risk estimates of CHD events and stroke in the blood pressure difference trials, observed and standardised for reduction in blood pressure, were similar in the three categories of trials (no vascular disease, history of CHD, and history of stroke), showing no difference in effect in people with or without vascular disease (see bmj.com). There was no heterogeneity.

Does blood pressure reduction alone explain the preventive effect of the drugs?

Blood pressure difference trials

Figure1 shows the relative risk estimates of CHD events and stroke in the blood pressure difference trials, standardised to a blood pressure reduction of 10 mm Hg systolic or 5 mm Hg diastolic, together with the corresponding relative risk estimates derived from the meta-analysis of cohort studies [25] in people aged 60-69 years, the average age at the time of a cardiovascular event in the trials. The estimates from the trials meta-analysis were a 22% (95% confidence interval 17% to 27%) reduction in CHD events (relative risk 0.78) and a 41% (33% to 48%) reduction in stroke (relative risk 0.59), similar to those from the cohort study meta-analysis, a 25% decrease in CHD events (relative risk 0.75) and a 36%

decrease in stroke (relative risk 0.64) for the same difference in blood pressure.

After only one year of follow-up the reduction in CHD events was 20% (9% to 29%) and the reduction in stroke was 32% (18% to 44%) for a blood pressure reduction of 10 mm Hg in systolic or 5 mm Hg diastolic, similar to the long term trial results (22% and 41%) and similar to the results expected from the cohort studies (25% and 36%) (see fig 1), indicating that the full potential effect of blood pressure reduction is achieved within a year.

In the single drug trials comparing a specified drug with placebo (or with a control group not receiving the study drug), reductions in CHD events and stroke were similar in magnitude for each of the five main classes of drug (see bmj.com). All the disease reductions were statistically significant but for angiotensin receptor blockers there were only four trials and hence insufficient statistical power to show an effect. No statistically significant heterogeneity for CHD events was observed across trials of the five drug classes (χ^2=2.0, df=5, P=0.86), but the reduction in incidence of stroke was smaller in trials of β blockers (17%) than in single drug trials of the other four classes of drug combined (29%; P=0.03).

Drug comparison trials

The summary relative risk estimates for CHD in the drug comparison trials comparing each of the five

Pretreatment diastolic blood pressure (mm Hg)	Coronary heart disease events				Strokes			
	No of trials	No of events	Relative risk (95% CI)	Relative risk (95% CI)	No of trials	No of events	Relative risk (95% CI)	Relative risk (95% CI)
70-74	5	663		0.79 (0.65 to 0.88)	2	284		0.64 (0.50 to 0.80)
75-79	21	3708		0.85 (0.76 to 0.94)	11	1394		0.76 (0.62 to 0.92)
80-84	8	1517		0.86 (0.73 to 1.01)	6	909		0.76 (0.66 to 0.88)
85-89	12	1462		0.84 (0.76 to 0.93)	10	1458		0.78 (0.66 to 0.92)
90-94	6	1358		0.88 (0.79 to 0.97)	7	1030		0.63 (0.56 to 0.72)
›95	9	255		0.74 (0.58 to 0.94)	9	332		0.54 (0.42 to 0.69)
Not reported	12	848		0.85 (0.75 to 0.97)	2	13		0.63 (0.21 to 1.92)
All trials	71	9811		0.84 (0.81 to 0.88)	45	5420		0.70 (0.64 to 0.76)
Pretreatment systolic blood pressure (mm Hg)								
110-119	2	320		0.78 (0.63 to 0.96)	0	0		
120-129	10	1176		0.77 (0.66 to 0.90)	2	27		0.56 (0.26 to 1.17)
130-139	18	3463		0.89 (0.80 to 0.99)	12	1301		0.75 (0.63 to 0.89)
140-149	7	1346		0.85 (0.76 to 0.94)	8	1709		0.77 (0.62 to 0.95)
150-159	11	1295		0.86 (0.77 to 0.96)	11	1339		0.69 (0.60 to 0.80)
160-169	4	460		0.79 (0.66 to 0.95)	4	479		0.66 (0.50 to 0.87)
›170	5	268		0.86 (0.68 to 1.09)	5	261		0.58 (0.46 to 0.74)
Not reported	15	1483		0.83 (0.75 to 0.93)	4	304		0.62 (0.49 to 0.77)
All trials	71	9811		0.84 (0.81 to 0.88)	45	5420		0.70 (0.64 to 0.76)

0.5   0.7   1   1.4   2   0.5   0.7   1   1.4   2

Treatment better   Placebo better   Treatment better   Placebo better
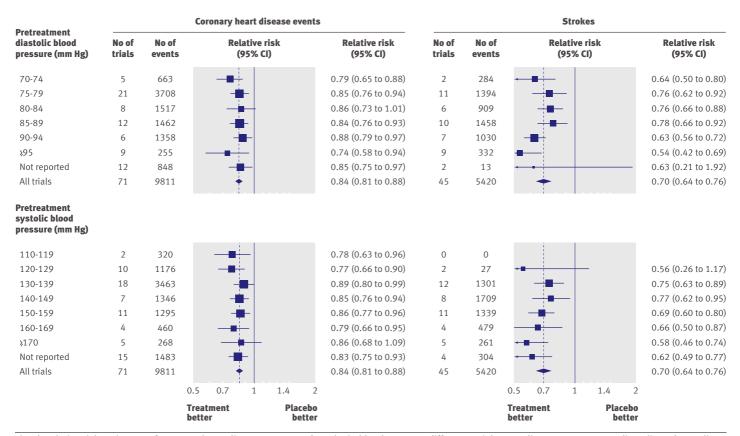
Fig 2 | Relative risk estimates of coronary heart disease events and stroke in blood pressure difference trials according to pretreatment diastolic and systolic blood pressures (taken as average in placebo group over course of trial). (Totals are less than the sum of the individual categories because some trials include more than one category)
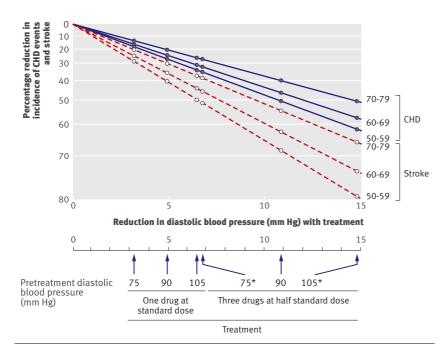
Fig 3 | Reduction in incidence of coronary heart disease events and stroke in relation to reduction in diastolic blood pressure according to drug dose, number of drugs, pretreatment diastolic blood pressure, and age. *Blood pressure reductions are more uncertain and hence also reductions in disease incidence

classes of drug with drugs from the other classes were close to 1.0, indicating no advantage of any one drug over others in the prevention of CHD. The differences between classes of drug in average blood pressure reductions were close to zero (see bmj.com), and the differences in use of add-on drugs were negligible (≤0.03 drugs per participant). The different classes of drug therefore reduced blood pressure by about the same extent and reduced CHD by about the same extent, providing evidence of a lack of preventive effect attributable to mechanisms other than lowering blood pressure.

In the drug comparison trials the overall risk reduction in CHD events with thiazides was similar to that of other classes of drug (see bmj.com). There was, however, an increased risk of sudden cardiac death from using thiazides in very high dose, concealed in the summary results because few of the thiazide trials used very high doses (four times standard) and because sudden cardiac deaths were a small proportion of all coronary heart disease events.

The summary relative risk estimates for stroke in the drug comparison trials were close to 1.0, with two exceptions, a greater preventive effect of calcium channel blockers than other drugs and a lesser effect of β blockers.[3-7 9 10 30] The greater preventive effect of calcium channel blockers than other drugs (relative risk 0.91, 95% confidence interval 0.84 to 0.98; P=0.01) was not materially altered after adjustment for the small difference in blood pressure reduction between the groups (relative risk 0.92, 0.85 to 0.98), and is equivalent to a reduction in risk of stroke of 33% rather than

27%, the overall summary estimate. The observed lesser effect of β blockers than other drugs in preventing stroke (relative risk 1.18, 1.03 to 1.36; P=0.02) is equivalent to a 19% reduction in risk of stroke rather than 27%. The observed lesser effect of β blockers, however, rested on trials comparing calcium channel blockers with β blockers.[w136-w140] Exclusion of the results from these trials weakened the evidence favouring a disadvantage of β blockers over the three other classes (relative risk 1.11, 0.86 to 1.44; P=0.40) but had little effect on the strength of evidence favouring an advantage of calcium channel blockers over the three other classes of drug (relative risk 0.93, 0.86 to 1.01; P=0.07).

