RESEARCH

The *BMJ* is an Open Access journal. We set no word limits on *BMJ* research articles, but they are abridged for print. The full text of each *BMJ* research article is freely available on bmj.com

Weighing the benefits and harms of mammography screening

The controversy over the value of mammography screening for breast cancer has continued for years, with a Cochrane review in 2009 concluding: "It is thus not clear whether screening does more good than harm." This week two modelling studies attempt to clarify the balance of benefits and harms from screening.

James Raftery and Maria Chorozoglou (p 14) have updated the life table analysis of the 1986 Forrest report (which led to the introduction of the UK breast screening programme) and combined harms and benefits of screening into a single measure (quality adjusted life years (QALYs)). They found negative net QALYs in the early years after the start of screening (up to 10 years after), after which positive net QALYs accrued, but at a slower rate than reported in the original Forrest report.

The main potential harms of mammography screening are overdiagnosis and overtreatment (detection and removal of lesions that would not have progressed to invasive cancer), and Arnaud Seigneurin and colleagues (p 15) designed a stochastic simulation model to estimate the degree of overdiagnosis among women aged 50-70 in a region of France. They calculated that overdiagnosis accounted for only 1.5% of all cases of invasive cancer and for 28% of all cases of carcinoma in situ (though carcinoma in situ accounted for less than 15% of all incident cases of breast cancer). They conclude that overdiagnosis of invasive cancers was smaller than expected—especially compared with studies that looked at incidence rates before and after implementation of breast screening, which have reported 30-50% overdiagnosis for invasive cancers.

SPL

However, as the accompanying editorial by Allan Hackshaw points out (p 7), the results of such modelling studies depend on the reliability of the parameters used and the underlying assumptions of the model. Readers' responses to the full studies published on bmj.com show just how controversial these can be.

Research online: For these and other new research articles see www.bmj.com/research

Influence of experience on performance of individual surgeons in thyroid surgery In this prospective cross sectional study by Antoine Duclos and colleagues, patients were at increased risk of permanent complications after a thyroidectomy when operated on by inexperienced surgeons or those who had spent the longest time in practice since graduation. Surgeons aged 35-50 years provided the safest care (doi:10.1136/bmj.d8041).

Effects of glucagon-like peptide-1 receptor agonists on weight loss Treatment with glucagon-like peptide-1 receptor agonists for at least 20 weeks leads to weight loss in obese or overweight patients with or without type 2 diabetes mellitus, according to results of a meta-analysis by Tina Vilsbøll and colleagues. The effect of the drugs might be more pronounced in patients without diabetes, and they also reduce systolic and diastolic blood pressure and total cholesterol, say the authors (doi:10.1136/bmj.d7771).

Do hip and knee replacements last longer with bisphosphonates?

SPL

Another study generating many responses online (http://bit.ly/wWXMmM) is Daniel Prieto-Alhambra and colleagues' cohort study of bisphosphonates in patients who've undergone total arthroplasty of the knee or hip (p 17). These operations are very common, but so is subsequent revision surgery—as the authors say in the full paper online, about one in 75 patients needed a revision of their prosthesis within three years in 2003—and loosening of the implant, which occurs when the bone supporting it is resorbed, is the most common cause. Bisphosphonates, which reduce bone resorption, are therefore of interest as a potential way to prolong the life of a hip or knee implant, reducing the need for revision surgery.

The retrospective study took data from the United Kingdom's General Practice Research Database on patients followed up for a median of 3.5 years after primary hip or knee replacement, and concluded that treatment with bisphosphonates did indeed show a strong association with implants lasting longer. However, as the authors state, randomised trials will be needed to confirm such an association, as this study was observational, although the analysis was adjusted for several confounding factors. Responses to the paper online reflect further on possible confounding factors in the study, as well as questioning whether the long term adverse effects of bisphosphonates might outweigh the apparent short term benefits, and speculating on whether the time at which patients receive bisphosphonates is important to the association with implant survival.

