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

THIS WEEK'S RESEARCH QUESTIONS

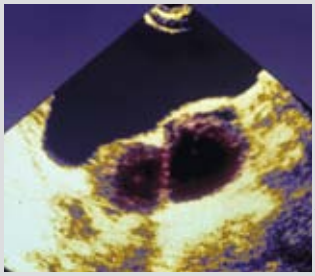
91 Is treatment of periodontal disease during pregnancy associated with a reduced rate of preterm births?

92 Are longlasting nets treated with insecticide effective in reducing incidence of visceral leishmaniasis in India and Nepal?

93 How well does a new QRISK model estimate lifetime risk of cardiovascular disease?

94 Can simple ultrasound rules be used as a triage test to predict benignity or malignancy in an adnexal mass?

Evaluating adnexal masses



BSIP, VEM/SPL

Whether an adnexal mass is malignant or not affects the choice of treatment, but making this decision requires an experienced ultrasound examiner. Dirk Timmerman and colleagues have tested a decision making tool that should make this process a bit easier for less experienced ultrasound examiners (p 94).

The rules they devised comprised five ultrasonic features (including shape, size, solidity, and results of colour Doppler examination) to predict a malignant tumour (M features) and five to predict a benign tumour (B features). If one or more M features were present in the absence of a B feature, the mass was classified as malignant, whereas if one or more B features were present in the absence of an M feature, it was classified as benign. Findings using these rules were then compared with whether the excised adnexal mass was classified as benign or malignant on histological assessment.

The rules correctly discriminated between benign and malignant masses in 77% of cases. When conclusive, the rules performed as well as subjective assessment by an experienced examiner. The authors suggest that their ultrasound rules should be used as a triage test and subjective assessment by an experienced ultrasound examiner should be used as a second stage test in those masses for which the simple rules yield an inconclusive result.

Periodontal disease, pregnancy, and preemies

Can a woman's oral hygiene affect her unborn baby? All sorts of research on the effect of periodontal disease on preterm birth have been published in the past 16 years, with observational studies suggesting that pregnant patients with periodontal disease have an increased risk of preterm birth, but the largest randomised study to date said that treating periodontal disease had no effect on the rate of preterm birth and low birthweight infants.

Now Nikolaos Polyzos and colleagues have published a meta-analysis of studies on treatment of periodontal disease during pregnancy, which they hope will put the issue to rest (p 91). Only five of the 11 trials they identified were of sufficiently high methodological quality, and these trials provided clear evidence that treatment of periodontal disease has no effect on the rate of preterm births, low birthweight infants, spontaneous abortions and stillbirths, or overall adverse pregnancy outcomes. The six low quality trials, however, indicated that treatment did have a beneficial effect.

Writing in a linked editorial (p 59), George Macones says: "Disappointingly, despite years of basic, clinical, and translational research, no robust data support the treatment of any infection to reduce preterm birth or improve pregnancy outcomes . . . It may be time to re-examine some basic assumptions about the cause of adverse pregnancy outcomes and explore new mechanisms and treatments." However, both Professor Macones and the study authors emphasise that maintenance of oral health should be encouraged as part of routine preventive care, irrespective of its effect on pregnancy outcomes.

An algorithm for lifetime cardiovascular risk

Julia Hippisley-Cox and colleagues have developed QRISK, an algorithm for lifetime cardiovascular risk (p 93). Patients with high lifetime risk were more likely to be younger, male, from non-white ethnic groups, and have a positive family history of premature coronary heart disease than those identified as being at high risk with the 10 year QRISK2 score. But in younger patients, the risk-benefit ratio of medical interventions might be narrow, and the authors admit that we need more data on their cost effectiveness and acceptability.

Editorialists Rod Jackson and colleagues (p 62) aren't convinced that lifetime risk estimation adds much, but they like the graphs produced by the calculator at www.qrisk.org/lifetime. These predict patients' cumulative cardiovascular risk throughout their lifetime on the basis of both current risk profiles and what will happen if their risk profiles improve.