Should the use of blood pressure lowering drugs be limited to people with "high" blood pressure?

The relative risk estimates of CHD events and stroke in the blood pressure difference trials were similar across all levels of pretreatment blood pressure down to 110 mm Hg systolic and 70 mm Hg diastolic, below which there were too few data (fig 2). At each blood pressure level the relative risk reductions were statistically significant and consistent with the summary relative risk estimates for all the trials: 0.84 for CHD events and 0.70 for stroke (see bmj.com). A metaregression analysis showed no significant trend in proportional disease reduction with lower pretreatment blood pressure, indicating a constant proportional effect. The trial results mirror those in cohort studies,[25 28 29] which show a proportional reduction in risk that is constant over all measured levels of blood pressure—that is, the same in people with lower and higher blood pressures.

There was no heterogeneity across the relative risk estimates for CHD events according to pretreatment diastolic blood pressure (χ^2=3.9, df=6, P=0.69; see bmj.com). There was, however, heterogeneity for stroke (χ^2=19, df=6, P=0.004), owing to a greater risk reduction in trials with the highest pretreatment blood pressure (≥95 mm Hg), which arose because of more intensive treatment in these trials. The same applied to the analysis based on systolic blood pressure (CHD, χ^2=3.7, df=7, P=0.82; stroke, χ^2=12.24, df=6, P=0.06; see bmj.com).

What is the quantitative effect of one or more blood pressure lowering drugs on lowering blood pressure and preventing CHD events and stroke?

The effect of taking blood pressure lowering drugs in reducing the incidence of CHD and stroke according to number of drugs used, dose of drugs, and age cannot be estimated accurately from the blood pressure differences trials (alone). This is because about a quarter of treated participants stopped taking their allocated drugs, individual trials used varying doses of drugs, use of combination drug therapy was limited, and the age range was relatively narrow. All this can be overcome by doing a two stage analysis (see bmj.com), in which the effect of drugs in lowering blood pressure is determined from mainly short term trials and this is used with cohort

study evidence on the effect of differences in blood pressure on risk of CHD events and stroke.

Figure 3 shows the resulting estimates. The observed reductions in CHD events and stroke in the blood pressure difference of single drug trials (mean difference between randomised groups 1.0 drug per participant) and of combination drug therapy (mean difference 2.0 drugs), were similar to the predicted values shown in figure 3 taking into account pretreatment blood pressure, drug dose, and age, after adjustment for non-adherence to allocated treatment (see bmj.com). The trial results from the present meta-analysis therefore validate the estimates in figure 3.

Figure 3 shows that one drug at standard dose reduces the incidence of CHD by about 24% and of stroke by 35% in people aged 60-69 with a diastolic blood pressure of 90 mm Hg (fig 3). Three drugs at half standard dose about doubles this effect, reducing the incidence of CHD by about 45% and of stroke by 60% (fig 3). At higher blood pressure (180/105 mm Hg) and at lower blood pressure (120/75 mm Hg) the effect of one drug at standard dose is about 7-9 percentage points greater and smaller, respectively, and of three drugs at half standard dose about 12-14 percentage points greater and smaller. The proportional effect of age is relatively small; in people 10 years older the effect of one drug at standard dose is only 3 percentage points lower on average, and of three drugs at half standard dose 5 percentage points lower. Because mortality from CHD and stroke approximately trebles with each 10 year increase in age, the absolute gain from blood pressure reduction is greater at older ages.

Heart failure

Heart failure (17 872 episodes) was recorded in 64 blood pressure difference trials and 31 drug comparison trials. Heterogeneity existed across the results of the trials of β blockers and heart failure (P=0.008), explained by the observation that β blockers without cardioselective or α blocking (vasodilatory) properties (such as propranolol) lacked a preventive effect on heart failure (relative risk 1.01, 95% confidence interval 0.76 to 1.35), but β blockers with one or other of these properties had a preventive effect (0.77, 0.69 to 0.87; P=0.01 for difference).

Calcium channel blockers reduced heart failure in the blood pressure difference trials by 19% (P=0.007), although the drug comparison trials showed that they were statistically significantly less effective in doing so than the other four classes of drugs (relative risk 1.22, 1.10 to 1.35; P<0.001). Each of the other four classes of drug significantly reduced the incidence of heart failure in the blood pressure difference trials (P<0.001) by 24% on average, with no significant differences in effect between them either in the blood pressure difference trials or the drug comparison trials (see bmj.com). The effect of calcium channel blockers in reducing heart failure in the blood pressure difference trials (19%) was therefore not much less than that of the other classes of drug (24%).

DISCUSSION

This, the largest meta-analysis of randomised trials of blood pressure reduction, shows that lowering systolic blood pressure by 10 mm Hg or diastolic blood pressure by 5 mm Hg using any of the main classes of blood pressure lowering drugs, reduces CHD events (fatal and non-fatal) by about a quarter and stroke by about a third, regardless of the presence or absence of vascular disease and of pretreatment blood pressure. Heart failure is also reduced by about a quarter.

β blockers in people with CHD

Our results confirm that there is a special protective effect of β blockers in preventing CHD events in people with a history of CHD over and above their blood pressure lowing effect. This special effect was limited to a few years after an acute myocardial infarction. The overall protective effect was about double that of β blockers in people with CHD but no recent infarct or in people without CHD and that of other drugs regardless of history of CHD. This analysis was possible because the trials in which participants were recruited immediately after an acute infarct had short durations of follow-up (one or two years). The dichotomy of the trial data on β blockers into short term trials of acute infarct and trials of non-acute CHD provided the opportunity to show that the special effect of β blockers was a short term effect, avoiding the dilution of effect that would have occurred had the acute infarct trials continued for many years.

Preventive effect in people with and without cardiovascular disease

With the exception of the special short term effect of β blockers in acute myocardial infarction, our results show that the preventive effect of all classes of blood pressure lowering drugs is the same or similar in people with and without a history of cardiovascular disease (fig 1), so there is no reason to use these drugs for secondary prevention but not for primary prevention. The preventive effect of blood pressure reduction was rapid, the full potential effect being achieved within a year.

Quantitative linking of blood pressure reduction and disease prevention

An important result from our analysis is that results from the meta-analysis of trials of drugs on blood pressure reduction linked to the cohort studies meta-analysis (differences in risk of CHD events and stroke for specified differences in blood pressure) accurately predict the results of the present meta-analysis indicating that blood pressure reduction in itself explains the preventive effect of the drugs. With the possible minor additional effect of calcium channel blockers in preventing stroke the five classes of drugs were equally effective in lowering blood pressure and equally effective in preventing CHD events and stroke.[3-7 9 10 30] A possible explanation for the greater effect of calcium channel blockers on the risk of stroke is the observation

that, although the different classes of blood pressure lowering drugs reduce peripheral arterial pressure to a similar extent,[27] the reduction in central aortic pressure appears greater with calcium channel blockers and lower with β blockers than with the other three classes of drug.[31-33] But it is not a persuasive argument because any additional reduction in central aortic pressure should also confer greater prevention of CHD than with other drugs but this was not observed. Thus with the exception of β blockers after acute myocardial infarction and the minor difference in the effect of calcium channel blockers in reducing the risk of stroke, blood pressure reduction explains the action of the drugs in preventing CHD and stroke. The results thus exclude the blood pressure lowering drugs in general having material pleiotropic effects.

While our results do not exclude possible differences in efficacy between drugs within a class this is unlikely. Any such differences are likely to be small and clinically unimportant because (β blockers and heart failure apart) for each class of drug there was no significant heterogeneity between trials of the individual drugs studied, either for blood pressure reduction[27] or for reduction in disease events. Trial results that suggest greater or lesser effects of some drugs can be explained by chance alone.