Possible net harms of breast cancer screening: updated modelling of Forrest report

James Raftery, Maria Chorozoglou

EDITORIAL by Hackshaw
RESEARCH p 15

Wessex Institute, Faculty of Medicine, University of Southampton, Southampton SO16 7PX, UK
Correspondence to: J P Raftery
j.p.raftery@soton.ac.uk

Cite this as: *BMJ* 2011;343:d7627
doi: 10.1136/bmj.d7627

This is a summary of a paper that was published on bmj.com as *BMJ* 2011;343:d7627

STUDY QUESTION

Could screening for breast cancer really cause more harm than good?

SUMMARY ANSWER

Harms as a result of screening largely offset the benefits up to 10 years, after which benefits accumulate rapidly but much less than originally expected.

WHAT IS KNOWN AND WHAT THIS PAPER ADDS

In the 1980s the Forrest report led to the introduction of the United Kingdom's breast screening programme. A subsequent Cochrane review concluded that: "for every 2000 women invited for screening throughout 10 years, one will have her life prolonged, and 10 healthy women, who would not have been diagnosed if there had not been screening, will be diagnosed as breast cancer patients and will be treated unnecessarily. Furthermore, more than 200 women will experience important psychological distress for many months because of false positive findings. It is thus not clear whether screening does more good than harm." An updated analysis of the Forrest report combining the harms and benefits into a single measure (quality adjusted life years (QALYs)), showed negative net QALYs in the early years after screening start, after which positive net QALYs accrue but at a slower rate than initially expected.

Participants and setting

Computer modelling followed two identical cohorts comprising 100 000 healthy women aged 50. One cohort was invited for mammographic breast cancer screening, the other was not. We applied English rates for breast cancer mortality and surgery for 1985—that is, before screening—to both cohorts but with reduced mortality and increased surgery in the screened cohort. The changes in mortality and surgery were based on data from meta-analyses of all relevant randomised trials.

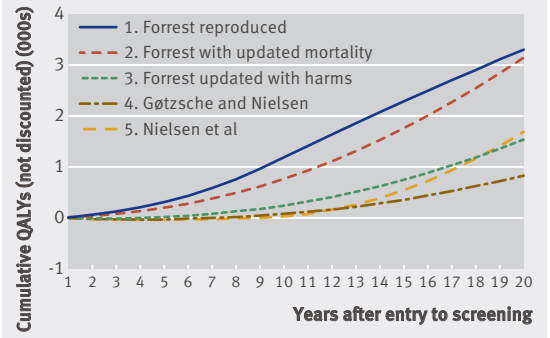
Design, size, and duration

We included life table estimates of women alive by cohort each year for 20 years and the numbers having false positive diagnoses and surgery, with linked losses in quality of life. The losses in quality of life from false positive results and surgery were based on data from literature reviews and recent trials.

Main results and the role of chance

Five scenarios explored the cumulative net QALYs gained from screening. Two scenarios (1 and 2) reproduced and updated the estimates from the Forrest report for QALYs, providing a baseline against which to assess the effect of adding harms. Scenario 3 (updated Forrest including harms) showed that adding harms halved the QALY gains. Scenario 4, using a published best estimate of the reduction in mortality from breast cancer, had negative QALYs for the first seven years after screening. Scenario 5, which assumed a greater reduction in mortality in women aged 60-70, led to higher net QALYs but only in the longer term. The different timing of harms (early) and benefits (late) was notable. Sensitivity analysis explored the effects of varying key parameter values. While the QALY totals varied, the pattern of low or negative net QALYs in the early years after the introduction of screening did not.

Net QALYs after start of breast screening with different modelling scenarios



Bias, confounding, and other reasons for caution

The greatest uncertainty applies to the duration of losses of quality of life from surgery, which we assumed to be permanent. Limiting that loss to shorter periods led to higher net QALYs but only in the longer term. The applicability of randomised trials completed decades ago could be queried, but the effect of screening on mortality from breast cancer is likely to have lessened. As we used recent English estimates for false positive results and the loss of quality of life from surgery, the results take into account recent changes including the trend towards less extreme surgery. Inclusion of the effects of radiotherapy and chemotherapy would lead to greater negative QALYs.

Generalisability to other populations

Screening an older cohort would increase the net QALYs because the incidence of breast cancer rises with age, but with the same pattern of early harms and delayed gains.

