European Journal of Cancer 58 (2016) 17–29

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejcancer.com

Review

Subjective assessment versus ultrasound models to diagnose ovarian cancer: A systematic review and meta-analysis

E.M.J. Meys [a], J. Kaijser [b,1], R.F.P.M. Kruitwagen [a], B.F.M. Slangen [a], B. Van Calster [c], B. Aertgeerts [d], J.Y. Verbakel [d,e], D. Timmerman [b], T. Van Gorp [a,*]

[a] *Department of Obstetrics and Gynaecology, Maastricht University Medical Centre+ (MUMC+), GROW – School for Oncology and Developmental Biology, P.Debyelaan 25, 6202 AZ, Maastricht, The Netherlands*
[b] *Department of Obstetrics and Gynaecology and Leuven Cancer Institute, University Hospital KU Leuven, Herestraat 49, Leuven 3000, Belgium*
[c] *Department of Development and Regeneration, KU Leuven, Herestraat 49, Leuven 3000, Belgium*
[d] *Department of General Practice, KU Leuven, Kapucijnenvoer 33, Leuven 3000, Belgium*
[e] *Nuffield Department of Primary Care Health Sciences, University of Oxford, Woodstock Road, Oxford OX2 6GG, UK*

Received 30 September 2015; received in revised form 8 January 2016; accepted 14 January 2016
Available online xxx

KEYWORDS
Ovarian neoplasms;
Ovarian cancer;
Ultrasonography;
Sensitivity and specificity;
Systematic review;
Meta-analysis

Abstract ***Introduction:*** Many national guidelines concerning the management of ovarian cancer currently advocate the risk of malignancy index (RMI) to characterise ovarian pathology. However, other methods, such as subjective assessment, International Ovarian Tumour Analysis (IOTA) simple ultrasound-based rules (simple rules) and IOTA logistic regression model 2 (LR2) seem to be superior to the RMI.

Our objective was to compare the diagnostic accuracy of subjective assessment, simple rules, LR2 and RMI for differentiating benign from malignant adnexal masses prior to surgery.

Materials and methods: MEDLINE, EMBASE and CENTRAL were searched (January 1990–August 2015). Eligibility criteria were prospective diagnostic studies designed to preoperatively predict ovarian cancer in women with an adnexal mass.

Results: We analysed 47 articles, enrolling 19,674 adnexal tumours; 13,953 (70.9%) benign and 5721 (29.1%) malignant. Subjective assessment by experts performed best with a pooled

* *Corresponding author:* Tel.: +31 43 387 4767.
E-mail addresses: evelyne.meys@mumc.nl (E.M.J. Meys), j.kaijser@ikazia.nl (J. Kaijser), r.kruitwagen@mumc.nl (R.F.P.M. Kruitwagen), brigitte.slangen@mumc.nl (B.F.M. Slangen), ben.vancalster@med.kuleuven.be (B. Van Calster), bert.aertgeerts@med.kuleuven.be (B. Aertgeerts), jan.verbakel@med.kuleuven.be (J.Y. Verbakel), dirk.timmerman@uzleuven.be (D. Timmerman), Toon.van.gorp@mumc.nl (T. Van Gorp).
[1] Present address: Montessoriweg 1, 3083 AN, Rotterdam, The Netherlands.

http://dx.doi.org/10.1016/j.ejca.2016.01.007

sensitivity of 0.93 (95% confidence interval [CI] 0.92–0.95) and specificity of 0.89 (95% CI 0.86–0.92). Simple rules (classifying inconclusives as malignant) (sensitivity 0.93 [95% CI 0.91–0.95] and specificity 0.80 [95% CI 0.77–0.82]) and LR2 (sensitivity 0.93 [95% CI 0.89–0.95] and specificity 0.84 [95% CI 0.78–0.89]) outperformed RMI (sensitivity 0.75 [95% CI 0.72–0.79], specificity 0.92 [95% CI 0.88–0.94]). A two-step strategy using simple rules, when inconclusive added by subjective assessment, matched test performance of subjective assessment by expert examiners (sensitivity 0.91 [95% CI 0.89–0.93] and specificity 0.91 [95% CI 0.87–0.94]).

Conclusions: A two-step strategy of simple rules with subjective assessment for inconclusive tumours yielded best results and matched test performance of expert ultrasound examiners. The LR2 model can be used as an alternative if an expert is not available.

1. Introduction

1.1. Rationale and objectives

In order to ensure that ovarian cancer patients receive appropriate treatment, an accurate characterisation of any adnexal mass that needs surgery is pivotal to improve the outcome of this disease. Subjective assessment by experienced examiners, also called 'pattern recognition', is generally accepted to be the best way to classify adnexal masses prior to surgery. Several individual reports have demonstrated that subjective assessment is superior to the use of scoring systems and mathematical models, such as International Ovarian Tumour Analysis (IOTA) simple ultrasound-based rules (simple rules), IOTA logistic regression model 2 (LR2) or the risk of malignancy index (RMI) [1–4]. However, both LR2 and simple rules closely approximate the performance of subjective assessment by expert examiners [5,6]. An advantage of these models over subjective assessment is their objectivity and simplicity which facilitates their use by ultrasonographers with different backgrounds and various levels of experience [7–10]. Despite accumulating and compelling evidence in favour of both subjective assessment and the ultrasound-based models such as simple rules and LR2, many national guidelines concerning the management of ovarian masses still advocate the use of RMI in the classification of adnexal masses. Consequently, the RMI is still the most commonly used model in clinical practice.

Several reviews have critically appraised the evidence relating to this subject [5,6,11–16]. However, none of these has provided a meta-analysis on the test performance of subjective assessment of adnexal tumours, while in general this method is considered the most accurate way to distinguish benign from malignant adnexal tumours. The aim of this meta-analysis was to compare the diagnostic accuracy of subjective assessment, simple rules, LR2 and RMI for the pre-operative differentiation of benign and malignant adnexal masses.

2. Methods

2.1. Protocol and registration

All methods described in this manuscript were determined in advance and recorded in a study protocol (Prospero CRD42013004334, http://www.crd.york.ac.uk/PROSPERO). The conduct of this systematic review and meta-analysis was done in accordance with prevailing guidelines (http://www.prisma-statement.org and http://srdta.cochrane.org/handbook-dta-reviews).-

2.2. Eligibility criteria

Eligible studies had to evaluate diagnostic accuracy of subjective assessment, simple rules, LR2 and/or RMI for the characterisation of adnexal tumours in women scheduled for surgery (in order to obtain a final histological diagnosis). Regarding subjective assessment, studies were only eligible when the diagnosis of the tumour was based purely on the ultrasonographic interpretation of the examiner (whether or not complemented with clinical information, such as medical history).

The simple rules comprise two strategies; simple rules supplemented with subjective assessment in case the simple rules could not be applied, or classification of all masses in which simple rules could not be applied as malignant [17]. Studies evaluating either of these strategies or both were eligible.

Three principal variants of the RMI have been described (RMI-I, II and III) which differ according to points attributed to the different ultrasound variables and the menopausal status of the patient [18–20]. All studies regarding one or more of these three versions were eligible.

Furthermore, eligible studies had to contain sufficient data to extract 2 × 2 contingency tables of diagnostic test performance.