In the blood pressure difference trials the use of add-on treatment was the same on average in the treated and placebo groups (overall difference 0.3 drugs per participant). Over all the trials, 25% of participants allocated active treatment stopped taking their tablets; this non-adherence did not bias comparisons between the classes of drug because the proportions who stopped were similar for each class. The non-adherence underestimates the effect of taking the drugs on disease prevention but does not underestimate the effect of a specified blood pressure reduction from the drugs on disease prevention because the calculation of the difference in blood pressure took non-adherence into account. Thus the observations that a blood pressure reduction of 10 mm Hg systolic or 5 mm Hg diastolic, however achieved, reduced CHD

events by 22% and stroke by 41% in the trials are unbiased estimates of efficacy.

In the drug comparison trials the differences in use of additional drugs between the groups were small (≤0.3 drugs per participants in trials comparing each class of drug with any other drug). That there were no material differences in blood pressure between the groups and no material difference in the incidence of CHD or stroke permits the conclusion that the preventive effects of each class of drug are mediated through blood pressure reduction alone, corroborating the conclusion that the drugs had no pleiotropic effects based on the similarity in predicted and observed results from the drug difference trials (fig 1).

Proportional disease reduction for a given blood pressure reduction independent of pretreatment blood pressure

Our results indicate that the use of blood pressure lowering drugs should not be limited to people with high blood pressure. The proportional reduction in disease events for a given blood pressure reduction was the same irrespective of pretreatment blood pressure, down to 70 mm Hg or lower for diastolic blood pressure, as expected from the results of epidemiological cohort studies that showed a constant proportional change in risk for a specified change in blood pressure from any level of pretreatment blood pressure.[25 28 29] This result supports a "lower the better" approach to blood pressure reduction. It means that there is medical benefit in lowering a person's blood pressure whatever the blood pressure, with the logically inescapable conclusion that there is then little or no gain in measuring a person's blood pressure—a conclusion that will undoubtedly stimulate discussion since it is at variance with a 100 years of medical practice.

From drugs to blood pressure reduction to disease prevention: a quantitative summary

Figure 3, based on meta-analyses of trials of blood pressure lowering drugs and blood pressure and cohort studies of blood pressure and cardiovascular disease, permits the prediction of disease prevention given the determining factors—namely, number and dose of drugs used, pretreatment blood pressure, and age. Importantly the analysis of the randomised trials of blood pressure reduction on disease presented in this paper confirm these predictions. The advantage of figure 3 is that it provides information on the expected effects of treatment over a wider range of age and drug regimens than can be obtained from the trials themselves.

Our estimates of the proportional reduction in risk of CHD events and stroke vary according to age. In a recent meta-analysis of 31 trials,[34] using individual patient data or unpublished tabular data in prespecified categories, age had no material influence on attenuating the effect of blood pressure reduction in preventing cardiovascular disease. However, their results did show an attenuating effect of age; the risk of cardiovascular disease was reduced by 24% per 5 mm Hg reduction in systolic blood pressure for a 15 year increase in age

---

WHAT IS ALREADY KNOWN ON THIS TOPIC

The different classes of blood pressure lowering drugs at standard doses, or the same multiple of standard dose, lower blood pressure to a similar extent

Blood pressure lowering drugs reduce the risk of coronary heart disease (CHD) events and stroke in people with a history of vascular disease and in those with high blood pressure

WHAT THIS STUDY ADDS

The effect of blood pressure lowering drugs in reducing the risk of disease is entirely or largely due to blood pressure reduction, with one main exception, a special extra effect of β blockers in people who have had a recent myocardial infarction

The proportional reduction in CHD events and stroke for a given reduction in blood pressure, an approximate halving in risk for each 10 mm Hg diastolic reduction, is the same in people with and without a history of vascular disease and in people without high blood pressure as well as in those with high blood pressure

There is benefit in people lowering blood pressure in anyone at sufficient cardiovascular risk whatever their blood pressure, so avoiding the need to measure blood pressure routinely

---

(11.9% cardiovascular disease prevention reduced to 9.1%), although this was not statistically significant.[30] This estimate was close to the 20% expected decrease from the results of the cohort study meta-analysis we used.[25] The 24% estimate from the trial meta-analysis was probably real but was not statistically significant because the blood pressure reductions observed in the trial were relatively small and the reductions in cardiovascular disease were therefore also small. The important conclusion is that the cohort studies and the trial data are consistent in showing an age modifying effect on prevention of CHD events and stroke in relation to reductions in blood pressure.

Strengths and limitations of the study

Having individual patient data from the trials would have provided more detail on the effects of blood pressure reduction in relation to pretreatment blood pressure and age. This was, however, not a serious limitation since the trials varied sufficiently for pretreatment blood pressure to be informative. For age, the observation that in the age group covered by the trials (60-69) the results were as expected from the cohort studies indicates that the synthesis of these two sources of data overcomes this limitation from the trial meta-analysis. That our meta-analysis was based on trials in which design varied in many ways may be considered a limitation. The meta-analysis was, however, sensitive enough to show that the trial results were as expected from cohort studies and it is therefore unlikely that random or systematic error in the analysis would produce essentially identical quantitative results when dichotomised in different ways, such as with or without cardiovascular disease. Indeed, the consistency of our results in the face of such variable trial designs reinforces, not diminishes, the validity of the conclusions. There are scarce direct data to show an additive effect of different combinations of three blood pressure lowering drugs on blood pressure but it is reasonable to conclude this given that it is true for combinations of two drugs.[27]

A strength of our analyses is that, based as they are on relative reductions in risk, they are generally applicable irrespective of the incidence of cardiovascular disease. However the preventive potential needs to be assessed in terms of the absolute risk reduction. To do this our estimates of relative risk reduction can be converted to absolute risk reductions by multiplying them by the incidence in a specified population. For example at age 65 the 10 year risk of myocardial infarction (fatal or non-fatal) in England and Wales was estimated at about 10% in men and 5% in women.[35] Given an average blood pressure at that age of 150 mm Hg systolic and 90 mm Hg diastolic [35] the expected relative risk reduction using three drugs at half standard dose is 46% (see bmj.com), so the absolute risk reduction over 10 years in men is 4.6% (from 10% to 5.4%) and in women is 2.3% (from 5% to 2.7%). The corresponding absolute risk reduction for stroke is 2.9% in men and 2.3% in women, based on 10 year incidences of 5% in men and 4% in women.[35] For myocardial infarction and stroke combined, therefore, the absolute risk reduction in men is 7.5% and in women is 4.6%.

Our results are of public health importance. Blood pressure lowering treatment can reduce the incidence of CHD and stroke in the population by at least half in people at risk of CHD events or stroke for any reason including age, whatever a person's blood pressure. Consideration should therefore be given to replacing current policies that focus on routinely measuring blood pressure with policies that focus on routinely lowering blood pressure.

1 Psaty BM, Lumley T, Furberg CD, Schellenbaum G, Pahor M, Alderman MH, et al. Health outcomes associated with various antihypertensive therapies used as first-line agents: a network meta-analysis. *JAMA* 2003;289:2534-44.

2 Kendall MJ, Lynch KP, Hjalmarson Å, Kjekshus J. β-blockers and sudden cardiac death. *Ann Intern Med* 1995;123:358-67.

3 Elliott WJ, Bandari A. The role of calcium antagonists in stroke prevention. *J Clin Hypertens* 2005;7:5-8.

4 Blood Pressure Lowering Treatment Trialists' Collaboration. Effects of different blood pressure-lowering regimens on major cardiovascular events in individuals with and without diabetes mellitus. *Arch Intern Med* 2005;165:1410-9.

5 Blood Pressure Lowering Treatment Trialists' Collaboration. Effects of different blood-pressure-lowering regimens on major cardiovascular events: results of prospectively-designed overviews of randomised trials. *Lancet* 2003;362:1527-35.

6 Blood Pressure Lowering Treatment Trialists' Collaboration. Effects of ACE inhibitors, calcium antagonists, and other blood-pressure-lowering drugs: results of prospectively designed overviews of randomised trials. *Lancet* 2000;355:1955-64.

7 Chrysant SG, Chrysant GS. The pleiotropic effects of angiotensin receptor blockers. *J Clin Hypertens* 2006;8:261-8.

8 Blood Pressure Lowering Treatment Trialists' Collaboration. Blood pressure-dependent and independent effects of agents that inhibit the renin-angiotensin system. *J Hypertens* 2007;25:951-8.

9 Staessen JA, Wang J-W, Thijs L. Cardiovascular protection and blood pressure reduction: a meta-analysis. *Lancet* 2001;358:1305-15.