Study funding/potential competing interests

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

bmj.com/podcasts
Interview about the Forrest report update with James Raftery, director of Southampton University's Health Technology Assessment programme

Overdiagnosis from non-progressive cancer detected by screening mammography: stochastic simulation study with calibration to population based registry data

Arnaud Seigneurin,[1 2] Olivier François,[2] José Labarère,[2 3] Pierre Oudeville,[2] Jean Monlong,[2] Marc Colonna[1]

EDITORIAL by Hackshaw
RESEARCH p 14

[1]Registre du Cancer de l'Isère, Centre Hospitalier Universitaire de Grenoble, BP 217, Pavillon E, 38043 Grenoble Cedex 9, France
[2]Université Joseph Fourier Grenoble 1, CNRS, TIMC-IMAG UMR 5525, 38041 Grenoble
[3]Unité d'évaluation Médicale, Centre Hospitalier Universitaire de Grenoble, Grenoble
Correspondence to: A Seigneurin
aseigneurin@chu-grenoble.fr

Cite this as: *BMJ* 2011;343:d7017
doi: 10.1136/bmj.d7017

This is a summary of a paper that was published on bmj.com as *BMJ* 2011;343:d7017

STUDY QUESTION

What is the magnitude of overdiagnosis resulting from the detection of non-progressive cancer by screening mammography?

SUMMARY ANSWER

Overdiagnosis was of limited magnitude, ranging from 1.5% of all cases of invasive cancer to 28% of all cases of carcinoma in situ in a population offered organised and opportunistic mammography screening.

WHAT IS KNOWN AND WHAT THIS PAPER ADDS

Mammography screening is associated with reduced mortality rates from breast cancer in women aged 50-70 but also exposes women to harm, including false positive results, low dose radiation, and overdiagnosis. This stochastic simulation model, designed to replicate standardised incidence rates of breast cancer, showed that overdiagnosis of invasive cancers was smaller than expected.

Participants and setting

We studied the incidence of breast cancers among women aged 50-69 between 1991 and 2006, living in Isère, France, a French administrative region with nearly 1.2 million inhabitants.

Design

We designed a stochastic simulation model to replicate standardised incidence rates of breast cancer and estimated overdiagnosis using an approximate Bayesian computation approach. The model components included the lifetime probability of breast cancer, the natural course of breast cancer, and participation in organised and opportunistic mammography screening. A total of simulated 100 000 datasets were simulated, each of them comprising 245 000 women.

Primary outcome

The primary outcome was overdiagnosis resulting from the detection of in situ and invasive non-progressive cancers by mammography screening.

Main results and the role of chance

Of 100 000 simulated datasets, we retained 500 (0.5%) with the smallest difference between the simulated and observed values of standardised incidence rates of breast cancer. Overdiagnosis accounted for 1.5% of all cases of invasive cancer (95% credibility interval 0.3% to 2.9%) and 28.0% of all cases of carcinoma in situ (2.2% to 59.8%) detected either clinically or by screening mammography. Because carcinoma in situ accounted for less than 15% of all incident cases of breast cancer, the estimate of overdiagnosis was less precise for this subgroup of cancer. When the analysis was restricted to the cancers detected by screening mammography only, the estimates of overdiagnosis were 3.3% (0.7% to 6.5%) and 31.9% (2.9% to 62.3%) for invasive cancer and carcinomas in situ, respectively.

Bias, confounding, and other reasons for caution

This simulation study accounted for important biases that can affect estimates of overdiagnosis. Firstly, the issue of opportunistic screening, which could contribute to underestimating overdiagnosis, was examined by simulating the probability of undergoing screening mammography on either an organised or opportunistic basis. Secondly, because estimates of overdiagnosis might be affected by secular changes in background risk of breast cancer, the model allowed for the possibility of an increasing linear trend in the lifetime probability of breast cancer. Thirdly, the model was adjusted for lead time by simulating the length of preclinical phases with various distributions. Overdiagnosis resulting from progressive cancers censored because of competing causes of death was not in the scope of this study.

