Economic evaluation of an expert examiner and different ultrasound models in the diagnosis of ovarian cancer

European Journal of Cancer 100 (2018) 55–64

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejcancer.com

Original Research

Economic evaluation of an expert examiner and different ultrasound models in the diagnosis of ovarian cancer

Evelyne M.J. Meys [a,b], Lara S. Jeelof [a], Bram L.T. Ramaekers [c,d], Carmen D. Dirksen [c,d], Loes F.S. Kooreman [b,e], Brigitte F.M. Slangen [a,b], Roy F.P.M. Kruitwagen [a,b], Toon Van Gorp [a,b,*]

[a] *Department of Obstetrics and Gynaecology, Maastricht University Medical Center+ (MUMC+), Maastricht, The Netherlands*
[b] *GROW: School for Oncology and Developmental Biology, P. Debyelaan 25, 6202 AZ, Maastricht, The Netherlands*
[c] *Department of Clinical Epidemiology and Medical Technology Assessment (KEMTA), Maastricht University Medical Center+ (MUMC+), Maastricht, The Netherlands*
[d] *CAPHRI, Care and Public Health Research Institute, PO Box 616, 6200 MD Maastricht, The Netherlands*
[e] *Department of Pathology, Maastricht University Medical Center+ (MUMC+), Maastricht, The Netherlands*

Received 30 December 2017; received in revised form 28 April 2018; accepted 11 May 2018
Available online 26 June 2018

KEYWORDS
Cost-effectiveness analysis;
Budget impact analysis;
Diagnosis;
Ovarian cancer;
Ultrasound;
RMI;
subjective assessment;
Simple ultrasound-based rules;
LR2 model;
ADNEX model

Abstract The Risk of Malignancy Index (RMI) is commonly used to diagnose adnexal masses. The aim of the present study was to determine the cost-effectiveness of the RMI compared with subjective assessment (SA) by an expert and the following novel ultrasound models:

- Simple rules (SR) added by SA (SR + SA);
- SR with inconclusive results diagnosed as malignant (SR + Mal);
- Logistic Regression model 2 (LR2); and
- Assessment of Different NEoplasias in the adneXa (ADNEX) model.

Cost-effectiveness and budget impact analyses were performed from a societal perspective. A decision tree was constructed, and short-term costs and effects were examined in women with adnexal masses. Sensitivity, specificity and the costs of diagnostic strategies were incorporated. Incremental cost-effectiveness ratios were expressed as costs/additional percentage of correctly diagnosed patients. Probabilistic and deterministic sensitivity analyses were performed.

Effectiveness was highest for SA (90.7% [95% confidence interval = 77.3–100]), with a cost saving of 5.0% (−€398 per patient [−€1403 to 549]) compared with the RMI. The costs of SR + SA were the lowest (€7180 [6072–8436]), resulting in a cost saving of 9.0% (−€709 per patient [−€1628 to 236]) compared with the RMI, with an effectiveness of 89.6% (75.8–100).

* *Corresponding author*: Maastricht University Medical Center+ (MUMC+), P. Debyelaan 25, 6202 AZ, Maastricht, The Netherlands. Fax: +31 43 387 4765.
E-mail address: Toon.van.gorp@mumc.nl (T. Van Gorp).

https://doi.org/10.1016/j.ejca.2018.05.003

SR + SA showed the highest probability of being the most cost-effective when willingness-to-pay was <€350 per additional percentage of correctly diagnosed patients. The RMI had low cost-effectiveness probabilities (<3%) and was inferior to SA, SR + SA and LR2. Budget impact in the Netherlands compared with that of the RMI varied between a cost saving of €4.67 million for SR + SA and additional costs of €3.83 million when implementing ADNEX (cut-off: 10%). The results were robust when tested in sensitivity analyses.

Although SA is the best strategy in terms of diagnostic accuracy, SR + SA might be preferred from a cost-effectiveness perspective.

1. Introduction

Adnexal masses occur frequently, and although they are mostly benign, approximately 22,000 patients are diagnosed with ovarian carcinoma in the United States each year [1]. In patients diagnosed with an adnexal mass, preoperative assessment by ultrasound is valuable for determining the optimal treatment. A high diagnostic accuracy of the method used to differentiate benign from malignant adnexal masses is necessary for optimal treatment planning. Low sensitivity leads to the misclassification of malignancies. These patients undergo a laparoscopy, often followed by a second operation to achieve full staging/debulking, resulting in a higher cost. Low specificity, on the other hand, causes patients to undergo unnecessary extensive surgery with prolonged rehabilitation.

Many current guidelines endorse the Risk of Malignancy Index (RMI) to distinguish benign from malignant tumours [2]. Although the RMI is easy to use, it has low sensitivity compared with other models [3,4]. Despite the development of various scoring systems, subjective assessment (SA) by an expert ultrasound examiner remains the best method currently available [3]. However, experts are scarce and expensive. Furthermore, when patients have to pay for an extra visit to an expert for ultrasound, the costs increase for the patient and often the employer. Whether the potential cost savings associated with a higher diagnostic accuracy, and thus an increase in correct diagnoses and treatments, can offset the increase in the cost of using experts remains to be investigated.

Other methods, such as simple ultrasound-based rules (simple rules, SR), Assessment of Different NEoplasias in the adneXa (ADNEX model) and Logistic Regression model 2 (LR2) have no additional operating costs compared with the RMI and have shown a higher diagnostic accuracy [3–8].

A cost-effectiveness analysis (CEA) is of utmost importance as a pivotal step to determine whether or not to abandon the RMI and implement other methods.

The aim of the present study was to determine the short-term cost-effectiveness and budget impact of the RMI compared with SA in the diagnosis of adnexal masses. The RMI was also compared with recently introduced ultrasound models for the diagnosis of adnexal masses, including SR added by SA (SR + SA), SR with inconclusive results diagnosed as malignant (SR + Mal), LR2 and the ADNEX model.

2. Materials and methods

2.1. Scope of the economic evaluation

Short-term CEA and budget impact analysis were performed by developing a decision tree using internationally accepted guidelines [9–11]. A detailed description of the methods (based on the Consolidated Health Economic Evaluation Reporting Standards checklist) is provided in Supplementary File S1 [11].