LATEST RESEARCH:

For this and other new research articles see http://www.bmj.com/channels/research.dtl

Efficacy of scorpion antivenom Being stung by the Indian red scorpion (*Mesobuthus tamulus*), the most lethal species of the Buthidae family in India, is bad news enough. But to make things worse, the efficacy of scorpion antivenom is unknown.

Himmatrao Saluba Bawaskar and Pramodini Himmatrao Bawaskar have conducted an open label randomised controlled trial in India to determine for sure the effects of the monospecific F(ab)2 scorpion antivenom on the cardiac pathophysiological effects of scorpion sting (doi:10.1136/bmj.c7136). They found that addition of scorpion antivenom to the α_1 blocker prazosin, which is widely used for the management of *M tamulus* sting, within six hours hastened recovery and shortened hospital stay compared with prazosin alone.

However, they point out that the total cost of treatment with antivenom approaches a month's salary for a labourer in rural India, whereas the cost of 10 tablets of 1 mg prazosin is a hundred times less.

AAMOD ZAMBRE

Obstetric outcomes after treatment of periodontal disease during pregnancy: systematic review and meta-analysis

Nikolaos P Polyzos,[1 2] Ilias P Polyzos,[3] Apostolos Zavos,[2] Antonis Valachis,[4] Davide Mauri,[5] Evangelos G Papanikolaou,[6] Spyridon Tzioras,[1] Daniel Weber,[7] Ioannis E Messinis[2]

EDITORIAL by Macones

[1]Section of Obstetrics and Gynaecology, Panhellenic Association for Continual Medical Research (PACMeR), Athens, Greece
[2]Department of Obstetrics and Gynaecology, University of Thessalia, Larisa, Greece
[3]Department of Oral and Maxillofacial Surgery, School of Dentistry, University of Manchester, Manchester, UK
[4]Section of Public Health, Panhellenic Association for Continual Medical Research (PACMeR)
[5]Panhellenic Association for Continual Medical Research (PACMeR)
[6]Biogenesis Medical Centre, Thessaloniki, Greece
[7]Department of Obstetrics and Gynecology, Lancaster General Health, Lancaster, PA, USA
Correspondence to: N P Polyzos
n.polyzos@gmail.com

Cite this as: *BMJ* 2010;341:c7017
doi: 10.1136/bmj.c7017

This is a summary of a paper that was published on bmj.com as *BMJ* 2010;341:c7017

STUDY QUESTION Is treatment of periodontal disease with scaling and root planing during pregnancy associated with a reduction in the rate of preterm birth?

SUMMARY ANSWER Treatment of periodontal disease with scaling and root planing has no significant effect on the incidence of preterm birth.

WHAT IS KNOWN AND WHAT THIS PAPER ADDS Periodontal disease is associated with increased risk of preterm birth, and a causal relation may exist; however, existing reports on the effect of treatment with scaling and root planning on the incidence of preterm birth are conflicting. Randomised trials of low methodological quality tend to overestimate the effect of treatment, whereas high quality trials provide strong evidence that no significant effect of treatment exists.

Selection criteria for studies

We searched the Cochrane Central Trials Registry, ISI Web of Science, and Medline without language restriction to July 2010. In addition, we reviewed the references of all eligible trials and hand searched the last two year volumes of two key dentistry journals. We included only randomised controlled trials of pregnant women with documented periodontal disease. Patients were randomised to either treatment with scaling and root planing or no treatment.

Primary outcome(s)

We focused on the overall rate of preterm births (<37 weeks of gestation). Additionally, we assessed the incidence of low birthweight infants (<2500 g) and spontaneous abortions/stillbirths and the overall rate of adverse outcomes of pregnancy (preterm birth <37 weeks and spontaneous abortions/stillbirths).