Generalisability to other populations

We cannot exclude that the findings would be different in other countries or settings because of the specificities relative to the epidemiology of breast cancer and to the screening procedure in Isère. Indeed, Isère ranked among the regions with the highest incidence of breast cancer worldwide, with an estimated standardised incidence rate of 97.8 per 100 000 person years in 2003-6. The participation rate in organised screening was low, ranging from 25% to 30% in the early 1990s. In 2002, the programme was extended to women aged 50-74 and included clinical breast examination and two view mammography, and the screening interval was shortened from 30 to 24 months. Finally, opportunistic individual screening coexisted with the breast cancer screening programme.

Study funding/potential competing interests

This study was funded by grants from the Institut National du Cancer, Paris, France, and the Comité de l'Isère de la Ligue Nationale Contre le Cancer, Grenoble, France.

bmj.com/podcasts

Interview about the Forrest report update with James Raftery, director of Southampton University's Health Technology Assessment programme

Percentage estimates of overdiagnosis (95% credibility interval) among women aged 50-69 between 1991 and 2006 in Isère, France

	Carcinoma in situ	Invasive cancer
Among all cancers diagnosed	28.0 (2.2 to 59.8)	1.5 (0.3 to 2.9)
Among cancers screened	31.9 (2.9 to 62.3)	3.3 (0.7 to 6.5)

Immediate and late benefits of treating very elderly people with hypertension: results from active treatment extension to Hypertension in the Very Elderly randomised controlled trial

N Beckett,[1] R Peters,[2] J Tuomilehto,[3 4] C Swift,[5] P Sever,[6] J Potter,[7] T McCormack,[8] F Forette,[9] B Gil-Extremera,[10] D Dumitrascu,[11] J A Staessen,[12] L Thijs,[13] A Fletcher,[14] C Bulpitt,[1] for the HYVET Study Group

EDITORIAL by Mancia

[1]Department of Medicine, Imperial College London, London SW7 2AZ, UK
[2]Clinical Trials Unit, Imperial College London
[3]Department of Public Health, University of Helsinki, Helsinki, Finland
[4]South Ostrobothnia Central Hospital, Seinäjoki, Finland
[5]Clinical Age Research Unit, Department of Clinical Gerontology, King's College London, London
[6]International Centre for Circulatory Health, Imperial College London
[7]School of Medicine, Health Policy and Practice, University of East Anglia, Norfolk, UK
[8]Whitby Group Practice, Spring Vale Medical Centre, Whitby, UK
[9]Hôpital Broca, University Paris V, Paris, France
[10]FAC Medicina, Departmento Medicina Granada, Granada, Spain
[11]Spitalul Judetean Cluj, Clinica Medicală 2, Cluj, Romania
[12]Department of Epidemiology, Maastricht University, Netherlands
[13]Studies Coordinating Centre, Division of Hypertension and Cardiovascular Rehabilitation, Department of Cardiovascular Diseases, University of Leuven, Leuven, Belgium
[14]London School of Hygiene and Tropical Medicine, London
Correspondence to N Beckett
n.beckett@imperial.ac.uk

Cite this as: *BMJ* 2011;343:d7541
doi: 10.1136/bmj.d7541

This is a summary of a paper that was published on bmj.com as *BMJ* 2011;343:d7541

STUDY QUESTION

Does starting antihypertensive treatment in people aged 80 or over reduce ardiovascular events quickly enough to support starting treatment?

SUMMARY ANSWER

Treatment based on indapamide SR 1.5 mg reduced cardiovascular events within 12 months, supporting the early accrual of benefits from treating very elderly people who are hypertensive.

WHAT IS KNOWN AND WHAT THIS PAPER ADDS

HYVET suggested that treating people aged 80 or over with sustained systolic blood pressures of 160 mm Hg or above was beneficial. This open label active treatment extension found that, in patients previously taking placebo, cardiovascular benefits can be achieved within one year, supporting the need to start treatment in such people.

Design

This was a one year, open label, active treatment extension of the randomised, placebo controlled Hypertension in the Very Elderly Trial (HYVET). Active treatment used in the main trial was indapamide SR 1.5 mg, with the addition of perindopril 2-4 mg as required to achieve the target blood pressure. At the start of the one year extension, all participants were restarted on indapamide. Treatment was titrated to achieve the target of systolic blood pressure below 150 mm Hg and diastolic pressure below 80 mm Hg.