The population consisted of women at least 18 years of age who presented to the gynaecology department of a hospital in the Netherlands in 2014 and were diagnosed with an adnexal mass requiring surgery. Epidemiologic data were obtained from the nationwide network and registry of histopathology and cytopathology in the Netherlands (PALGA, http://www.palgaopenbaredatabank.nl/), which registers pathology reports generated by all pathology departments nationwide [12]. Reports on all adnexal masses excised in 2014 (benign, borderline and malignant) were evaluated. Borderline tumours were considered to be malignant.

The time horizon of the analysis was from the time of detection of the mass until the time of recovery after the appropriate surgical intervention. Most of the consequences of misdiagnosis by ultrasound occur during this time period. Discounts for future costs were not applied because of the short time frame of less than 1 year. The analysis was performed from a societal perspective.

The costs of six diagnostic strategies were compared. These included the reference strategy (RMI) and five promising methods with high diagnostic accuracy that have the potential to replace the RMI, namely, SA, International Ovarian Tumor Analysis Group (IOTA) SR as a first step, SR + SA, SR + Mal, ADNEX and the LR2 model [5–7,13]. The ADNEX and LR2 models

Table 1
Model parameters and distribution.

Parameters	Estimate	SE (%) *upper/lower*	Distribution	Source
Epidemiology				
Prevalence	28%	2.4	Beta	[4]
Population	6493	*5926/7537*	Beta PERT	[PALGA]
Sensitivity RMI	71%	8.7	Beta	[7]
Specificity RMI	81%	5.6	Beta	[7]
Sensitivity SA	90%	6.1	Beta	[7]
Specificity SA	91%	4.1	Beta	[7]
Sensitivity SR + SA	89%	6.6	Beta	[7]
Specificity SR + SA	90%	4.6	Beta	[7]
Sensitivity SR + Mal	93%	5.6	Beta	[7]
Specificity SR + Mal	68%	4.6	Beta	[7]
Sensitivity ADNEX	98%	3.6	Beta	[7]
Specificity ADNEX	62%	6.6	Beta	[7]
Sensitivity LR2	93%	5.6	Beta	[7]
Specificity LR2	79%	5.6	Beta	[7]
Medical cost				
Outpatient clinic visit	In confidence	–	Beta PERT	[*]
CA125	In confidence	–	Beta PERT	[*]
Adnexal ultrasound	In confidence	–	Beta PERT	[*]
Laparoscopy	In confidence	–	Beta PERT	[*]
Laparotomy	In confidence	–	Beta PERT	[*]
Indirect cost				
Travelling expenses (per km)	€ 0.19	–	Fixed	[14]
Parking cost	€ 3.00	–	Fixed	[14]
Time investment companion (per hour)	€ 14.00	–	Fixed	[14]
Absenteeism cost				
Outpatient clinic	€ 94.8	–	Fixed	[12, 13]
Recovery laparoscopy	€ 2401.60	–	Fixed	[12, 13]
Recovery laparotomy benign	€ 6004.00			
Recovery laparotomy malignant	€ 8405.60	–	Fixed	[12, 13]

SE, standard Error; RMI, Risk of Malignancy Index; SA, subjective assessment; SR + SA, simple rules with subjective assessment for inconclusive test results; SR + Mal, simple rules for which inconclusive results were assumed to be malignant; ADNEX, Assessment of Different NEoplasias in the adneXa; LR2, logistic regression model 2; PALGA, 'Pathologisch-Anatomisch Landelijk Geautomatiseerd Archief' (Dutch pathological anatomical nationwide automated archive); PERT, program evaluation and review technique.
[*] Cost prices are obtained in confidence from the financial department of the Maastricht University Medical Center+, Maastricht, the Netherlands.

were evaluated using different cut-off values (i.e. 10%, 20%, 30% and 40%).

2.2. Outcomes

The sensitivity and specificity for the six strategies were combined with the prevalence to calculate the correct diagnoses and misdiagnoses in terms of true positives/negatives and false positives/negatives, respectively. The sensitivity and specificity of the various strategies considered are presented in Table 1 [8]. For the CEA, the incremental cost-effectiveness ratio (ICER) was expressed as the costs per additional percentage of correctly diagnosed patients. The budget impact analysis examines the expected expenditures/savings associated with the adoption of the various strategies compared with those of the RMI.

2.3. Resources and costs

A decision tree was constructed (Fig. 1) to evaluate the costs of each strategy using Excel 2010 software (Microsoft Corporation, Redmond, WA, USA). The selected model structure reflects the current treatment practices and efficacy of the RMI for the diagnosis of an adnexal mass and compares these with those of the other models (also see Supplement S1, item 13). In brief, patients were seen twice at the outpatient clinic and, if relevant, underwent additional testing (expert ultrasound or CA-125 measurement). The patients were then referred for surgery. All patients returned to the outpatient clinic for follow-up examinations. If the diagnostic strategy yielded false-negative test results, patients were scheduled for a second operation (staging or debulking surgery) and another follow-up visit at the outpatient clinic.

Three types of costs were considered: the costs from a hospital perspective (costs of outpatient clinic visits, additional testing such as CA-125 or expert ultrasound, surgery and hospitalisation after surgery), costs for the patient and her family (travel and parking expenses) and expenses in other sectors (absenteeism-related costs) [14–16]. All costs were reported for 2014 in Euros.

