UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE MEDICINA
PROGRAMA DE PÓS-GRADUAÇÃO
EM EPIDEMIOLOGIA

REVISÃO SISTEMÁTICA COM ENFOQUE DIAGNÓSTICO E TERAPÊUTICO DAS TUMORAÇÕES OVARIANAS

Tese de Doutorado

Porto Alegre
2008

LÍDIA ROSI DE FREITAS MEDEIROS

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE MEDICINA
PROGRAMA DE PÓS-GRADUAÇÃO EM EPIDEMIOLOGIA

TESE DE DOUTORADO

REVISÃO SISTEMÁTICA COM ENFOQUE DIAGNÓSTICO E TERAPÊUTICO DAS TUMORAÇÕES OVARIANAS

Orientadora: Profa. Dra. MARY CLARISSE BOZZETTI

A apresentação desta tese é exigência do Programa de Pós-graduação em Medicina: Epidemiologia, Universidade Federal do Rio Grande do Sul, para obtenção do título de Doutor.

Porto Alegre 2008

CATALOGAÇÃO NA PUBLICAÇÃO

M488r Medeiros, Lídia Rosi de Freitas
Revisão sistemática com enfoque diagnóstico e terapêutico das tumorações ovarianas / Lídia Rosi de Freitas Medeiros. – 2008.
288f. : il.

Tese (doutorado) – Universidade Federal do Rio Grande do Sul. Faculdade de Medicina. Programa de Pós-Graduação em Epidemiologia. Porto Alegre, 2008.
Orientação: Mary Clarisse Bozzetti.

1. Neoplasias ovarianas. 2. Laparoscopia. 3. Laparotomia. 4. Diagnóstico. 5. Terapêutica. I. Bozzetti, Mary Clarisse. II. Título.

CDU 618.11-006

Mara Lúcia Araújo Meireles
CRB-10/1003

Creio que a educação libertadora
implica a iluminação da realidade,
mas os iluminados são os dois agentes do processo,
os educadores e os educandos.

(Paulo Freire)

BANCA EXAMINADORA

Prof. Bernardo Gacia de Oliveira Soares, Professor Convidado do Departamento de Medicina de Urgência e Pesquisador da Cochare, da Universidade Federal de São Paulo

Prof. Dr. Edison Capp, Professor Adjunto do Departamento de Ginecologia e Obstetrícia da Universidade Federal do Rio Grande do Sul.

Profa. Dra. Sandra Costa Fuchs, Professora Adjunta do Departamento de Medicina Social UFRGS e do Curso de Pós-Graduação em Epidemiologia e de Clínica Médica da Universidade Federal do Rio Grande do Sul.

Dedico esta tese de doutorado a minha mãe **Cecília**, pois me ensinou a importância do ser e do saber, que os limites da vida são relativos e que todos os sonhos são possíveis desde que a força para realizá-los brote do coração.

Agradecimentos

Agradeço, com muito, reconhecimento, a todos os que, de uma maneira ou outra, direta ou indiretamente, possibilitaram que este trabalho fosse realizado. De maneira particular, sou grata

- à Professora Doutora **Mary Clarisse Bozzetti**, Professora Adjunta do Departamento de Medicina Social da Universidade Federal do Rio Grande do Sul e do Curso de Pós-Graduação em Epidemiologia da UFRGS, por sua amizade, disponibilidade, paciência, compreensão, incentivo e que, por seu entusiasmo à ciência e à pesquisa, me auxiliou na elaboração da tese;
- ao Professor Doutor **Bruce Duncan**, Professor Adjunto de Departamento de Medicina Social da UFRGS e coodernador do Curso de Pós-Graduação em Epidemiologia da UFRGS, pela disponibilidade, estímulo e compreensão;
- à Professora Doutora **Maria Inês Schmidt**, Professora Adjunta do Departamento de Medicina Social UFRGS e do Curso de Pós-Graduação em Epidemiologia da UFRGS, por sua amizade, incentivo, apoio e compreensão;
- ao Professor Doutor **Airton Stein**, que conduziu meus primeiros passos no aprofundamento da ciência e me proporcionou ter uma visão ampla do conhecimento científico;
- à Doutora **Sandra Costa Fuchs**, Professora Adjunta do Departamento de Medicina Social UFRGS e do Curso de Pós-Graduação em Epidemiologia e de Clínica Médica da UFRGS, pela oportunidade de crescimento científico, pelas palavras e atitudes de incentivo e apoio e por ter confiado e acreditado;
- à Doutora **Daniela Dornelles Rosa**, pela amizade, compreensão, paciência, incentivo e por ter me salvado a vida com sua atitude rápida mesmo estando tão distante;
- ao Doutor **Alberto Agusto Alves Rosa**, pela amizade, desprendimento e por ter me escutado, acreditado nas minhas queixas e me encaminhando a uma equipe maravilhosa que salvou a minha vida e tornou possível a apresentação deste trabalho;