Main results and role of chance

Eleven studies (with 6558 women) met our inclusion criteria. We considered five trials to be of high methodological quality (low risk of bias), whereas the rest were low quality (high or unclear risk of bias). Results among low quality and high quality trials were consistently diverse; low quality trials supported a beneficial effect of treatment, and high quality trials provided clear evidence that no such effect exists. Among high quality studies, we found no significant effect of treatment on the overall rate of preterm birth (odds ratio 1.15, 95% confidence interval 0.95 to 1.40; P=0.15). Furthermore, treatment did not reduce the rate of low birthweight infants (odds ratio 1.07, 0.85 to 1.36; P=0.55), spontaneous abortions/stillbirths (0.79, 0.51 to 1.22; P=0.28), or overall adverse pregnancy outcomes (preterm births <37 weeks, spontaneous abortions/stillbirths) (1.09, 0.91 to 1.30; P=0.34).

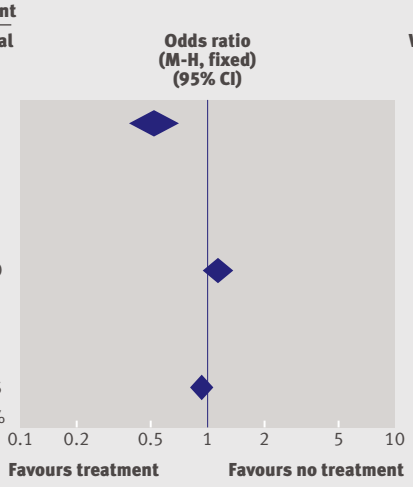
Bias, confounding, and other reasons for caution

A contour enhanced funnel plot showed that the chance of publication bias was high, and Harbord's test suggested that this meta-analysis might be affected by small study effect bias. The definition of periodontal disease differed among eligible trials. However, on the basis of trials' inclusion criteria, the five high quality trials had comparable inclusion criteria as regards the severity of disease

Study funding/potential competing interests

This research received no specific funding from any funding agency in the public, commercial, or non-profit sectors.

META-ANALYSIS PLOT FOR PRETERM BIRTH <37 WEEKS OF GESTATION



M-H=Mantel-Haenszel fixed effects model

Longlasting insecticidal nets for prevention of *Leishmania donovani* infection in India and Nepal: paired cluster randomised trial

Albert Picado,[1] Shri Prakash Singh,[2] Suman Rijal,[3] Shyam Sundar,[2] Bart Ostyn,[5] François Chappuis,[4] Surendra Uranw,[3] Kamlesh Gidwani,[2] Basudha Khanal,[3] Madhukar Rai,[2] Ishwari Sharma Paudel,[3] Murari Lal Das,[3] Rajiv Kumar,[2] Pankaj Srivastava,[2] Jean Claude Dujardin,[5] Veerle Vanlerberghe,[5] Elisabeth Wreford Andersen,[1] Clive Richard Davies,[1] Marleen Boelaert[5]

EDITORIAL by Desjeux

[1]London School of Hygiene and Tropical Medicine, London WC1E 7HT, UK
[2]Banaras Hindu University, Varanasi 221005, India
[3]BP Koirala Institute of Health Sciences, PO Box 7053, Kathmandu, Nepal
[4]Geneva University Hospitals, Geneva, Switzerland
[5]Institute of Tropical Medicine, 2000 Antwerp, Belgium
Correspondence to: M Boelaert
mboelaert@itg.be

Cite this as: *BMJ* 2010;341:c6760
doi: 10.1136/bmj.c6760

This is a summary of a paper that was published on bmj.com as *BMJ* 2010;341:c6760

STUDY QUESTION Are longlasting nets treated with insecticide (deltamethrin) effective in reducing incidence of visceral leishmaniasis (kala-azar) in India and Nepal?

SUMMARY ANSWER Village-wide distribution of longlasting insecticidal nets in India and Nepal did not confer additional protection against *Leishmania donovani* infection or disease compared with existing preventive practice (irregular insecticide residual spraying and use of untreated nets).

WHAT IS KNOWN AND WHAT THIS PAPER ADDS Observational evidence in East Africa and Asia suggested that people sleeping under nets were protected against visceral leishmaniasis, a potentially fatal parasitic disease transmitted by sandflies. This cluster randomised controlled trial showed that longlasting insecticidal nets were not effective in reducing the incidence of *L donovani* infection and visceral leishmaniasis.