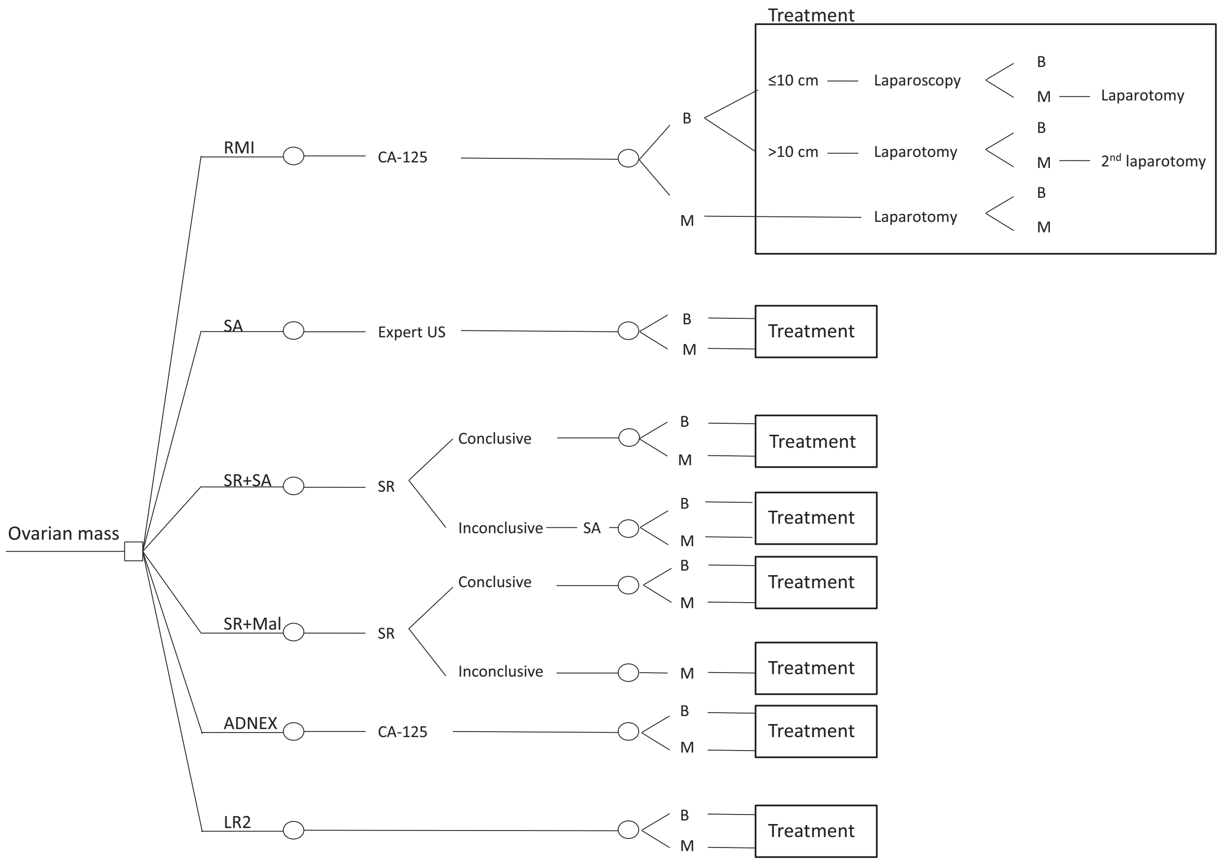

Fig. 1. Decision tree. RMI, Risk of Malignancy Index; SA, subjective assessment; SR + SA, simple ultrasound-based rules added by subjective assessment in case of inconclusive results; SR + Mal, simple ultrasound-based rules in which inconclusive results are diagnosed as malignant; ADNEX, Assessment of Different NEoplasias in the adneXa; LR2, Logistic Regression model 2; US, ultrasound; B, benign; M, malignant.

Dutch reference prices are available only for certain provisions at specific departments and do not include the subjects of the present study. Therefore, the hospital-related costs were determined by an internal cost-price investigation at the Maastricht University Medical Center+ (MUMC+). Start-up costs for the different strategies were excluded to ensure that the models were evaluated and compared under steady-state conditions.

Structural or other assumptions underpinning the decision-analytical model used can be found under item 16 in Supplement S1.

2.4. Analytical methods

A baseline analysis was performed. Absolute costs and effects (percentage of true positives and true negatives) were calculated for each strategy and compared with those of the RMI to determine incremental costs and effects.

Deterministic sensitivity analyses were performed for various parameters (sensitivity, specificity, prevalence, hospital costs, unit costs for the expert ultrasound and absenteeism-related costs). Probabilistic sensitivity analyses (PSAs) were performed using the Monte Carlo simulation (5000 iterations) to reflect parameter uncertainty [17]. The parameters, distribution and their respective sources (i.e. published literature and an internal cost-price investigation) are presented in Table 1 [3,8,14–16]. All analyses were performed using Excel (Microsoft Office 2010).

The ICER between the RMI, as the reference test, and each of the strategies was calculated based on the PSA. Cost and effectiveness were plotted for each strategy in a cost-effectiveness plane, and cost-effectiveness acceptability curves (CEACs) were drawn.

Threshold analyses were performed to identify optimal combinations of sensitivity and specificity in terms of cost-effectiveness. Optimal cut-off values were also estimated for ADNEX and LR2 in terms of cost-effectiveness.

3. Results

3.1. Epidemiology results

The study included 5175 reports on benign pathology and 1318 on malignant pathology from the PALGA database for the year 2014 (see also Supplementary File S1, item 18). All calculations were performed for this cohort of 6493 women. We found that 94.2% of benign masses were ≤10 cm, indicating treatment by laparoscopy, whereas laparotomy was indicated in 5.8% of benign tumours. Furthermore, 20.0% of women undergoing surgery for an adnexal mass were retired, resulting in lack of absenteeism-related costs in this population.

3.2. Costs and effectiveness

The mean costs for the RMI were €7889 (95% confidence interval [CI] = €6613–9288) (Table 2). The mean costs for SA were €7491 (95% CI = €6387–8705), which amounts to a cost saving of 5.0% (−€398 per patient [95% CI = −€1403 to 549]) compared with the RMI. The ADNEX model was associated with the highest costs (at a cut-off of 10%) because of the low specificity of 62%, which resulted in higher costs of surgery and increased absenteeism at work (more 'unnecessary' laparotomies with a prolonged recovery). At a cut-off of 40%, however, sensitivity and specificity were more balanced (85% [95% CI = 75.6–91.3] and 83% [95% CI = 77.5–87.5], respectively), resulting in a cost saving of −€281 (95% CI = −€1315 to 748). SR + SA was the least expensive method; it showed a sensitivity and specificity of 89% and 90%, respectively, and prevented incorrect diagnoses at a rate 11.4% (95% CI = −0.2 to 22.7) higher than that of the RMI, with a cost-saving effect of 9.0% (−€709 per patient [95% CI = −€1628 to 236]).