10 Staessen JA, Wang J-G, Thijs L. Cardiovascular prevention and blood pressure reduction: a quantitative overview updated until 1 March 2003. *J Hypertens* 2003;21:1055-76.

11 Staessen JA, Li Y, Thijs L, Wang J-G. Blood pressure reduction and cardiovascular prevention: an update including the 2003-2004 secondary prevention trials. *Hypertens Res* 2005;28:385-407.

12 Zhang H, Thijs L, Staessen JA. Blood pressure lowering for primary and secondary prevention of stroke. *Hypertension* 2006;48:187-95.

13 Verdecchia P, Reboldi G, Angeli F, Gattobigio R, Bentivoglio M, Thijs L, et al. Angiotensin-converting enzyme inhibitors and calcium channel blockers for coronary heart disease and stroke prevention. *Hypertension* 2005;46:386-92.

14 Lindholm LH, Carlberg B, Samuelsson O. Should β blockers remain first choice in the treatment of primary hypertension? A meta-analysis. *Lancet* 2005;366:1545-53.

15 Carlberg B, Samuelsson O, Lindholm LH. Atenolol in hypertension: is it a wise choice? *Lancet* 2004;364:1684-9.

16 Wiysonge CS, Bradley H, Mayosi BM, Maroney R, Mbewu A, Opie LH, et al. Beta-blockers for hypertension. *Cochrane Database Syst Rev* 2007;(1):CD002003.

17 Bradley HA, Wiysonge CS, Volmink JA, Mayosi BM, Opie LH. How strong is the evidence of use of beta-blockers as first-line therapy for hypertension? Systematic review and meta-analysis. *J Hypertens* 2006;24:2131-41.

18 Khan N, McAlister FA. Re-examining the efficacy of β-blockers for the treatment of hypertension: a meta-analysis. *CMAJ* 2006;12:1737-42.

19 Verma S, Strauss M. Angiotensin receptor blockers and myocardial infarction. *BMJ* 2004;329:1248-9.
20 Pahor M, Psaty BM, Alderman MH, Applegate WB, Williamson JD, Cavazzini C, et al. Health outcomes associated with calcium antagonists compared with other first-line antihypertensive therapies: a meta-analysis of randomised controlled trials. *Lancet* 2000;356:1949-54.
21 National Institute for Health and Clinical Excellence. *Management of hypertension in adults in primary care. NICE clinical guideline 34.* London; NICE, Jun 2006.
22 Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure. *The seventh report of the Joint National Committee on Detection, Evaluation, and Treatment of High Blood Pressure.* NIH Publication No 04-5230. Aug 2004. www.nhlbi.nih.gov/guidelines/hypertension/jnc7full.pdf. 2007.
23 Williams B, Poulter NR, Brown MJ, Davis M, McInnes GT, Potter JF, et al. British Hypertension Society guidelines for hypertension management 2004 (BHS-IV): summary. *BMJ* 2004;328:634-9.
24 Task Force for the Management of Arterial Hypertension of the European Society of Hypertension (ESH) and of the European Society of Cardiology (ESC). 2007 guidelines for the management of arterial hypertension. *J Hypertens* 2007;25:1105-87.
25 Prospective Studies Collaboration. Age-specific relevance of usual blood pressure to vascular mortality: a meta-analysis of individual data for one million adults in 61 prospective studies. *Lancet* 2002;360:1903-13.
26 DerSimonian R, Laird N. Meta-analysis in clinical trials. *Controlled Clin Trials* 1986;7:177-88.
27 Law MR, Wald NJ, Morris JK, Jordan RE. Value of low dose combination treatment with blood pressure lowering drugs: analysis of 354 randomised trials. *BMJ* 2003;326:1427-31.
28 MacMahon S, Peto R, Cutler J, Collins R, Sorlie P, Neaton J, et al. Blood pressure, stroke and coronary heart disease. Part 1, prolonged differences in blood pressure: prospective observational studies corrected for the regression dilution bias. *Lancet* 1990;335:765-74.
29 Asia Pacific Cohort Studies Collaboration. Blood pressure and cardiovascular disease in the Asia Pacific region. *J Hypertens* 2003;21:707-16.
30 Jamerson K, Weber MA, Bakrias GL, Dahlof B, Pitt B, Shi V, et al. Benazepril plus amlodipine or hydrochlorothiazide for hypertension in high risk patients. *N Engl J Med* 2008;359:2417-28.
31 Deary AJ, Schumann AL, Murfet H, Haydock S, Foo RS, Brown MJ. Influence of drugs and gender on the arterial pulse wave and natriuretic peptide secretion in untreated patients with essential hypertension. *Clin Sci* 2002;103:493-9.
32 The CAFÉ Investigators, for the Anglo-Scandinavian Cardiac Outcomes Trial (ASCOT) Investigators. Differential impact of blood pressure-lowering drugs on central aortic pressure and clinical outcome. Principal results of the Conduit Artery Function Evaluation (CAFÉ) Study. *Circulation* 2006;113:1213-25.
33 Asmar RG, London GM, O'Rourke ME, Safar ME. Improvement in blood pressure, arterial stiffness and wave reflections with a very-low-dose perindopril/indapamide combination in hypertensive patients: a comparison with atenolol. *Hypertension* 2001;38:922-6.
34 Blood Pressure Lowering Treatment Trialists' Collaboration. Effects of different regimens to lower blood pressure on major cardiovascular events in older and younger adults: meta-analysis of randomised trials. *BMJ* 2008;336:1121-3.
35 Law M, Wald N, Morris J. Lowering blood pressure to prevent myocardial infarction and strokes: a new preventive strategy. *Health Technol Assess* 2003;7(31).

Effect of virtual reality training on laparoscopic surgery: randomised controlled trial

Christian R Larsen,[1] Jette L Soerensen,[2] Teodor P Grantcharov,[3] Torur Dalsgaard,[4] Lars Schouenborg,[4] Christian Ottosen,[4] Torben V Schroeder,[5] Bent S Ottesen[6]

ABSTRACT

Objective To assess the effect of virtual reality training on an actual laparoscopic operation.

Design Prospective randomised controlled and blinded trial.

Setting Seven gynaecological departments in the Zeeland region of Denmark.

Participants 24 first and second year registrars specialising in gynaecology and obstetrics.

Interventions Proficiency based virtual reality simulator training in laparoscopic salpingectomy and standard clinical education (controls).

Main outcome measure The main outcome measure was technical performance assessed by two independent observers blinded to trainee and training status using a previously validated general and task specific rating scale. The secondary outcome measure was operation time in minutes.

Results The simulator trained group (n=11) reached a median total score of 33 points (interquartile range 32-36 points), equivalent to the experience gained after 20-50 laparoscopic procedures, whereas the control group (n=10) reached a median total score of 23 (22-27) points, equivalent to the experience gained from fewer than five procedures (P<0.001). The median total operation time in the simulator trained group was 12 minutes (interquartile range 10-14 minutes) and in the control group was 24 (20-29) minutes (P<0.001). The observers' inter-rater agreement was 0.79.

Conclusion Skills in laparoscopic surgery can be increased in a clinically relevant manner using proficiency based virtual reality simulator training. The performance level of novices was increased to that of intermediately experienced laparoscopists and operation time was halved. Simulator training should be considered before trainees carry out laparoscopic procedures.

Trial registration ClinicalTrials.gov NCT00311792.

INTRODUCTION

Laparoscopy has become the standard approach for many conditions in most surgical specialties.[1-3] It is, however, associated with a longer operating time and a higher rate of complications during the learning curve of the surgeons. The possibility of overcoming problems during the learning curve by appropriate training and ensuring that surgeons perform a sufficient number of procedures has also been documented.[4]

The technical skills needed for laparoscopic surgery are fundamentally different from those for traditional open surgery, leading to a prolonged learning curve. The primary obstacles in learning laparoscopy are psychomotor and perceptual. The unique nature of

laparoscopic surgery combined with an increasing focus on patients' safety and rights, the present decrease in working hours, and concern over costs of operating theatre time challenge the traditional surgical approach and contribute to a growing need for novel methods to train laparoscopic surgeons.[5] Although virtual reality simulation has the potential to offer important advantages in the area of training for new skills and procedures, evidence on the transfer of skills from the simulated environment to the operating theatre is still limited.[6 7] We investigated the impact of training using a virtual reality simulator on the quality of skills acquired for a key gynaecological procedure.