Design

A paired cluster randomised controlled trial designed to detect a 50% reduction in incidence of *L donovani* infection compared 13 intervention clusters (treated nets) with 13 control clusters (no treated nets). In each country, clusters were paired on the basis of past incidence of visceral leishmaniasis. Block randomisation was used to allocate clusters to two groups. The intervention was then randomly allocated to one of the groups by tossing a coin. In December 2006, longlasting insecticidal nets (PermaNet 2.0) were distributed in the intervention clusters. The control clusters were allowed to continue using untreated nets. Field workers monitored individual use of nets and official spraying activities.

Participants and setting

The 26 study clusters were located in Muzaffarpur district in India (n=16) and Saptari, Sunsari, and Morang districts in Nepal (n=10). All individuals who lived for at least six months a year in the clusters were eligible, but blood sampling was restricted to those aged above 2 years.

Primary outcomes

Infection was determined by direct agglutination tests at 12 and 24 months after the intervention in those who had negative results (titre <1:1600) at baseline. Incident cases of visceral leishmaniasis and malaria were assessed during quarterly house to house surveys.

Main results and the role of chance

Out of 15 504 eligible people, 12 691 (82%) were included in the analysis of the main outcome. There was no significant difference in the risk of seroconversion over 24 months in intervention (5.4%; 347/6372) compared with control clusters (5.5%; 345/6319) (risk ratio 0.90, 95% confidence interval 0.49 to 1.65) or in clinical visceral leishmaniasis (0.99, 0.46 to 1.40). Adjustment for covariates did not alter these conclusions. There was a significant reduction of malaria in intervention clusters (0.46, 0.28 to 0.77) in the adjusted model (table).

Harms No harms were reported.

Bias, confounding, and other reasons for caution

The risk of *L donovani* infection was 10% less in intervention clusters, a small and not significant effect, though its wide confidence interval does not rule out a potential beneficial effect. When we accounted for the use of untreated nets and spraying the conclusions were the same. A substantial amount of *L donovani* transmission probably occurs outside the house, where presumably treated nets would have less impact on preventing sandfly-human contact.

Generalisability to other populations

As part of the ongoing initiative to eliminate visceral leishmaniasis in South Asia, distribution of insecticidal nets might not have a major impact on incidence in areas similar to those studied. In other areas, such as Iran and Sudan, the vector is different so our results might not be applicable.

Study funding/potential competing interests:
This study was funded by the European Commission (contract No INCO-CT 2005-01537, KALANET project).

Trial registration number Clinical Trials NCT 2005-015374.

EFFECT OF LONGLASTING INSECTICIDAL NETS ON VISCERAL LEISHMANIASIS IN INDIA AND NEPAL

Variable	Treated nets	Control	Risk ratio (95% CI), P value Unadjusted	Adjusted*
Primary outcome				
No in serology study	6372 (13 clusters)	6319 (13 clusters)	—	—
No of seroconversions (%)	347 (5.4)	345 (5.5)	0.90 (0.49 to 1.65), 0.71	0.89 (0.48 to 1.64), 0.68
Secondary outcomes				
Total No of participants	9829 (13 clusters)	9981 (13 clusters)	—	—
Visceral leishmaniasis (%)	37 (0.38)	40 (0.40)	0.99 (0.46 to 2.16), 0.99	1.15 (0.61 to 2.16), 0.64
Malaria (%)	88 (0.90)	137 (1.37)	0.63 (0.29 to 1.36), 0.21	0.46 (0.28 to 0.77), 0.01

*Simultaneously adjusted for age group, sex, times sprayed, and socioeconomic status.