The RMI was dominated by SA, SR + SA and LR2 (more costly, less effective) to different extents (Fig. 2). CEACs showed that the SR + SA strategy had the highest probability of being the most cost-effective for a range of willingness-to-pay thresholds (Fig. 3). When the willingness-to-pay for an increase in the number of correct diagnoses was zero, SR + SA was the most cost-effective method, with a probability of 58.1% versus 26.0% for the LR2 (cut-off: 10%). This value decreased to 43.0% and 3.0%, respectively, at a threshold of €1000. At a willingness-to-pay >€350, the probability that SA was the most cost-effective method for the given data exceeded the possibility that SR + SA was the most cost-effective method. The RMI, SR + Mal and ADNEX (cut-off: 10%) showed low cost-effectiveness probabilities (all <3%, independent of the ceiling ratio).

The maximal achievable budget impact in the Netherlands compared with that of the current strategy varied between a cost saving of €4.67 million for SR + SA and additional costs of €3.83 million when implementing the ADNEX model (at a cut-off of 10%) (Table 2).

3.3. Sensitivity analyses

The mean costs based on the PSA were comparable to those of the base case analysis. Deterministic sensitivity analyses showed uncertainty about absenteeism-related costs, and hospital costs had the greatest impact on the cost-effectiveness outcome (Fig. 4). The ICER for all models was least sensitive to changes in the sensitivity of the diagnostic intervention. In general, the conclusions in terms of cost-effectiveness were similar when sensitivity analyses were performed on diagnostic test accuracy, prevalence, absenteeism-related costs and total hospital costs, as well as costs for the expert adnexal ultrasound.

3.4. Threshold analyses

For all interventions, the RMI was dominant when the sensitivity of the diagnostic intervention was <21% or

Table 2
Estimated cost and effectiveness for various ultrasound models for 2014 in the Netherlands compared to the RMI.

Triage strategy	Base case analysis		Expected outcomes (95% CI) based on PSA					
	Costs in €	Cost on population level (million €)	Mean cost in €	iCost in €	Effectiveness in %	iEffectiveness in %	Probability increased effectiveness	ICER
RMI	7890	51.23	7889 (6613–9288)	–	78.3 (63.4–92.3)	–	–	–
SA	7484	48.59	7491 (6387–8705)	-398 (-1403 to 549)	90.7 (77.3–100)	12.4 (10.2–13.9)	98.6%	Dominated
SR + SA	7170	46.56	7180 (6072–8436)	-709 (-1628 to 236)	89.6 (75.8–100)	11.4 (−0.2 to 22.7)	97.3%	Dominated
SR + Mal	8215	53.34	8214 (6954–9569)	325 (−603 to 1303)	75.1 (62.3–87.4)	−3.2 (−14.9 to 8.4)	29.4%	Dominant
ADNEX (10%)	8480	55.06	8482 (7062–9988)	594 (−495 to 1740)	72.1 (57.3–86.7)	−6.2 (−19.8 to 7.4)	17.9%	Dominant
ADNEX (40%)	7605	49.38	7608 (6658–8611)	−281 (−1315 to 748)	83.6 (75.9–91.3)	5.3 (−0.1 to 18.1)	87.2%	Dominated
LR2 (10%)	7600	49.35	7599 (6356–8928)	−290 (−1859 to 1237)	82.8 (68.4–96.5)	4.6 (−8.1 to 16.7)	76.6%	Dominated
LR2 (40%)	7274	47.23	7278 (6311–8224)	−611 (−1620 to 334)	87.4 (79.6–95.1)	9.1 (2.5–21.6)	98.8%	Dominated

95% CI, 95% confidence interval; RMI, Risk of Malignancy Index; SA, subjective assessment; SR + SA, simple rules with subjective assessment for inconclusive test results; SR + Mal, simple rules for which inconclusive results were assumed to be malignant; ADNEX, Assessment of Different NEoplasias in the adneXa (at a cut-off of 10% and 40%, resp.); LR2, logistic regression model 2 (at a cut-off of 10% and 40%, respectively); PSA, probabilistic sensitivity analysis; iCost, incremental cost when the intervention was compared to the control (RMI); iEffectiveness, incremental effectiveness when the intervention was compared to the control (RMI); ICER,: incremental cost-effectiveness ratio.
Effectiveness is defined as the percentage of cases correctly diagnosed by each method (true positives and true negatives based on sensitivity and specificity).