- Ao Doutor **Paulo Ricardo Avancini Caramori**, por ter entendido o meu desespero e acreditado que algo estava errado, por ter me avaliado com o seu saber científico, por ter feito um procedimento médico, que me salvou a vida e me trouxe de volta a vontade de viver e crescer – sem ele esta apresentação seria impossível;
- ao Doutor **Diovane Berlezze**, cardiologista, querido amigo que, com seu saber, também me salvou a vida, me equilibrando cardiologicamente, me devolvendo a qualidade de vida e a vontade de viver – sem ele a apresentação deste trabalho seria impossível;
- ao Doutor **Guilherme Sudbrack**, que com o seu saber, baseado nas melhores evidências científicas, escutou as minhas queixas, encontrou explicações fisiopatológicas, tratou a minha dor, eliminou meu sofrimento, me devolvendo a vida e a vontade de viver – sem ele a apresentação deste trabalho seria impossível;
- à Doutora **Karin Mombach**, que me escutou e me devolveu o equilíbrio para poder seguir em frente – sem ela a apresentação deste trabalho seria impossível;
- ao meu colega e amigo de jornada **Dr. Marco Aurélio Sbroglio**, pelo incentivo, pelo carinho, paciência e apoio ao longo de 22 anos de profissão;
- à Doutora **Maria Inês da Rosa**, colega do doutorado que me deu seu carinho, apoio e incentivo;
- às colegas de doutorado, Doutoras **Roselaine Zanini** e **Anaelena de Bragança,** amigas de valor que muito contribuíram para o meu crescimento, me auxiliaram quando precisei e me deram seu incentivo;
- ao colega **Fernando Bernd**, cuja paixão, dedicação e conhecimentos motivaram minha opção pela especialidade;
- aos **professores** do Curso de Pós-Graduação em Epidemiologia da UFRGS, pelos conhecimentos transmitidos, fundamentais para o constante aperfeiçoamento profissional e pessoal;
- aos **colegas** do Curso de Pós-Graduação em Epidemiologia da UFRGS, pelo incentivo e carinho;
- ao professores Doutores **Kurt Semm**, **Jacques Donnez** e **Liselote Mettler**, pelos ensinamentos na área da cirurgia laparoscópica ginecológica;

- à Professora Doutora **Maria do Horto Motta**, que fez a revisão do texto, com excelentes críticas e sugestões, pelo seu carinho, pelas palavras de apoio e incentivo e por sua amizade;
- aos **funcionários** da Secretaria do Curso de Pós-Graduação em Epidemiologia da UFRGS, à Carmen, ao Rodrigo, pela eficiência no atendimento e pela ajuda na resolução de problemas;
- à **Clair Azevedo**, responsável pela editoração do texto, feita com inegável capacidade, pela disponibilidade, por seu carinho e amizade;
- à **Maria de Fátima Guterres** por seu carinho, compreensão e apoio nos momentos mais difíceis da minha vida;
- ao meu irmão **Vilmo Luiz Medeiros,** que com seu carinho e apoio, mesmo a distância, tornou a realização deste projeto viável de maneira correta.

Muito obrigada a todos queridos amigos.