METHODS

From September 2006 to June 2007 trainees in gynaecological specialty training years 1 and 2, with no experience of advanced laparoscopy (coordination of more than one instrument), were included in the study. They came from seven of nine gynaecology departments in the Zeeland region of Denmark. To ensure that the trainees' baseline characteristics were similar within and between each group, we chose a stratified randomisation based on experience of simple laparoscopy (one instrument). The Clinical Trial Unit at Copenhagen University independently randomised the trainees by computer to intervention or control groups. Randomisation was concealed. All involved departments, supervisors, and theatre staff were blinded to the trainee's group, and the assessors of outcome were blinded to the trainee and their group. The control group was to continue standard clinical education.

The virtual reality laparoscopy simulator program (LapSim Gyn v 3.0.1; Surgical Science, Gothenburg, Sweden) was run on an IBM T42 computer in a docking station (IBM, Armonk, NY) using an interface with a diathermy pedal (Virtual Laparoscopic Interface; Immersion, San Jose, CA). The operations were recorded on DVD for later blinded evaluation. During the operation an observer (CRL or designated TD) recorded the handling of instruments, any involvement of the supervisor, whether the standard procedure for the operation was followed, and whether the recording was done correctly, finalised, and assessed.

The intervention group undertook a training programme in the simulator, comprising training in two basic skills ("lifting and grasping" and "cutting") and one procedure specific task of right sided salpingectomy with preservation of the ovary. The sessions were repeated until the expert criterion level was reached in two consecutive and independent simulations.

The trainees performed the salpingectomy at their gynaecological department, supervised by a senior colleague. The supervisors were allowed to give oral instructions only.

Outcome measures

The primary outcome measure was technical performance, measured as total score (10-50 points) using the objective structured assessment of laparoscopic salpingectomy, comprising a general rating scale and a task specific rating scale.[8] Two independent observers blinded to trainee and allocated group assessed the recorded operations. The secondary outcome measure was operating time in minutes. The reliability of the assessment was determined by the inter-rater agreement and γ coefficient. We present outcomes as medians and interquartile ranges.

Statistical analysis

The power calculation was based on a previous validation study on the procedure specific scale of the objective structured assessment of laparoscopic salpingectomy, showing a difference of six points between novice laparoscopists (0-5 procedures) and intermediately experienced laparoscopists (30-50 procedures).[8] An improvement of skills to the level of 30 or more points was considered acceptable. We therefore chose the minimal relevant difference to be six points. We determined that with an α of 0.05 (two sided) and a power of 80% we required 18 or more trainees. To compensate for possible drop outs, we added a third, totalling 24 trainees.

We present cumulated scores as medians (average score of two observers), compared using non-parametrical analysis (Mann-Whitney U test). We considered a two tailed P value of 0.05 or less to be statistically significant and an inter-rater agreement of 0.8 or more and γ coefficient of 0.8 or more to be acceptable. Analysis was done using SPSS 13.0 for Windows.

RESULTS

The first 24 of 30 eligible trainees were enrolled; 22 (90%) were women, representing the current sex distribution among trainees in obstetrics and gynaecology in

Follow trainee surgeon Sofie Leisby through laparascopic surgery—from practising in VR to a real life procedure—in a video to accompany this study on bmj.com.

Impact of virtual reality simulator training on surgical performance and operation time. Values are medians (ranges; interquartile ranges) unless stated otherwise

Outcome measure	Simulator trained group (n=11)	Control group (n=10)	P value*
Surgical performance:			
Total score (points)	33 (25-39; 32-36)	23 (21-28; 22-27)	<0.001
% reaching ≥30 points	82	0	
Operation time:			
Total time (minutes)	12 (6-24; 10-14)	24 (14-38; 20-29)	<0.001

Inter-rater agreement 0.79. γ coefficient 0.83 (95% confidence interval 0.68 to 0.98).
*Mann-Whitney U test.

WHAT IS ALREADY KNOWN ON THIS TOPIC

The European Working Time Directive has put extra pressure on surgical training programmes

Virtual reality simulators could contribute to the training of core skills for laparoscopy

High grade evidence of the effect of virtual reality simulator training on real operations is sparse

WHAT THIS STUDY ADDS

Training using a virtual reality simulator improved performance in a laparoscopic procedure

Denmark (see bmj.com). The average age was 32.8 years (range 26-42 years), and 23 trainees were right handed. Eleven were randomised to virtual reality training in laparoscopic salpingectomy and 10 were randomised to traditional clinical education (see bmj.com). Two trainees were subsequently excluded from the simulator group and one from the control group.

The median total score on the general and task specific rating scale reached 33 points (interquartile range 32-36 points) in the simulator group and 23 (22-27 points) in the control group (P<0.001, table). The median time to complete the procedure was 12 minutes (interquartile range 10-14 minutes) in the simulator group and 24 (20-29 minutes) in the control group (P<0.001, table).

The median number of simulated salpingectomies needed to reach the proficiency level in the simulator group was 28 (24-32 salpingectomies). The control group was offered simulator training after the study operation; nine of the 11 trainees volunteered and a median of 26 (23-32) simulated operations were needed to reach the proficiency level (P=0.70). A mean 7 hours and 15 minutes (5h 30 min-8h 0 min) was spent on training in the simulator group and 7 hours and 0 min (5h 15 min-7h 45 min) in the control group (P=0.65; see bmj.com). The baseline score was 8 (5-15) in the simulator group and 9 (7-19) in the control group after training (P=0.70; see bmj.com).

The inter-rater agreement was 0.79. The γ coefficient reached 0.83 (95% confidence interval 0.69 to 0.99).

DISCUSSION

Proficiency based virtual reality training in laparoscopic salpingectomy compared with standard clinical education was associated with a clinically important improvement of actual operative skills. The learning curve in the operating theatre was also shorter. On the rating scale used in this study, novices (<5 procedures) scored a median 24 points and intermediately experienced trainees (20-50 procedures) a median 33 points compared with 39 points for experts.[8] After training in a specific procedure to a predefined (proficiency based) level trainees progressed to the performance of an intermediately experienced gynaecologist. By using simulator training it might be possible to bypass the early learning curve and its associated increased rate of complications.[9]

Although operating time might be greater with novice surgeons, the outcomes of a supervised operation ought to be the same. The time to complete the laparoscopic salpingectomy was reduced by half.

The present results emphasise that by using virtual reality simulator training the surgical community can meet the need for proficiency based basic training in laparoscopy. Criterion based procedural training using a virtual reality simulator can also help compensate for reduced working hours by quickly advancing trainees to a higher level of performance. To achieve an average of 28 salpingectomies can take a year or more in clinical practice, compared with eight hours of intensive training using the simulator.

To date no published studies on the transfer of technical skills from simulator to real operations had exceeded grade 2a evidence; in our study the level of evidence is 1b. The conclusion of a meta-analysis was that only few studies possess the necessary quality, that two studies showed a positive effect[10 11] (real operation) and one study no effect[12] (simulated operation) of simulator based training. Another common feature of previous studies was that they were carried out using basic skills rather than procedure specific simulation. We used a procedural simulator, which provides training in psychomotor and cognitive skills. There are probably several reasons for the significant impact on performance and time in our study compared with the other studies. Firstly, the simulator provides a realistic graphic presentation of anatomy in the surgical field and an immediate feedback system. Secondly, using predefined goals (expert proficiency level) encouraged the trainees to rehearse until they reached the maximal effect of training. Thirdly, we studied highly motivated trainees who needed to learn laparoscopic skills.

We measured the impact of simulator training on salpingectomy, a key operation possessing all the core skills needed for most laparoscopic procedures. We did not test external validity and reproducibility beyond the specialty of gynaecology.

The internal consistency of the trial could have been higher if the trainees had operated in the same theatre, using the same technical equipment, and with the same supervisor and staff. However, by showing the effects of simulator training in settings closely resembling a regional simulator training course the external validity was improved. The primary investigator helped the trainee to use the simulator and introduced the different training modules but did not teach laparoscopic techniques. The feedback on performance was based on assessment in the simulator. A designated supporter at the training session could, however, be a source of bias. Finally, performing laparoscopic surgery also consists of identifying diseased anatomy, communication, teamwork, decision making,[13] leadership, alternative plans, and conversion to open surgery if needed.[13-15] These non-technical skills are trained in the currently existing virtual reality systems to a limited degree only.