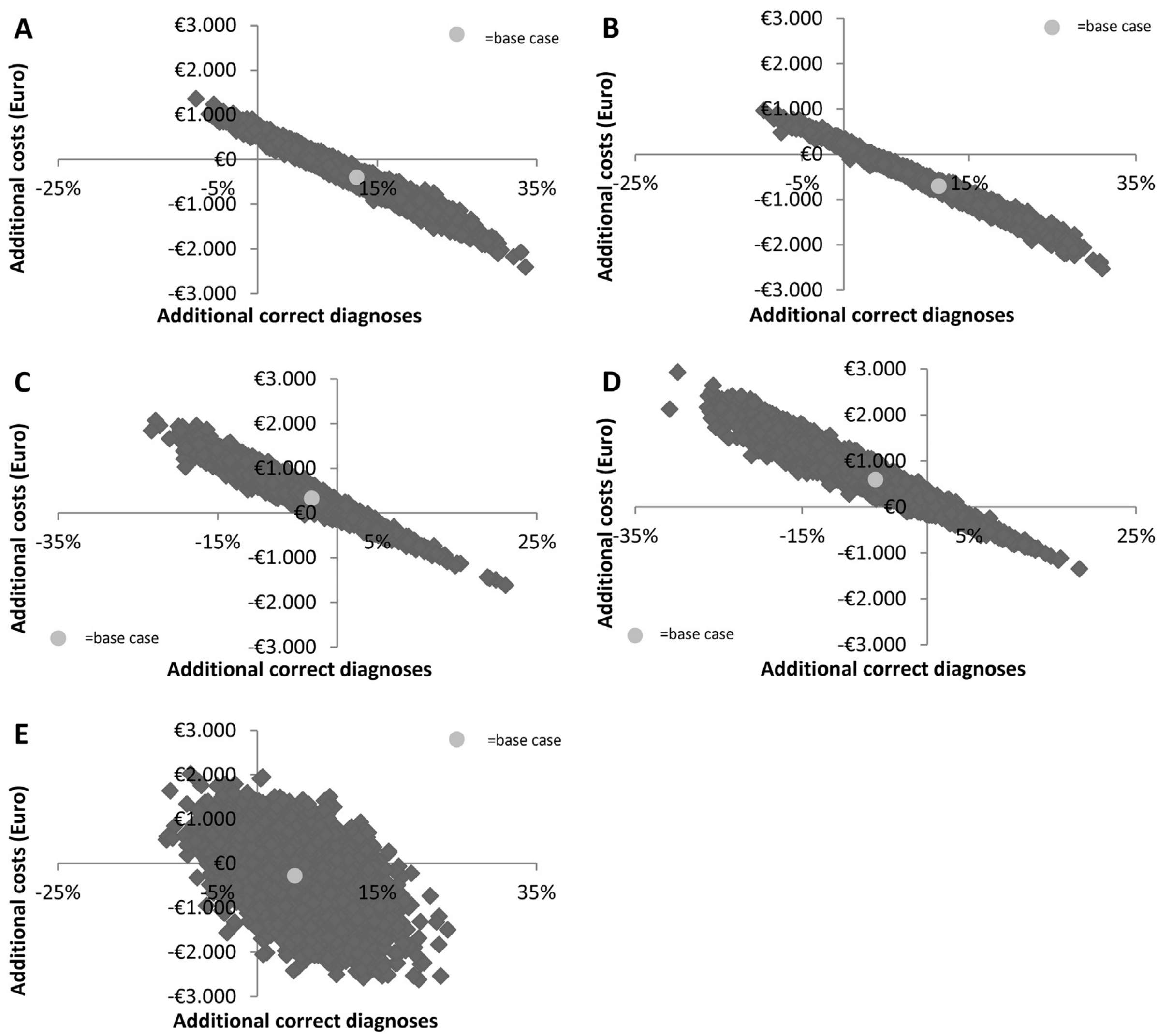

Fig. 2. Cost effectiveness plane (CE-plane) showing the incremental cost-effectiveness ratios (ICERs) when RMI was compared with (A) subjective assessment, (B) simple rules with subjective assessment for inconclusive test results, (C) simple rules for which inconclusive results were assumed to be malignant, (D) ADNEX model (at a cut-off of 10%) and (E) LR2 model (at a cut-off of 10%). The horizontal axis displays the additional correct diagnoses, indicated as percentage true positive plus true negative test results (ranging from −100% to 100%). The vertical axis of the plane depicts the additional costs in Euro (positive or negative). Each point in the CE-plane represents one of the 5000 ICER values derived from the Monte Carlo analysis. The ICER estimates of both SA and SA + SR are mainly located in the right quadrant of the CE-plane, with the majority in the lower right quadrant. This implicates that both methods are more effective than RMI, and probably yield a decrease in costs. RMI, risk of malignancy index; SA, subjective assessment; SR + SA, simple ultrasound-based rules added by subjective assessment in case of inconclusive results; ADNEX, Assessment of Different NEoplasias in the adneXa; LR2, logistic regression model 2.

the specificity was <70%, regardless of the willingness-to-pay threshold (Supplementary File S2, Fig. A–E). The RMI was inferior to both SA and SR + SA for a (balanced) combination of sensitivity and specificity of ≥80% and ≥78%, respectively.

The optimal cut-off of the ADNEX model in terms of cost-effectiveness was 65% (Supplementary File S3, Fig. A and Supplementary File S4, Fig. A–C), with an effect of 86.2% against a cost price of €7411. The RMI was inferior to the ADNEX model at a cut-off of ≥27% (Supplementary File S3, Fig. A). The optimal cut-off for the LR2 was 35% (Supplementary File S3, Fig. B and Supplementary File S4, Fig. D–F), with a total effect of 87.8% against a cost price of €7238. The RMI was inferior to the LR2 at cut-off values of ≥9% and <90%.

4. Discussion

4.1. Main findings

The present study analysed the cost-effectiveness and budget impact of the RMI for the diagnosis of ovarian

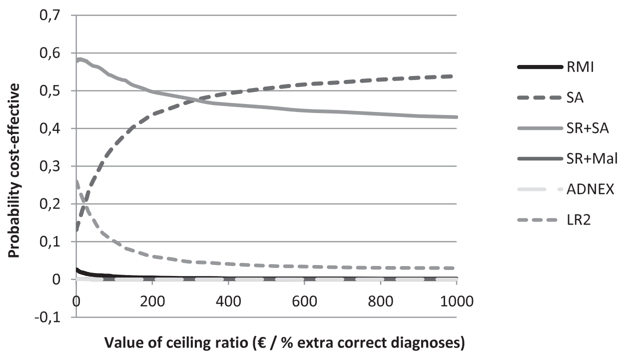

Fig. 3. Cost-effectiveness acceptability curves (CEACs) summarising evidence in support of the various ultrasound methods being cost-effective for all potential (relevant) threshold ratios decision-makers could use (representing the willingness-to-pay for a unit of health gain, in this case 1% increase in the amount of correct diagnoses). RMI, Risk of Malignancy Index; SA, subjective assessment; SR + SA, simple rules with subjective assessment for inconclusive test results; SR + Mal, simple rules for which inconclusive results were assumed to be malignant; ADNEX, Assessment of Different NEoplasias in the adneXa (at a cut-off of 10%); LR2, logistic regression model 2 (at a cut-off of 10%).

carcinoma in the Netherlands and compared these with those of SA and various recently introduced ultrasound models. SA and SR + SA were superior to the RMI in terms of cost-effectiveness, similar to the LR2 and ADNEX models at specific cut-off values ($\geq$9% and $\geq$27%, respectively). The implementation of SA in clinical practice would result in a cost saving of €2.64 million in the Netherlands (based on 6493 estimated adnexal masses requiring surgery annually) and prevent 12.4% of incorrect diagnoses.