Sumário

Lista de Abreviaturas

Resumo

Abstract

Lista de Quadro e Lista de Figura

Apresentação

1 INTRODUÇÃO 16

2 REVISÃO DA LITERATURA 19

2.1 Revisão sistemática 19

2.1.1 Revisão sistemática com enfoque diagnóstico 26

2.1.1.1 Histórico e definição 26

2.1.1.2 Estratégias de busca de artigos com enfoque diagnóstico 27

2.1.1.3 Seleção dos estudos 29

2.1.1.4 Metanálise nos estudos de acurácia diagnóstica 29

2.1.1.4.1 Somatório da sensibilidade e especificidade na metanálise diagnóstica 30

2.1.1.4.2 DOR (*diagnostic Odds Ratio*) 32

2.1.2 Revisão sistemática de estudos observacionais com enfoque terapêutico (qualitativa) 34

2.2 Tumores ovarianos malignos 36

2.2.1 Dados epidemiológicos 36

2.2.2 Classificação histológica dos tumores de ovário 37

2.2.3 Diagnóstico das tumorações ovarianas 39

2.2.3.1 Ultra-sonografia transvaginal e com Doppler colorido 40

2.2.3.2 Dosagem de marcador CA 125 41

2.2.4 Manejo clínico e cirúrgico das tumorações ovarianas 42

2.2.4.1 Laparoscopia *versus* laparotomia nas tumorações ovarianas malignas estágios iniciais pela FIGO (Ia, Ib e Ic) 42

2.2.5 Síntese da revisão de literatura 45

3 REFERÊNCIAS BIBLIOGRÁFICAS 47

4 OBJETIVOS 58
4.1 Objetivo geral 58
4.2 Objetivos específicos 58
5 ARTIGOS 59
5.1 Artigo 1: Laparoscopy *versus* laparotomy for figo stage i ovarian cancer 60
5.2 Artigo 2: Accuracy of frozen section analysis in the diagnosis of ovarian tumors: a systematic quantitative review 106
5.3 Artigo 3: Accuracy of ultrasonography with color doppler in ovarian tumor: a systematic quantitative review 133
5.4 Artigo 4: Accuracy of CA 125 in ovarian tumor: a systematic quantitative review 157
6 CONCLUSÕES E CONSIDERAÇÕES FINAIS 181
7 ANEXOS 184

Lista de Abreviaturas

AUC	*area under the curve*
CA-125	*Cancer antigen number 125*
CO_2	Dióxido de carbono
CONSORT	*Consolidated Standards of Reporting Trials*
DOR	*Diagnostic Odds ratio*
DP	Desvio padrão
E	Especificidade
ECR	Ensaio clínico randomizado
EMBASE	*Excerpta Medica Database*
EMTREE	Descritor de vocabulário do EMBASE (Thesaurus)
FIGO	*International Federation of Gynecology and Obstetrics*
IC 95%	Intervalo de confiança de 95%
IP	Índice de pulsatilidade
IR	Índice de resistência
LILACS	Literatura Latino-Americana em Ciências da Saúde
MEDLINE	*Medical Literature Retrieval System Online*
MeSH	*Medical Subject Headings*
MHz	Mega Hertz
MOOSE	*Meta-analysis of Observational Studies in Epidemiology*
mm	Milímetros
NOS	*Newcastle-Ottawa Scale*
OC 125	Antígeno monoclonal cancer 125
OR	*Odds ratio*
PUBMED	Banco de dados de informações bibliográficas na área da ciência, desenvolvido pelo *National Center for Biotechnology Information* (NCBI) *at the National Library of Medicine*
QUORUM	*Quality of Reporting of Meta-analysis*
ROC	*Receiver operator characteristic*
RS	Revisão sistemática
S	Sensibilidade
SE	*Stardand error*
SROC	*Summary of receiver operator characteristic*
STARD	*Standards for Reporting of Diagnostic Accuracy*
STROBE	*Strengthening the Reporting of Observational Studies in Epidemiology*
T	Teste T
RCT	*Randomized control trial*
U/ml	Unidade internacional por mililitro
VPP	Valor preditivo positivo
VPN	Valor preditivo negativo