Conclusion

It is possible to transfer skills acquired during proficiency based training using a virtual reality simulator

to a real operation. Training in proficiency based skills should be incorporated in a comprehensive surgical training and assessment curriculum for all residents before they operate on real patients.

1   Bruhat MA, Pouly JL. Endoscopic treatment of ectopic pregnancies. *Curr Opin Obstet Gynecol* 1993;5:260-6.
2   Keus F, Broeders IA, van Laarhoven CJ. Gallstone disease: surgical aspects of symptomatic cholecystolithiasis and acute cholecystitis. *Best Pract Res Clin Gastroenterol* 2006;20:1031-51.
3   Rosenmuller M, Haapamaki MM, Nordin P, Stenlund H, Nilsson E. Cholecystectomy in Sweden 2000-2003: a nationwide study on procedures, patient characteristics, and mortality. *BMC Gastroenterol* 2007;7:35.
4   Carlsson S, Nilsson A, Wiklund PN. Postoperative urinary continence after robot-assisted laparoscopic radical prostatectomy. *Scand J Urol Nephrol* 2006;40:103-7.
5   Grantcharov TP, Reznick RK. Teaching procedural skills. *BMJ* 2008;336:1129-31.
6   Issenberg SB, McGaghie WC, Petrusa ER, Lee GD, Scalese RJ. Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. *Med Teach* 2005;27:10-28.
7   Sutherland LM, Middleton PF, Anthony A, Hamdorf J, Cregan P, Scott D, et al. Surgical simulation: a systematic review. *Ann Surg* 2006;243:291-300.
8   Larsen CR, Grantcharov TP, Schouenborg L, Soerensen JL, Ottosen C, Ottesen BS. Objective assessment of surgical competence in gynaecological laparoscopy: development and validation of a procedure specific rating scale. *Br J Obstet Gynaecol* 2008;115:908-16.
9   Southern Surgeons Club. A prospective analysis of 1518 laparoscopic cholecystectomies. The Southern Surgeons Club. *N Engl J Med* 1991;324:1073-8.
10   Grantcharov TP, Kristiansen VB, Bendix J, Bardram L, Rosenberg J, Funch-Jensen P. Randomized clinical trial of virtual reality simulation for laparoscopic skills training. *Br J Surg* 2004;91:146-50.
11   Seymour NE, Gallagher AG, Roman SA, O'Brien MK, Bansal VK, Andersen DK, et al. Virtual reality training improves operating room performance: results of a randomized, double-blinded study. *Ann Surg* 2002;236:458-63.
12   Ahlberg G, Heikkinen T, Iselius L, Leijonmarck CE, Rutqvist J, Arvidsson D. Does training in a virtual reality simulator improve surgical performance? *Surg Endosc* 2002;16:126-9.
13   Kneebone RL, Nestel D, Chrzanowska J, Barnet AE, Darzi A. Innovative training for new surgical roles—the place of evaluation. *Med Educ* 2006;40:987-94.
14   Moorthy K, Munz Y, Sarker SK, Darzi A. Objective assessment of technical skills in surgery. *BMJ* 2003;327:1032-7.
15   Lingard L, Reznick R, DeVito I, Espin S. Forming professional identities on the health care team: discursive constructions of the 'other' in the operating room. *Med Educ* 2002;36:728-34.

Accepted: 21 January 2009

Reporting of sample size calculation in randomised controlled trials: review

Pierre Charles,[1,2,3] Bruno Giraudeau,[1,4,5,6] Agnes Dechartres,[1,2,3] Gabriel Baron,[1,2,3] Philippe Ravaud[1,2,3]

[1]INSERM, U738, Paris, France
[2]Université Paris 7 Denis Diderot, UFR de Médecine, Paris
[3]AP-HP, Hôpital Bichat, Département d'Epidémiologie, Biostatistique et Recherche Clinique, Paris
[4]INSERM Centre d'Investigation Clinique 202, Tours, France
[5]Université François Rabelais, Tours
[6]CHRU de Tours, Tours
Correspondence to: P Ravaud, philippe.ravaud@bch.aphp.fr

Cite this as: *BMJ* 2009;338:b1732
doi:10.1136/bmj.b1732

This article is an abridged version of a paper that was published on bmj.com. Cite this article as: *BMJ* 2009;338:b1732

ABSTRACT

Objectives To assess quality of reporting of sample size calculation, ascertain accuracy of calculations, and determine the relevance of assumptions made when calculating sample size in randomised controlled trials.

Design Review.

Data sources We searched MEDLINE for all primary reports of two arm parallel group randomised controlled trials of superiority with a single primary outcome published in six high impact factor general medical journals between 1 January 2005 and 31 December 2006. All extra material related to design of trials (other articles, online material, online trial registration) was systematically assessed. Data extracted by use of a standardised form included parameters required for sample-size calculation and corresponding data reported in results sections of articles. We checked completeness of reporting of the sample size calculation, systematically replicated the sample-size calculation to assess its correctness, then quantified discrepancies between a priori hypothesised parameters necessary for calculation and a posteriori estimates.

Results Of the 215 selected articles, 10 (5%) did not report any sample size calculation and 92 (43%) did not report all the required parameters. The difference between the sample size reported in the article and the replicated sample size calculation was greater than 10% in 47 (30%) of the 157 reports that gave enough data to recalculate the sample size. The difference between the assumptions for the control group and the observed data was greater than 30% in 31% (n=45) of articles and greater than 50% in 17% (n=24). Only 73 trials (34%) reported all data required to calculate the sample size, had an accurate calculation, and used accurate assumptions for the control group.

Conclusions Sample size calculation is still inadequately reported, often erroneous, and based on assumptions that are frequently inaccurate. Such a situation raises questions about how sample size is calculated in randomised controlled trials.

INTRODUCTION

The importance of sample size determination in randomised controlled trials has been widely asserted, and according to the CONSORT statement must be reported in published articles.[1-4] An a priori sample size calculation will determinate the number of participants needed to detect a clinically relevant treatment effect.[5 6]

The conventional approach is to calculate sample size with four parameters: type I error, power, assumptions in the control group (response rate and standard deviation), and expected treatment effect.[5] Type I error and power are usually fixed at conventional levels (5% for type I error, 80% or 90% for power). Assumptions related to the control group are often prespecified on the basis of previously observed data or published results, and the expected treatment effect is expected to be hypothesised as a clinically meaningful effect. The uncertainty related to the rate of events or the standard deviation in the control group[7 8] and to treatment effect could lead to lower than intended power.[6]

We aimed to assess the quality of reporting sample size calculation in published reports of randomised controlled trials, the accuracy of the calculations, and the accuracy of the a priori assumptions.

MATERIALS AND METHODS

Search strategy

We searched MEDLINE via PubMed with the search terms "randomized controlled trials" and "randomised controlled trials" for articles published in six general journals with high impact factors: *New England Journal of Medicine, Journal of the American Medical Association (JAMA), The Lancet, Annals of Internal Medicine, BMJ*, and *PLoS Medicine* between 1 January 2005 and 31 December 2006.

Selection of relevant articles

We included all two arm, parallel group superiority randomised controlled trials with a single primary outcome. We selected the first report that presented the results for the primary outcome.

Data abstraction

For all selected articles, we systematically retrieved and assessed the full published report, any extra material or appendices available online, the study design article, if cited, and the details of online registration of the trial, if mentioned. We recorded the following data.

In the full text of the articles

General characteristics of the studies: including the medical area, whether the trial was multicentre, type of treatment, type of primary endpoint, and funding source.

Details of the a priori sample size calculation: we noted whether the sample size calculation was reported and, if so, the target sample size. We also collected all the parameters used for the calculation. Justification for assumptions made was also recorded.

Observed data as reported in the results section: number of patients randomised and analysed was recorded, and results for the control group. We also noted whether the results of the trial were statistically significant for the primary outcome.

In the online extra material or study design article

We recorded the target sample size and all the required parameters for sample size calculation if different from those reported in the article.

In the trial registration website

We noted the target sample size and all the required parameters for sample-size calculation.

One of us independently completed all data extractions. A second member of the team reviewed a random sample of 30 articles for quality assurance. The κ statistic provided a measure of interobserver agreement. The reviewers were not blinded to the journal name and authors.