However, the cost-saving effect of SR + SA was greater, with a reasonable probability that SR + SA is the most cost-effective method for a range of acceptable willingness-to-pay thresholds.

4.2. Interpretation

Potential future levels of uptake are core components of budget impact. There are usually limited data available on this item. We conducted a survey among gynaecologists to evaluate the use of ultrasound models in the Netherlands (unpublished data from Meys E, Jeelof LS, Lambrechts S, Slangen BFM, Kruitwagen RFPM, Van Gorp T, Evaluation of expert ultrasound in the diagnosis of adnexal masses by patients and gynaecologists: a feasibility study, *submitted*). The results showed that, despite Dutch guidelines recommending the use of the RMI, 26.5% of gynaecologists in the Netherlands currently do not use the RMI at all. Therefore, the cost-saving effect in clinical practice may be smaller than estimated. However, a change in the guidelines leading to the implementation of SA or SR + SA would mean a substantial cost saving (against higher efficacy) compared with the current situation.

Furthermore, the introduction of either of the intervention strategies would result in changes in disease outcome; if a malignant mass is not identified as such and is treated as a benign mass (false-negative result), laparoscopic surgery increases the risk of spillage of cyst fluid, which can negatively affect the prognosis of the patient [18]. On the other hand, laparoscopy has certain advantages over laparotomy such as decreased postoperative pain and recovery time and reduced adhesion formation [19,20]. Consequently, false-positive results have an impact on patient outcomes and the use of health care services, and these changes contribute substantially to cost reduction. Nonetheless, appropriate training of gynaecologists is necessary before implementing either method. The implementation of SA requires a certain level of expertise to achieve good diagnostic accuracy. Training experts is expensive and time consuming, and it is therefore nearly impossible to have an expert available in each centre. SR as a first step to triage patients requiring SA could be the solution to this problem. If only 19% of patients (with inconclusive results for SR) are referred to a centre with an expert in SA on staff, it would result in time and resource savings for both the expert ultrasound examiner and the patient. This improvement was reflected in the sensitivity analyses, which demonstrated that the unit price of the expert ultrasound had a marked effect on the ICER when SA was compared with the RMI, whereas the effects were smaller for SR + SA. However, a previous study showed that SR is easily misinterpreted when implemented by untrained examiners; therefore, training in both IOTA terminology and SR is necessary before SR can be introduced into guidelines and daily clinical practice [21]. Nonetheless, one could argue that a certain level of training in ultrasound and corresponding terminology is indispensable, regardless of the ultrasound method used. However, the ADNEX model only contains six ultrasound items, which are generally thought to be easier to interpret, and thus easier to teach, than the items used in SR. There are currently no studies to confirm this assumption.

Methods such as the ADNEX model or LR2 can also be used to triage patients with a difficult to diagnose mass who should receive SA; however, according to the CEACs, these methods showed poor chances of being the most cost-effective at a cut-off of 10% (Fig. 3). At higher cut-off values, these models perform much better and are more cost-effective than the RMI. Furthermore, although the ADNEX model is slightly more expensive, its added value lies in the multiclass nature of the model (i.e. the ability of the model to differentiate between benign, borderline, stage I, stage II–IV and metastatic tumours). This added value is difficult to express into costs and is therefore not included in our analysis. If this had been the case, the ADNEX model may have been more appealing than SA or SR + SA.