Derivation, validation, and evaluation of a new QRISK model to estimate lifetime risk of cardiovascular disease: cohort study using QResearch database

Julia Hippisley-Cox,[1] Carol Coupland,[1] John Robson,[2] Peter Brindle[3]

EDITORIAL by Jackson, Kerr, and Wells

[1]Division of Primary Care, University Park, Nottingham NG2 7RD, UK
[2]Centre for Health Sciences, Queen Mary's School of Medicine and Dentistry, London E1 2AT, UK
[3]Avon Primary Care Research Collaborative, NHS Bristol, South Plaza, Bristol BS1 3NX, UK
Correspondence to: Julia Hippisley-Cox
juliahippisleycox@gmail.com

Cite this as: *BMJ* 2010;341:c6624
doi: 10.1136/bmj.c6624

This is a summary of a paper that was published on bmj.com as *BMJ* 2010;341:c6624

STUDY QUESTION How well does a new QRISK model to estimate lifetime risk of cardiovascular disease work?

SUMMARY ANSWER Compared with the standard QRISK2 model of 10 year risk, the lifetime approach identified younger patients who were more likely to be men, from non-white ethnic groups, and more likely to have a family history of early coronary heart disease and who would not otherwise be identified.

WHAT IS KNOWN AND WHAT THIS PAPER ADDS Risk prediction algorithms that estimate 10 year risk of cardiovascular disease to identify patients at high risk may miss younger people who, despite a low 10 year risk, have a high risk relative to their peers. A lifetime approach identifies such patients, but the net potential benefit of such an approach needs careful consideration.

Design, size, and duration

This prospective cohort study used routinely collected data from 563 general practices contributing to the QResearch database between 1994 and 2010. We included patients aged 30–84 who were free of cardiovascular disease and not taking statins—2 343 759 in the derivation dataset and 1 267 159 in the validation dataset. Cox proportional hazards models in the derivation cohort were used to derive risk equations for cardiovascular disease, accounting for non-cardiovascular death as a competing risk. Measures of calibration and discrimination were used to assess the model's performance in the validation cohort. The outcomes were individual estimates of lifetime risk of cardiovascular disease accounting for smoking status, ethnic group, systolic blood pressure, blood cholesterol, body mass index, family history of premature coronary heart disease, deprivation, treated hypertension, rheumatoid arthritis, chronic renal disease, diabetes, and atrial fibrillation.

CHARACTERISTICS OF 10% OF PATIENTS AT HIGHEST RISK OF CARDIOVASCULAR DISEASE BASED ON 10 YEAR RISK AND ON LIFETIME RISK

	Men		Women	
Predictor variables	10 year risk (n=69 794)	Lifetime risk (n=93 426)	10 year risk (n=56 921)	Lifetime risk (n=33 290)
Mean (SD) age (years)	71.7 (7.6)	45.2 (10.6)	76.7 (6.0)	50.1 (11.7)
Mean (SD) Townsend score	−0.1 (3.5)	−0.5 (3.4)	0.3(3.5)	0.5 (3.5)
No (%) of current smokers	20 300 (29.1)	30 466 (32.6)	11 845 (20.8)	10 639 (32.0)
No (%) in age band (years):				
30–44	178 (0.3)	49 009 (52.5)	76 (0.1)	11 593 (34.8)
45–64	11 523 (16.5)	39 570 (42.4)	2 376 (4.2)	17 490 (52.5)
65–74	30 753 (44.1)	4 469 (4.8)	13 728 (24.1)	3 725 (11.2)
75–84	27 340 (39.2)	378 (0.4)	40 741 (71.6)	482 (1.4)
No (%) in ethnic group:				
White or not recorded	68 652 (98.4)	83 411 (89.3)	56 326 (99.0)	29 054 (87.3)
Non-white	1 142 (1.6)	10 015 (10.7)	595 (1.0)	4 236 (12.7)
No (%) with clinical condition:				
Treated hypertension	14 196 (20.3)	10 066 (10.8)	14 833 (26.1)	7 865 (23.6)
Type 2 diabetes	6 907 (9.9)	3 425 (3.7)	5 369 (9.4)	1 784 (5.4)
Family history of early heart disease*	9 037 (12.9)	53 600 (57.4)	6 005 (10.5)	27 601 (82.9)
Atrial fibrillation	2 712 (3.9)	792 (0.8)	2 327 (4.1)	381 (1.1)
Chronic renal disease	417 (0.6)	155 (0.2)	270 (0.5)	91 (0.3)
Mean (SD) BMI (kg/m^2)	26.8 (20.8)	28.5 (4.5)	26.5 (4.7)	29.4 (5.4)
Mean (SD) total:HDL cholesterol ratio	4.7 (1.5)	5.6 (1.6)	4.4 (1.4)	5.1 (1.4)
Mean (SD) systolic BP (mm Hg)	150 (21)	138 (20)	154 (22)	144 (22)

*Coronary heart disease in a first degree relative aged <60 years.
BMI=body mass index. HDL=high density lipoprotein. BP=blood pressure.