Resumo

Introdução: Em cirurgia ginecológica, o tratamento dos tumores ovarianos é tema que suscita inúmeras controvérsias quanto à melhor abordagem – se por laparoscopia ou por laparotomia. Entretanto, nos casos de neoplasia maligna de ovário, o manejo por essa via torna-se inadequado, em decorrência das pequenas incisões que dificultam os cuidados operatórios preconizados para essas neoplasias. Quando uma tumoração ovariana é detectada, faz-se necessário estabelecer suas características de benignidade ou malignidade, sendo que determinados exames, de certa forma, guiam a conduta terapêutica a ser adotada, apesar de não possuírem 100% de acurácia. Esses exames são a ultra-sonografia transvaginal com Doppler colorido, o marcador tumoral CA 125 e o exame anatomopatológico de congelação durante o ato cirúrgico. Devido às incertezas que suscitam controvérsias quanto à melhor abordagem inicial para o manejo operatório das tumorações ovarianas, se por laparoscopia ou por laparotomia, planejamos realizar estudos de revisão sistemática. A revisão sistemática com enfoque terapêutico, comparou laparoscopia e laparotomia, na abordagem cirúrgica do câncer ovariano, em seus estágios iniciais (Ia, Ib, Ic) pela Federação Internacional de Ginecologia e Obstetrícia (FIGO). Estudos de revisão sistemática, com enfoque diagnóstico, avaliaram a acurácia diagnóstica da ultra-sonografia com Doppler colorido, do CA 125 e do exame anatomo-patológico de congelação.

Metodologia: A revisão sistemática com enfoque terapêutico – laparoscopia *versus* laparotomia para câncer ovariano em estágio inicial (Ia, Ib,Ic) pela FIGO, foi realizada de maneira sistemática, entre os anos 1990 e 2007, com o auxílio do Grupo de Câncer Ginecológico da Cochrane, nos seguintes bancos de dados: Cochrane Central de Registro de Estudos Controlados (CENTRAL- janeiro de 1990 a novembro de 2007), MEDLINE (janeiro de 1990 a novembro de 2007), EMBASE (janeiro de 1990 a novembro de 2007), LILACS (janeiro de 1990 a novembro 2007), BIOLOGICAL ABSTRACTS (janeiro de 1990 a novembro 2007) e CANCERLIT (janeiro de 1990 a novembro 2007). Foram procuradas também publicações através

de busca manual, em periódicos médicos relevantes até novembro de 2007 e listas de referência de artigos bem como resumos de conferências médicas. Foram incluídos estudos relativos a pacientes, com comprovação histológica de câncer de ovário em seus estágios iniciais de acordo com a FIGO. Estudos comparando laparoscopia e laparotomia para manejo clínico dessas, começaram a ser realizados a partir de 1990, presumindo-se que devam existir muito poucos estudos randomizados com esse tipo de enfoque. Por tal razão, foram também considerados estudos de coorte e de caso-controle, comparando as duas técnicas, tendo sido excluídos somente estudos de séries de casos e retrospectivos. Os dados foram extraídos de forma independente por cinco revisores que aferiam a qualidade do estudo e a qualidade dos dados extraídos. A extração dos dados inclui avaliar as características clínicas de cada estudo, como tipo de participantes, intervenção e desfechos. A qualidade dos estudos não randomizados foi aquilatado por dois tipos de instrumentos de avaliação metodológica: o STROBE (*Strengthening the Reporting of Observational Studies in Epidemiology*) e o NOS (*Newcastle-Ottawa*) para estudos com enfoque observacional. A análise estatística foi realizada de acordo com o *guideline* desenvolvido pelo *Cochrane Gynaecological Cancer Group.* Heterogeneidade estatística entre os resultados dos diferentes estudos é examinada usando-se o teste Q_T (*Cochran)* para distribuição do χ^2 com N-1 graus de liberdade. Se não há evidência de heterogeneidade estatística, o valor é de P>0,10, e o modelo de efeitos fixos pode ser usado. Em caso de heterogeneidade importante P<0,10 deve-se explorar a qualidade metodológica e utilizar, para os cálculos estatísticos, o modelo de efeitos aleatórios. Outra alternativa, para quantificar o efeito da heterogeneidade é a inconsistência (I^2) com intervalo de confiança de 95%. Esta descreve a percentagem de variabilidade do efeito estimado devido à heterogeneidade. O valor de 0% indica que não há heterogeneidade e um valor maior que 50%, indica que existe uma heterogeneidade substancial. Se existir um somatório inapropriado dos dados devido a aspectos metodológicos, clínicos ou de heterogeneidade estatística, realiza-se uma revisão sistemática qualitativa sem realizar uma metanálise.