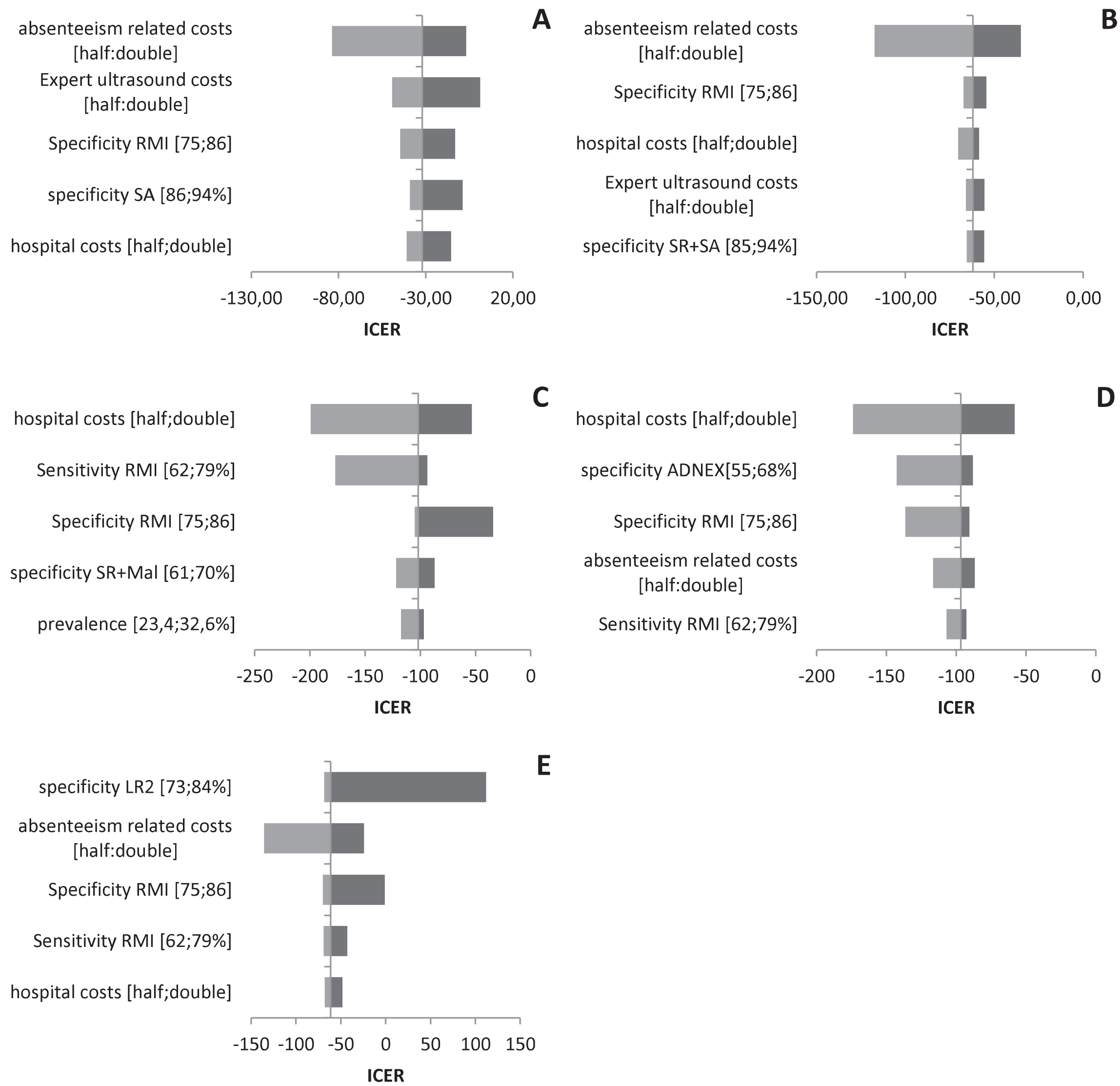

Fig. 4. Tornado plots summarising the effect of the one-way (deterministic) sensitivity analyses when RMI was compared with (A) subjective assessment, (B) simple rules with subjective assessment for inconclusive test results, (C) simple rules for which inconclusive results were assumed to be malignant, (D) ADNEX model and (E) LR2 model. Parameters were varied on their 95% confidence intervals (sensitivity, specificity and prevalence), and costs were decreased and increased by 50%. Longer bars represent parameters to which the model is more sensitive. Only the five parameters most sensitive to the outcome of the ICER (with the highest variations) are shown for each model. RMI, Risk of Malignancy Index; SA, subjective assessment; SR + SA, simple rules with subjective assessment for inconclusive test results; SR + Mal, simple rules for which inconclusive results were assumed to be malignant; ADNEX, Assessment of Different NEoplasias in the adneXa (at a cut-off of 10%); LR2, logistic regression model 2 (at a cut-off of 10%); ICER, incremental cost-effectiveness ratio.

4.3. Strengths and limitations

The present study analysed the cost-effectiveness of SA and added four recently introduced and promising ultrasound models to the comparison with the RMI. Together with recently published data on the diagnostic accuracy of the different methods, the present analysis provides a fairly complete overview of the feasibility of introducing new ultrasound methods for the diagnosis of ovarian carcinoma into (Dutch) clinical practice [3,4,8]. Although the study focused on the situation in the Netherlands, we think that it is widely generalisable, since the RMI is the world's most commonly used ultrasound method and the Dutch workup of an adnexal mass is comparable to those in most Western countries.

Cost-effectiveness studies addressing the diagnosis of adnexal masses were previously performed; however, a comparison of the six diagnostic strategies was not performed in these studies [22,23]. A study by Piovano *et al.* in Italy that investigated SR + SA exclusively explored the usefulness of adding tumour markers to SR for the diagnosis of adnexal masses and was limited to the cost of the diagnostic phase [24].

A limitation of cost-effectiveness analyses in general is the uncertainty surrounding the assumptions made. For instance, we used cost prices from an internal cost price investigation. The MUMC+ is a tertiary care centre. Cost prices in third line centres in the Netherlands are generally approximately 50% higher than those in second line centres [14]. Therefore, the use of cost prices from our centre may result in lower total costs than expected. Sensitivity analyses demonstrated that variation in hospital prices (between 50% and 150% of the original price) affected the outcome of the calculations, although the actual differences in the ICER of SA and SR + SA between these lower and upper limits of hospital prices were only 25.5 and 11.9, respectively (Fig. 4). Moreover, the PSA was performed twice with marginal differences between the two analyses, indicating that the results were robust over a range of plausible estimates of event probabilities and costs.

The present study was performed over a short time horizon. Although most of the effects are short-term, an improvement in the rate of correct diagnoses will also have long-term consequences, such as enhanced disease-free survival associated with improvements in referral [25–27]. The long-term effects of the conservative management of asymptomatic patients with a very small risk of malignancy were not included in the present analyses; however, these data are currently not available. The results of projects such as the IOTA 5 study are necessary to perform these analyses (clinicaltrials.gov, NCT01698632).

The present study was limited to a CEA with disease-specific outcomes, instead of generic outcomes such as Quality of Adjusted Life Years (QALYs). We considered performing a quality of life and QALY analysis, but refrained from these analyses because these utilities were not available. Furthermore, no relevant QALY differences are expected in the short-term.

The willingness-to-pay threshold for these outcomes is not known. However, willingness-to-pay should exceed €350 per percent increase in correctly diagnosed patients for SA to have a higher probability of cost-effectiveness than SR + SA. This is equivalent to a willingness-to-pay exceeding €35,000 per correctly diagnosed patient for SA to have a higher probability of cost-effectiveness, whereas the difference in efficacy between SA and SR + SA is small (increased effectiveness of 12.4% and 11.4%, respectively). Thus, there is a large probability that SR + SA is the best method from a cost-effectiveness perspective.

Finally, two- and three-step strategies based on SR, SA, ADNEX and LR2 were developed and have shown a good performance, although their cost-effectiveness has not been evaluated [28,29].

The present study examined the implications of replacing the currently used RMI with other ultrasound methods with better diagnostic accuracy for the differentiation of an ovarian tumour from a cost-effectiveness point of view. Although SA is the best strategy in terms of diagnostic accuracy, SR + SA might be the preferred method from a cost-effectiveness perspective. The assets released by the substitution of the RMI could be used for the training necessary for the implementation of these ultrasound methods.

Conflict of interest statement

None declared.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.ejca.2018.05.003.

References

[1] Siegel R, Ma J, Zou Z, Jemal A. Cancer statistics, 2014. CA Cancer J Clin 2014;64(1):9–29.

[2] Nederlandse Vereniging voor Obstetrie en Gynaecologie, Richtlijn het vergroot ovarium (*guideline 'The enlarged ovary'*). 2013. available from, nvog-documenten.nl/richtlijn/doc/download.php?id=932 [Accessed 26 April 2017].