Participants and setting

Eligibility for the main trial included being aged 80 years or over with sustained levels of systolic blood pressure of 160 mm Hg or more. To enter the extension, participants had to be on double blind treatment at the final visit of the main trial. The trial took place in hospital and general practice based centres mainly in eastern and western Europe, China, and Tunisia.

Primary outcome(s)

We collected the same endpoint data as during the main trial, including all strokes (fatal and non-fatal), total mortality, and cardiovascular events (cardiovascular mortality, non-fatal myocardial infarction, non-fatal stroke, and non-fatal heart failure).

Main results and the role of chance

Of the 1882 participants on double blind treatment at the end of the main trial (1009 on active trial treatment, 873 on placebo), 1712 (91%) consented to enter the one year extension. Of these, 788 (46%) were previously taking placebo and 924 (54%) active treatment. By six months and beyond, we found no statistical difference in blood pressure between people previously on active treatment and those previously on placebo (mean blood pressures 145.3/76 mm Hg and 146.6/76.6 mm Hg). During the one year extension, 47 participants died (11 from cardiovascular causes) and 1682 patient years of follow-up were accrued. Comparing people previously treated with active drug and those previously on placebo, we found no significant differences for stroke (the primary end point), cardiovascular events, or heart failure. We found significant differences for total mortality (P=0.016) and cardiovascular mortality (P=0.033).

Harms

No serious adverse drug reactions were reported. No difference in serious adverse events existed between people previously on active drug and those previously on placebo (46 *v* 53; P=0.12).

Bias, confounding, and other reasons for caution

The small number of events necessitates caution in generating firm conclusions from the results. This is an extension of a trial and thus a more select group.

Generalisability to other populations

The results apply to people aged 80 or over with sustained blood pressures of 160 mm Hg or above who are free of dementia and do not need regular nursing intervention or have many comorbidities.

Study funding/potential competing interests

HYVET was funded by grants from the British Heart Foundation and the Institute de Recherches Internationales Servier. NB and RP received salary support from the grant from Servier; CB, JP, TM, and FF have received consulting/speakers' fees from various companies that manufacture antihypertensive drugs.

Trial registration number Clinical trials NCT00122811.

Main fatal and fatal plus non-fatal outcomes for intention to treat analysis

End points	Rate per 1000 patient years (No of events)		Hazard ratio (95% CI)
	Previously on placebo	Previously on active treatment	
Fatal and non-fatal stroke	5.18 (4)	9.89 (9)	1.92 (0.59 to 6.22)
All cause mortality	38.8 (30)	18.6 (17)	0.48 (0.26 to 0.87)
Non-cardiovascular or unknown cause	14.2 (11)	6.6 (6)	0.46 (0.17 to 1.25)
Cardiovascular mortality	11.6 (9)	2.19 (2)	0.19 (0.04 to 0.87)
All heart failure	3.9 (3)	1.1 (1)	0.28 (0.03 to 2.73)
All cardiovascular events	16.9 (13)	13.2 (12)	0.78 (0.36 to 1.72)

Association between bisphosphonate use and implant survival after primary total arthroplasty of the knee or hip: population based retrospective cohort study

Daniel Prieto-Alhambra,[1 2 3 4 5] M Kassim Javaid,[1 6] Andrew Judge,[1] David Murray,[1] Andy Carr,[1] Cyrus Cooper,[1 6] Nigel K Arden[1 6]

[1]Oxford NIHR Musculoskeletal Biomedical Research Unit, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford OX3 7LD, UK
[2]Institut Català de la Salut, Barcelona, Spain
[3]Unitat de Recerca en Fisiopatologia Òssia i Articular (URFOA), Institut Municipal d'Investigació Mèdica (IMIM), Barcelona
[4]Department of Medicine, Universitat Autonoma de Barcelona, Barcelona
[5]Primary Care Research Institute, IDIAP (Institut D'Investigació en Atenció Primària) Jordi Gol, Barcelona
[6]MRC Lifecourse Epidemiology Unit, University of Southampton, Southampton General Hospital, Southampton, UK
Correspondence to: N Arden
nigel.arden@ndorms.ox.ac.uk

Cite this as: *BMJ* 2011;343:d7222
doi: 10.1136/bmj.d7222

This is a summary of a paper that was published on bmj.com as *BMJ* 2011;343:d7222

STUDY QUESTION

Can bisphosphonate use improve implant survival after total arthroplasty of the lower limb?