Main results and the role of chance

Comparison of the predicted and observed lifetime risks of cardiovascular disease in the validation cohort showed a slight under-prediction in those at low risk (ratio of predicted to observed lifetime risk in the lowest 10th of risk was 0.90 for men and 0.82 for women) but good calibration in the highest 10th of risk (predicted:observed ratio 1.01 for men and 1.02 for women).

The 90th centile values identified a lifetime risk of cardiovascular disease of 50% in the validation cohort. Of the 10% of patients in this cohort classified at highest risk with either the lifetime risk model or the 10 year risk model, only 14.5% were at high risk on both measures. Patients identified as having high lifetime risk were more likely to be younger, male, from non-white ethnic groups, and have a positive family history of premature coronary heart disease than those identified with the 10 year QRISK2 score (see table). See www.qrisk.org/lifetime for the QRISK Lifetime Risk calculator.

Bias, confounding, and other reasons for caution

Observational studies can be subject to bias and unmeasured confounding, including the effects of missing data.

Generalisability to other populations

Our study is likely to be generalisable since it was carried out in a large primary care population representative of where most patients in the UK are assessed.

Study funding/potential competing interests

The study had no external funding. JHC and CC work at the University of Nottingham and hold positions at ClinRisk, which produces software for clinical risk algorithms. JHC is also codirector of QResearch—a joint partnership between University of Nottingham and EMIS, a commercial software provider for primary care.

Simple ultrasound rules to distinguish between benign and malignant adnexal masses before surgery: prospective validation by IOTA group

Dirk Timmerman,[1] Lieveke Ameye,[2] Daniela Fischerova,[3] Elisabeth Epstein,[4] Gian Benedetto Melis,[5] Stefano Guerriero,[5] Caroline Van Holsbeke,[6] Luca Savelli,[7] Robert Fruscio,[8] Andrea Alberto Lissoni,[8] Antonia Carla Testa,[9] Joan Veldman,[1] Ignace Vergote,[1] Sabine Van Huffel,[2] Tom Bourne,[10 1] Lil Valentin[11]

[1]Department of Obstetrics and Gynaecology, University Hospitals KU Leuven, 3000 Leuven, Belgium
[2]Department of Electrical Engineering, Katholieke Universiteit Leuven, Belgium
[3]Oncogynaecological Center, First Faculty of Medicine and General University Hospital of Charles University, Prague, Czech Republic
[4]Department of Obstetrics and Gynaecology, Lund University, Lund, Sweden
[5]Department of Obstetrics and Gynaecology, University of Cagliari, Ospedale San Giovanni di Dio, Cagliari, Sardinia, Italy
[6]Department of Obstetrics and Gynaecology, Ziekenhuis Oost-Limburg, Genk, Belgium
[7]Reproductive Medicine Unit, Department of Obstetrics and Gynaecology, University of Bologna, Italy
[8]Clinica Ostetrica e Ginecologica, Ospedale S Gerardo, Università di Milano Bicocca, Monza, Italy
[9]Istituto di Clinica Ostetrica e Ginecologica, Università Cattolica di Sacro Cuore, Roma, Italy
[10]Department of Obstetrics and Gynaecology, Imperial College London, Queen Charlotte's and Chelsea Hospital, London, UK
[11]Department of Obstetrics and Gynaecology, Malmö University Hospital, Lund University, SE20502 Malmö, Sweden

Correspondence to: D Timmerman
dirk.timmerman@uzleuven.be

Cite this as: *BMJ* 2010;341:c6839
doi: 10.1136/bmj.c6839

This is a summary of a paper that was published on bmj.com as *BMJ* 2010;341:c6839

STUDY QUESTION Can simple ultrasound rules be used as a triage test to predict benignity/malignancy in an adnexal mass?