Data analysis

Replication of sample size calculation

We replicated the sample size calculation for each article that provided all the data needed for the calculation. If

Reporting of parameters required for a priori sample size calculation for the 215 articles

Parameter	Reporting frequency (%)
α risk	191 (93)
0.05	183 (96)
Two tailed test	119 (65)
One tailed test	7 (4)
Unspecified	57 (31)
0.025 for one tailed test	2 (1)
Adapted for interim analyses	6 (3)
Power	200 (98)
80%	107 (54)
85%	9 (5)
90%	66 (33)
95%	4 (2)
Other values	14 (7)
Assumptions for control group	165 (81)
Justification of assumptions	81 (49)
Results from previous trial	54 (67)
Preliminary study	15 (19)
Observational data	6 (7)
Results of systematic review	2 (3)
Others	4 (5)
Assumptions for the treatment effect	186 (91)
Justification of the assumptions	50 (27)
Analogy to another trial or treatment	41 (82)
Clinical relevance	7 (14)
Observational data	1 (2)
Results of a meta-analysis	1 (2)
All parameters required for sample size calculation	113 (53)

parameters for replicating the sample size were missing in the article and if the calculation was described elsewhere (in the online extra material or study design article) we used the parameters given in this supplemental material. If the missing values were only the α risk or whether the test was one or two tailed, we hypothesised an α risk of 0.05 with a two tailed test to replicate the calculation. The formulae used for the replication are provided and explained in appendix 1. If the absolute value of the standardised difference between the recalculated sample size and the reported sample size was greater than 10%, an independent statistician extracted the data from the full text independently and replicated the sample size calculation again. Any difference between the two calculations was resolved by consensus.

To assess the accuracy of a priori assumptions, we calculated relative differences between hypothesised parameters for the control group and estimated ones reported in the results sections. See bmj.com.

RESULTS

Selected articles

The electronic search yielded 1070 citations, of which 215 articles (appendix 2) met our criteria and were included in our analysis. See bmj.com.

Reporting of required parameters for a priori sample size calculation

Ten articles (5%) did not report any sample size calculation. Only 113 (53%) reported all the required parameters for the calculation (table).

The median of the expected treatment effect for dichotomous or time to event outcomes (relative difference of event rates) was 33.3% (IQR 24.8-50.0) and the median of the expected effect size for continuous outcomes was 0.53 (0.40-0.69).

The design of 35 of the 215 trials (16%) was described elsewhere. In two, the primary outcome described in the report differed from that in the design article. In 31 articles (89%), the data for sample size calculation were given. For 16 articles (52%) the reporting of the assumptions differed from the design article.

Reporting of sample size calculation in online trial registration database

Of the 215 selected articles, 113 (53%) reported registration of the trial in an online database. For 96 articles (85%), an expected sample size was given in the online database and was equal to the target sample size reported in the article in 46 of these articles (48%). The relative difference between the registered and reported sample size was greater than 10% in 18 articles (19%) and greater than 20% in five articles (5%). The parameters for the sample size calculation were not stated in the online registration databases for any of the trials.

WHAT IS ALREADY KNOWN ON THIS TOPIC

Planning and reporting of sample size calculation for randomised controlled trials is recommended by ICH E9 and the CONSORT statement

WHAT THIS STUDY ADDS

Sample size calculations are inadequately reported, often erroneous, and based on assumptions that are frequently inaccurate

These major issues question the foundation of sample size calculation and its reporting in randomised controlled trials

Replication of sample size calculation

We were able to replicate sample size calculations for 164 articles and able to compare our recalculated sample size and the target sample size for 157 articles, since seven did not report any target sample size. The sample size recalculation was equal to the authors' target sample size for 27 articles (17%) and close (absolute value of the difference <5%) for 76 (48%). See bmj.com.

Comparisons between a priori parameters and corresponding estimates in results section

A comparison between the a priori assumptions and observed data was feasible for 145 of the 157 articles reporting enough parameters to recalculate the sample size and reporting the results of the authors' calculations. For the control group, the difference between the assumptions and the observed data was greater than 30% for 45 articles (31%) and greater than 50% for 24 (17%). See bmj.com.

Overall, 73 articles (34%) reported enough parameters for us to replicate the sample size calculation, had an accurate calculation (the replicated sample size calculation differed by less than 10% from the reported target sample size), and had accurate assumptions for the control group (the differences between the a priori assumptions and their estimates was less than 30%).

DISCUSSION

Principal findings

In this survey of 215 reports published in 2005 and 2006 in six general medical journals with high impact factors, only about a third (n=73, 34%) adequately described sample size calculations—that is, they reported enough data to recalculate the sample size, the sample size calculation was accurate, and assumptions in the control group differed less than 30% from observed data. Our study raises two main issues. The first is the inadequate reporting and the errors in sample size calculations, which are surprising in high quality journals with a peer review process; the second is the large discrepancies between the assumptions and the data in the results, which raises a much more complex problem because investigators often have to calculate a sample size with insufficient data to estimate these assumptions.

Reporting of the sample size calculation has greatly increased in the past decades, from 4% of reports describing a calculation in 1980 to 83% of reports in 2002.[8 9] Our review highlights that some parameters for sample size calculation are frequently absent and that miscalculations occur.

We also found large discrepancies between values for assumed parameters in the control group used for sample size calculations and estimated ones from observed data. Assumed values were fixed at a higher or lower level than corresponding data in the results sections in roughly even proportions, a finding different from the results of a previous study. [10]

Our results suggest that researchers, reviewers, and editors do not take reporting of sample size determination seriously.[11] An effort should be made to increase transparency in sample size calculation or, if sample size calculation reporting is of little relevance in randomised controlled trials, perhaps it should be abandoned, as suggested by Bacchetti.[12]

Limitations

An important limitation of this study is that we could not directly assess whether assumptions had been manipulated to obtain feasible sample sizes because we used only published data.[6 13]

Implications

A major discrepancy exists between the importance given to sample size calculation by funding agencies, ethics review boards, journals, and investigators and the current practice of sample size calculation and reporting.[14] We therefore believe, as do others, that there is room for reflection on how sample size should be determined for randomised trials. After years of trials with supposedly inadequate sample sizes, it is time to develop and use new ways of planning sample sizes.

1   ICH Harmonised Tripartite Guideline. Statistical principles for clinical trials. International Conference on Harmonisation E9 Expert Working Group. *Stat Med* 1999;18:1905-42.
2   Moher D, Dulberg CS, Wells GA. Statistical power, sample size, and their reporting in randomized controlled trials. *JAMA* 1994;272:122-4.
3   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.
4   Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *Lancet* 2001;357:1191-4.
5   Machin D, Campbell M, Fayers P, Pinol A. Sample size tables for clinical studies. 2nd ed. Oxford: Blackwell Science, 1997.
6   Schulz KF, Grimes DA. Sample size calculations in randomised trials: mandatory and mystical. *Lancet* 2005;365:1348-53.
7   Weaver CS, Leonardi-Bee J, Bath-Hextall FJ, Bath PM. Sample size calculations in acute stroke trials: a systematic review of their reporting, characteristics, and relationship with outcome. *Stroke* 2004;35:1216-24.
8   Meinert CL, Tonascia S, Higgins K. Content of reports on clinical trials: a critical review. *Control Clin Trials* 1984;5:328-47.
9   Mills EJ, Wu P, Gagnier J, Devereaux PJ. The quality of randomized trial reporting in leading medical journals since the revised CONSORT statement. *Contemp Clin Trials* 2005;26:480-7.
10   Vickers AJ. Underpowering in randomized trials reporting a sample size calculation. *J Clin Epidemiol* 2003;56:717-20.
11   Altman DG, Moher D, Schulz KF. Peer review of statistics in medical research. Reporting power calculations is important. *BMJ* 2002;325:491.
12   Bacchetti P. Peer review of statistics in medical research: the other problem. *BMJ* 2002;324:1271-3.
13   Chan AW, Hrobjartsson A, Jorgensen KJ, Gotzsche PC, Altman DG. Discrepancies in sample size calculations and data analyses reported in randomised trials: comparison of publications with protocols. *BMJ* 2008;337:a2299.
14   Guyatt GH, Mills EJ, Elbourne D. In the era of systematic reviews, does the size of an individual trial still matter? *PLoS Med* 2008;5:e4.