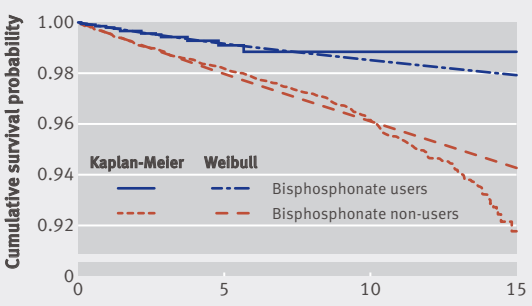
SUMMARY ANSWER

Bisphosphonate use was associated with an almost twofold increase in implant survival time after primary total arthroplasty of the knee or hip.

WHAT IS KNOWN AND WHAT THIS PAPER ADDS

Bisphosphonates have theoretical benefits on implant survival, but direct evidence is lacking. In this population based cohort study, we found that bisphosphonate use is associated with improved implant survival and reduced rates of revision.

Implant survival in bisphosphonate users versus non-users



Participants and setting

We included all patients undergoing primary total arthroplasty of the knee (n=18 726) or hip (n=23 269) in 1986-2006 within the United Kingdom's General Practice Research Database. We excluded patients younger than 40 years at surgery and those with a history of hip fracture or rheumatoid arthritis before surgery.

Design, size, and duration

We did a population based retrospective cohort study including 41 995 participants and followed up participants for a median of 3.5 years. We classified bisphosphonate users as participants with at least six prescriptions of bisphosphonates or at least six months' prescribed bisphosphonate treatment with more than 80% adherence before revision surgery. Propensity scores were used to adjust for confounding by indication.

Main results and the role of chance

Of 41 995 eligible participants, we identified 1912 (4.6%) bisphosphonate users. Overall, we recorded 511 (1.3%) revisions in bisphosphonate non-users, with a reduced rate in users (eight (0.8%) hip; three (0.3%) knee). For patients with at least five years' follow-up, users had a lower revision rate than non-users (0.93% (95% confidence interval 0.52% to 1.68%) *v* 1.96% (1.80% to 2.14%)). Bisphosphonates seemed to have a strongly protective effect on implant survival throughout the study (adjusted hazard ratio 0.54 (0.29 to 0.99), P=0.047), with a significant increase in median prosthesis survival (time ratio 1.96 (1.01 to 3.82); fig).

Bias, confounding, and other reasons for caution

The main limitation of this study was the observational nature of the data and the lack of validation for each individual event. Although we used standard methods to adjust for confounding, it is still unclear whether observational studies can accurately estimate the effects of drugs on outcomes. Hence, formal randomised controlled trials are needed to confirm these results.

Generalisability to other populations

The main advantage of population based cohort studies over randomised trials is that they provide data from a wider range of population groups. In addition, data from the General Practice Research Database accurately reflect primary care in the UK, because they are collected throughout practice.

Study funding/potential competing interests

This research was commissioned by the National Institute for Health Research (NIHR) under its Programme Grants for Applied Research and also received funding from: the NIHR Biomedical Research Unit into Musculoskeletal Disease, Nuffield Orthopaedic Centre, and University of Oxford; Institut Catala de la Salut-IDIAP Jordi Gol; and Merck, Sharpe and Dohme, Novartis, and Southampton Rheumatology Trust MKJ, NKA, and CC have received honorariums, held advisory board positions, and received consortium research grants from several pharmaceutical companies (see full article for details).

bmj.com

- Editor's choice: The trouble with medical devices (*BMJ* 2011;342:d3123)
- Editorial: The regulation of medical devices (*BMJ* 2011;342:d2822)

doc2doc

- Are we rigorous enough about our devices? http://bit.ly/z5ifmz