SUMMARY ANSWER Simple rules yielded a conclusive result (benign or malignant) in about 75% of adnexal masses, and when conclusive, or if used as a triage test, they performed as well as subjective assessment by an experienced ultrasound examiner.

WHAT IS KNOWN AND WHAT THIS PAPER ADDS Subjective assessment of ultrasound examination is the most reliable method to distinguish between benign and malignant adnexal masses before surgery, but it requires expertise. The use of the simple rules can change clinical practice by providing an accurate and instant classification of most adnexal masses while reducing the number of patients that need to be referred for expert scanning.

SIMPLE ULTRASOUND RULES

Ultrasonic features for predicting a malignant tumour (M)
M1—Irregular solid tumour
M2—Presence of ascites
M3—At least four papillary structures
M4—Irregular multilocular solid tumour with largest diameter ≥100 mm
M5—Very strong blood flow (colour score 4)

Ultrasonic features for predicting a benign tumour (B)
B1—Unilocular
B2—Presence of solid components, of which largest solid component has largest diameter <7 mm
B3—Presence of acoustic shadows
B4—Smooth multilocular tumour with largest diameter <100 mm
B5—No blood flow (colour score 1)

Rules
Rule 1—If one or more M features are present in absence of B feature, mass is classified as malignant.
Rule 2—If one or more B features are present in absence of M feature, mass is classified as benign.
Rule 3—If both M features and B features are present, or if no B or M features are present, result is inconclusive and second stage test is recommended

Participants and setting

This was an international multicentre study in 19 ultrasound centres in eight countries. The principal investigator at each centre examined women with an adnexal mass with ultrasound by using a standardised research protocol.

Design, size, and duration

We did a prospective temporal and external validation of the simple ultrasound rules (see box) to distinguish benign from malignant adnexal masses in 1938 women with an adnexal mass between November 2005 and October 2007. In adnexal masses for which the simple rules yielded an inconclusive result, we compared the performance of expert subjective assessment with the risk of malignancy index and two logistic regression models developed by the International Ovarian Tumor Analysis Group. The reference standard was the histological diagnosis and, in case of malignancy, the surgical stage.

Main results and the role of chance

The simple ultrasound rules yielded a conclusive result in 1501 (77%) masses, for which they resulted in a sensitivity of 92% (95% confidence interval 89% to 94%) and a specificity of 96% (94% to 97%). The corresponding sensitivity and specificity of subjective assessment were 91% (88% to 94%) and 96% (94% to 97%). In the 357 masses for which the simple rules yielded an inconclusive result and with available results of CA 125 measurements, the sensitivities were 89% (83% to 93%) for subjective assessment, 50% (42% to 58%) for the risk of malignancy index, 89% (83% to 93%) for logistic regression model 1, and 82% (75% to 87%) for logistic regression model 2; the corresponding specificities were 78% (72% to 83%), 84% (78% to 88%), 44% (38% to 51%), and 48% (42% to 55%). This means that in adnexal masses for which the rules yielded an inconclusive result, subjective evaluation of ultrasound findings by an experienced ultrasound examiner was the most accurate diagnostic test, whereas the risk of malignancy index and logistic regression models were not useful. Using the simple rules as a triage test and subjective assessment for those masses for which the simple rules yielded an inconclusive result, we obtained a sensitivity of 91% (88% to 93%) and a specificity of 93% (91% to 94%). This is the same diagnostic performance as obtained when using subjective assessment in all masses—sensitivity 90% (88% to 93%) and specificity 93% (91% to 94%).

Bias, confounding, and other reasons for caution

A limitation of the study is that all the examinations were done by experienced ultrasound examiners.

Generalisability to other populations

This was a prospective temporal and external validation study of the simple ultrasound rules in a large number of patients with adnexal masses, indicating that the rules are generalisable.

Study funding/potential competing interests

The study had no commercial funding.