From our archive
Childhood malignancies and x-ray exposure (1958)

The pre-natal and post-natal experiences of a large group of children who recently died of malignant diseases have been compared, point by point, with the experiences of a similar group of live children.

The frequency of three pre-natal events—namely, direct foetal irradiation, virus infections and threatened abortion—was significantly higher among the dead children than among the live children.

One other pre-natal influence—namely, excessive maternal age—appears to increase the risk of leukaemia in childhood and to be related to the fact that this disease and mongolism tend to occur together.

The frequency of three post-natal events—namely, x-ray exposures in infancy, acute pulmonary infections and severe injuries—was significantly higher for children who subsequently died of leukaemia than for other children . . .

Our final conclusions are that foetal irradiation does not account for the recent increase in childhood malignancies, but the finding of a case excess for this event does underline the need to use minimum doses for essential medical x-ray examinations and treatments.

Stewart A, Webb J, Hewitt D. A survey of childhood malignancies. *BMJ* 1958;i:1495-508, doi:10.1136/bmj.1.5086.1495

The entire archive of the *BMJ*, going back to 1840, is now available at www.bmj.com/archive.
Cite this as: *BMJ* 2009;338:b1268

p i c o

Long term monitoring in patients receiving treatment to lower blood pressure: analysis of data from placebo controlled randomised controlled trial

Katherine Keenan,[1] Andrew Hayen,[1] Bruce C Neal,[2] Les Irwig[1]

STUDY QUESTION How does increasing the time interval between blood pressure monitoring measurements affect the assessment of adequacy of long term treatment with blood pressure lowering drugs?

SUMMARY ANSWER Current monitoring strategies are a poor method of determining adequacy of long term treatment with blood pressure lowering drugs. Blood pressure monitoring intervals can be lengthened for most patients.

Participants and setting

1709 patients with a history of stroke or transient ischaemic attack drawn from 172 centres in Asia, Australasia, and Europe and treated with fixed doses of perindopril and indapamide as part of the perindopril protection against recurrent stroke study (PROGRESS).

Design, size, and duration

Estimates over 33 months of the probability that measured rises in an individual's blood pressure are a reflection of a true rise in blood pressure above a treatment threshold. Recorded blood pressure was the mean of two measurements taken five minutes apart with the patient seated, measured to the nearest 2 mm Hg. We assumed that, at three months after randomisation, each patient would have had a response to treatment and would start long term monitoring. We estimated long term change in blood pressure from a true baseline blood pressure from the difference between it and the subsequent follow-up "observed" blood pressure measurements at each time point (three months, six months, nine months, etc). We estimated the proportion of patients whose assumed true baseline systolic blood pressure of 130 or 120 mm Hg would truly increase by 10 or 20 mm Hg to a threshold of 140 mm Hg or above while receiving treatment. We also estimated the proportion of patients who would be observed to be above the thresholds and who actually had true blood pressure levels above the threshold (true positives) or below the threshold (false positives).

Main results and the role of chance

An observed increase in an individual's blood pressure above widely accepted treatment thresholds was much more likely to reflect usual day to day variability than a true increase in blood pressure. For example, if six months after achieving a systolic blood pressure of 130 mm Hg a patient recorded a measurement of 140 mm Hg or above, the reading was six times more likely to reflect usual day to day variability than to be truly above 140 mm Hg.

Bias, confounding, and other reasons for caution

The participants in this study were a selected group of patients, with many taking other blood pressure lowering drugs and other medication.

Generalisability to other populations

The amount of variability in observed blood pressure is likely to be greater in general population settings than in a randomised controlled trial like PROGRESS. Accordingly, the results presented here may overestimate the ability of current long term monitoring to detect changes in blood pressure.

Study funding/potential competing interests

This is a summary of a paper that was published on bmj.com as *BMJ* 2009;338:b1492

PROPORTION OF HIGH BLOOD PRESSURE MEASUREMENTS THAT TRULY EXCEEDED THRESHOLD OF 140 MM HG SYSTOLIC OR 90 MM HG DIASTOLIC			
Blood pressure initially achieved on treatment	Percentage of patients with follow-up measurements truly above threshold	Percentage of patients with follow-up measurements observed to be above threshold (true positive + false positive)	Ratio of false-positive to true-positive results
Systolic:			
120 mm Hg	0.02	3.4 (0.01+3.4)	>200
130 mm Hg	4.6	19.0 (2.7+16.3)	6.0
Diastolic:			
80 mm Hg	0.3	7.2 (0.2+7.0)	39
85 mm Hg	8.5	23.2 (5.2+18.0)	3.5

pico

Prevalence of severe congenital heart disease after folic acid fortification of grain products: time trend analysis in Quebec, Canada

Raluca Ionescu-Ittu,[1] Ariane J Marelli,[2][3] Andrew S Mackie,[4] Louise Pilote[2][5]

[1]Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, Canada

[2]McGill University, Montreal, Canada

[3]McGill Adult Unit for Congenital Heart Disease Excellence (MAUDE), McGill University Health Centre, Montreal, Canada

[4]University of Alberta, Division of Cardiology, Stollery Children's Hospital, Edmonton, Canada

[5]Division of General Internal Medicine, McGill University Health Centre, Montreal, Canada

Correspondence to: L Pilote, Royal Victoria Hospital, V Building, 687 Pine Avenue West, Room V2.17, Montreal, QC, Canada H3A 1A1
louise.pilote@mcgill.ca

STUDY QUESTION Did the folic acid fortification policies introduced in Quebec, Canada, in 1998 affect the birth prevalence of severe congenital heart defects?

SUMMARY ANSWER Mandatory fortification of grain products with folic acid was followed by a significant decrease in the birth prevalence of severe congenital heart defects (6.2% decrease per year). The change in time trend between the period before and the period after fortification was significant overall and for conotruncal and non-conotruncal defects.

Participants and setting

The study population included all infants (live births and stillbirths) born in Quebec, 1990-2005, with severe congenital heart defects identified from administrative databases with complete population coverage. The term "severe congenital heart defects" included tetralogy of Fallot, endocardial cushion defects, univentricular hearts, truncus arteriosus, and transposition complexes.

Design, size, and duration

Time trend analysis estimated the change in birth prevalence of severe congenital heart defects before and after the implementation of fortification. The birth prevalence of severe congenital heart defects was determined annually as the number of infants (live births and stillbirths) born with severe congenital heart defects per 1000 births. We estimated the time trends in the period before and after fortification by Poisson regression with the annual birth prevalence as the dependent variable and the calendar year (birth year) as the independent variable.

Main results and the role of chance

There was no change in the birth prevalence of severe congenital heart defects in the period before fortification (time trend rate ratio 1.01, 0.99 to 1.03), while the period after fortification was accompanied by a 6.2% decrease per year in birth prevalence (0.94, 0.90 to 0.97).

The interaction between period (before/after fortification) and calendar year was significant (P<0.001), suggesting that the decreasing trend after fortification did not occur by chance. We obtained similar results for the birth prevalence of conotruncal defects (time trend rate ratio 1.025, 0.997 to 1.053, before fortification, and 0.95, 0.91 to 0.99, after fortification) and non-conotruncal defects (0.98, 0.95 to 1.02, *v* 0.91, 0.86 to 0.97). The change in time trend between the two periods was significant for both conotruncal defects (P=0.004 for interaction) and non-conotruncal defects (P=0.02).

Bias, confounding, and other reasons for caution

We could not directly measure the changes over time in other risk factors for severe congenital heart defects, which might have confounded the observed temporal changes in birth prevalence. Despite the limited descriptive data available on concurrent secular trends in other risk factors, we believe that our results are not caused by chance because the timing of the observed effect coincides exactly with the timing of the fortification, there is biological plausibility for this association, and, with the exception of natural and elective abortions (on which we have no data from Quebec), all other factors known to increase the birth prevalence of severe congenital heart defects (older maternal age, more pregnancies in women with congenital heart defects, maternal medication use, or obesity) have gradually increased over the study period in Quebec and thus cannot explain the results.

Generalisability to other populations

Folic acid fortification policies might have a different impact in countries with other diets and different use of vitamin supplements.

Study funding/potential competing interests

TIME TRENDS IN BIRTH PREVALENCE OF SEVERE CONGENITAL HEART DISEASE BEFORE AND AFTER 1 JANUARY 1999