A revisão sistemática, com enfoque diagnóstico para avaliar a acurácia do exame de anatomopatológico de congelação, foi realizada entre os anos de 1984 a 2003, nos mesmos bancos de dados descritos na primeira revisão sistemática. Foram procuradas também publicações através de busca manual, em periódicos

médicos relevantes e listas de referência de artigos bem como resumos de conferências médicas. Foram usados os seguintes termos médicos para busca dos artigos: "*ovarian neoplasm"* e "*frozen section"* combinados com os seguintes termos diagnósticos ("sensitivity and specificity"). Foram incluídos estudos observacionais que comparam o resultado desse exame com o resultado final da histologia no exame de parafina nesta revisão sistemática. O diagnóstico final das lesões ovarianas foi comparado de quatro maneiras: (1) concordância entre lesões benignas *versus borderline* ou maligna, (2) concordância entre lesões malignas *versus* benignas, (3) concordância entre lesões *borderline versus* benignas e (4) concordância entre lesões *borderline versus* malignas. Os dados foram extraídos de forma independente por quatro revisores que aferiam a qualidade do estudo e a qualidade dos dados extraídos. A extração dos dados inclui avaliar as características clínicas de cada estudo. A qualidade metodológica de cada estudo foi aferida através do *Oxford Centre for Evidence Based Medicine Level of Classification* e através do escore de qualidade metodológica de estudos com enfoque diagnóstico. Na análise estatística a concordância entre o exame anatomopatológico de congelação e o exame de parafina foi calculado através do coeficiente *Kappa*. Para cada estudo foi construída uma tabela de contingência (2 x 2). Foram calculados a sensibilidade, os falso-positivos, a especificidade, a razão de verossimilhança e a probabilidade pós-teste de cada estudo com intervalos de confiança de 95%. A associação entre sensibilidade e especificidade foi calculada através do coeficiente de correlação de *Spearman.* Quando não há correlação, o somatório da sensibilidade e especificidade pode ser calculado, desde que haja 2 categorias de resultados (positiva e negativa). A probabilidade pós-teste foi estimada multiplicando-se a prevalência pela razão de verossimilhança. A heterogeneidade da sensibilidade e especificidade dos estudos foi acessada usando o teste Q_T (*Cochran)* para distribuição do χ^2 com N-1 graus de liberdade.

A revisão sistemática, com enfoque diagnóstico para avaliar a acurácia da ultra-sonografia com Doppler colorido, foi realizada entre os anos de 1990 a 2007, nos mesmos bancos de dados descritos na primeira revisão sistemática. Foram procuradas também publicações através de busca manual, em periódicos médicos relevantes e listas de referência de artigos bem como resumos de conferências médicas. Foram usados os seguintes termos médicos para busca dos artigos: "*ovarian neoplasm"* e "*transvaginal ultrasound with color Doppler*" combinados com

os seguintes termos diagnóstico (“sensitivity and specificity”). Foram incluídos estudos observacionais que comparam o resultado desse exame com o resultado final da histologia no exame de parafina nesta revisão sistemática. Lesões malignas são suspeitas quando o valor do índice de resistência for menor que 0,5. O diagnóstico final das lesões ovarianas foi comparado somente de utilizando a concordância entre lesões *borderline* ou malignas *versus* benignas.