[3] Meys EM, Kaijser J, Kruitwagen RF, et al. Subjective assessment versus ultrasound models to diagnose ovarian cancer: a systematic review and meta-analysis. Eur J Cancer 2016;58:17–29.

[4] Kaijser J, Sayasneh A, Van Hoorde K, et al. Presurgical diagnosis of adnexal tumours using mathematical models and scoring systems: a systematic review and meta-analysis. Hum Reprod Update 2014;20(3):449–62.

[5] Timmerman D, Testa AC, Bourne T, et al. Simple ultrasound-based rules for the diagnosis of ovarian cancer. Ultrasound Obstet Gynecol 2008;31(6):681–90.

[6] Timmerman D, Testa AC, Bourne T, et al. Logistic regression model to distinguish between the benign and malignant adnexal mass before surgery: a multicenter study by the International Ovarian Tumor Analysis Group. J Clin Oncol 2005;23(34):8794–801.

[7] Van Calster B, Van Hoorde K, Valentin L, et al. Evaluating the risk of ovarian cancer before surgery using the ADNEX model to differentiate between benign, borderline, early and advanced stage invasive, and secondary metastatic tumours: prospective multicentre diagnostic study. BMJ 2014;349:g5920.

[8] Meys EMJ, Jeelof LS, Achten NMJ, et al. Estimating risk of malignancy in adnexal masses: external validation of the ADNEX model and comparison with other frequently used ultrasound methods. Ultrasound Obstet Gynecol 2017;49(6):784–92.

[9] Sullivan SD, Mauskopf JA, Augustovski F, et al. Budget impact analysis-principles of good practice: report of the ISPOR 2012 budget impact analysis good practice II task force. Value Health 2014;17(1):5–14.

[10] Tan SSB-FC, Hakkaart-van Roijen L. Handleiding voor kostenonderzoek: methoden en referentieprijzen voor economische evaluaties in de gezondheidszorg. Tijdschrift voor gezondheidswetenschappen. 2012;90(6):367–72. https://doi.org/10.1007/s12508-12012-10128-12503.

[11] Husereau D, Drummond M, Petrou S, et al. Consolidated health economic evaluation reporting standards (CHEERS) statement. Value Health 2013;16(2):e1–5.

[12] Casparie M, Tiebosch AT, Burger G, et al. Pathology databanking and biobanking in The Netherlands, a central role for PALGA, the nationwide histopathology and cytopathology data network and archive. Cell Oncol 2007;29(1):19–24.

[13] Jacobs I, Oram D, Fairbanks J, Turner J, Frost C, Grudzinskas JG. A risk of malignancy index incorporating CA 125, ultrasound and menopausal status for the accurate preoperative diagnosis of ovarian cancer. Br J Obstet Gynaecol 1990; 97(10):922–9.
[14] Diemen Institute for Medical Technology Assessment Erasmus Universiteit Rotterdam, Kostenhandleiding: Methodologie van kostenonderzoek en referentieprijzen voor economische evaluaties in de gezondheidszorg. 2016.
[15] CBS. Arbeidskosten [Employment costs]. 2014. http://statline.cbs.nl [Accessed 26 April 2017].
[16] Federal government. Income tax. 2014. https://wwwrijksoverheidnl/onderwerpen/inkomstenbelasting [Accessed 26 April 2017].
[17] Briggs AH, Claxton K, Sculpher MJ. Decision modelling for health economic evaluation. Oxford: Oxford University Press; 2006.
[18] Vergote I, De Brabanter J, Fyles A, et al. Prognostic importance of degree of differentiation and cyst rupture in stage I invasive epithelial ovarian carcinoma. Lancet 2001;357(9251):176–82.
[19] ten Broek RP, Issa Y, van Santbrink EJ, et al. Burden of adhesions in abdominal and pelvic surgery: systematic review and met-analysis. BMJ 2013;347:f5588.
[20] Weber S, McCann CK, Boruta DM, Schorge JO, Growdon WB. Laparoscopic surgical staging of early ovarian cancer. Rev Obstet Gynecol 2011;4(3–4):117–22.
[21] Meys E, Rutten I, Kruitwagen R, et al. Simple rules, not So simple: the Use of international ovarian tumor analysis (IOTA) terminology and simple rules in Inexperienced hands in a prospective multicenter cohort study. Ultraschall Med 2017 Aug 23. https://doi.org/10.1055/s-0043–113819 [Epub ahead of print].
[22] Havrilesky LJ, Dinan M, Sfakianos GP, et al. Costs, effectiveness, and workload impact of management strategies for women with an adnexal mass. J Natl Cancer Inst 2015;107(1):322.
[23] Forde GK, Hornberger J, Michalopoulos S, Bristow RE. Cost-effectiveness analysis of a multivariate index assay compared to modified American College of Obstetricians and Gynecologists criteria and CA-125 in the triage of women with adnexal masses. Curr Med Res Opin 2016;32(2):321–9.
[24] Piovano E, Cavallero C, Fuso L, et al. Diagnostic accuracy and cost-effectiveness of different strategies to triage adnexal masses: a prospective study. Ultrasound Obstet Gynecol 2016;Oct 5. https://doi.org/10.1002/uog.17320 [Epub ahead of print].
[25] Giede KC, Kieser K, Dodge J, Rosen B. Who should operate on patients with ovarian cancer? An evidence-based review. Gynecol Oncol 2005;99(2):447–61.
[26] Vernooij F, Heintz P, Witteveen E, van der Graaf Y. The outcomes of ovarian cancer treatment are better when provided by gynecologic oncologists and in specialized hospitals: a systematic review. Gynecol Oncol 2007;105(3):801–12.
[27] Woo YL, Kyrgiou M, Bryant A, Everett T, Dickinson HO. Centralisation of services for gynaecological cancer. Cochrane Database Syst Rev 2012;(3):CD007945.
[28] Ameye L, Timmerman D, Valentin L, et al. Clinically oriented three-step strategy for assessment of adnexal pathology. Ultrasound Obstet Gynecol 2012;40(5):582–91.
[29] Sayasneh A, Kaijser J, Preisler J, et al. A multicenter prospective external validation of the diagnostic performance of IOTA simple descriptors and rules to characterize ovarian masses. Gynecol Oncol 2013;130(1):140–6.