A revisão sistemática, com enfoque diagnóstico para avaliar a acurácia do CA 125, foi realizada entre os anos de 1990 a 2007, nos mesmos bancos de dados descritos na primeira revisão sistemática. Foram procuradas também publicações através de busca manual, em periódicos médicos relevantes e listas de referência de artigos bem como resumos de conferências médicas. Foram usados os seguintes termos médicos para busca dos artigos: “*ovarian neoplasm”* e “*CA125*” combinados com os seguintes termos diagnósticos (“sensitivity and specificity”). Foram incluídos estudos observacionais que comparam o resultado desse exame com o resultado final da histologia no exame de parafina nesta revisão sistemática. Lesos malignas são suspeitas quando o nível sérico do CA 125 ≥ 35 U/ml. O diagnóstico final das lesões ovarianas foi comparado somente de uma maneira: concordância entre lesões *borderline* ou malignas *versus* benignas.

As duas revisões sistemáticas acima (ultra-sonografia com Doppler colorido e CA 125) tiveram seus dados extraídos de forma independente por 4 revisores que aferiam a qualidade do estudo e a qualidade dos dados extraídos. A extração dos dados inclui avaliar as características clínicas de cada estudo. A qualidade metodológica de cada estudo foi aferida através do *Oxford Centre for Evidence Based Medicine Level of Classification* e através do escore de qualidade metodológica de estudos com enfoque diagnóstico. Na análise estatística a concordância entre a ultra-sonografia com Doppler e o CA 125 com o exame de parafina foi calculado através do coeficiente *Kappa*. Para cada estudo foi construída uma tabela de contingência (2 x 2). Foi calculada: a sensibilidade, o falso positivo, a especificidade com intervalo de confiança de 95%. O sumário da sensibilidade e especificidade foram calculadas somando-se a sensibilidade e especificidade de cada estudo, multiplicando-se pelo número de pacientes de cada estudo, dividindo-se pelo número total de pacientes de todos os estudos, o intervalo de confiança é calculado usando-se o mesmo método. A associação entre sensibilidade e especificidade foi calculada através do coeficiente de correlação de *Spearman.*

Quando não há correlação o somatório da sensibilidade e especificidade pode ser calculada, desde que haja 2 categorias de resultados (positiva e negativa). Quando há correlação calcula-se o sumário da curva ROC (SROC) usando-se os dados dos limiares, usando método de *Littenberg and Moses.* Foi calculada a *odds ratio* diagnóstica (DOR) que pode descrever as diferentes combinações entre sensibilidade e especificidade. Um resultado de *odds* diagnóstica corresponde a um balanceamento entre sensibilidade e especificidade descrito pelo SROC. Foi calculada a área sobre a curva que pode sumarizar a capacidade inerente de um teste discriminatório para doença ou não doença. A heterogeneidade da sensibilidade e especificidade dos estudos foi acessada usando o teste Q_T (*Cochran*) para distribuição do χ^2 com N-1 graus de liberdade.

Resultados: A revisão sistemática com enfoque terapêutico – l laparoscopia versus laparotomia para câncer ovariano em estágio inicial (Ia, Ib,Ic) pela FIGO não encontrou nenhum estudo randomizado. Somente, três estudos observacionais, mas com pobre qualidade metodológica. Portanto, tornou-se impossível realizar a metanálise devido a diferença entre os estudos. Estes três estudos observacionais foram classificados pelos autores como um sendo de uma coorte prospectiva e outros dois estudos como sendo de casos e controles. Entretanto, ao analisar-se os artigos, o desenho metodológico dos dois artigos classificados como de casos e controles eram na realidade estudos de coorte. Não podemos mudar a classificação da publicação dada pelos autores e mantivemos como de casos e controles, mas salientamos que na realidade esses todos os três estudos são na realidade de coorte prospectiva.

A revisão sistemática com enfoque diagnóstico do exame de anatomopatológico de congelação selecionou 14 estudos para serem analisados que incluíram 3.659 mulheres. Lesões ovarianas benignas *vs* lesões limítrofes (*bordeline*) ou malignas de ovário possuem, no somatório final de todos os estudos inclusos, uma razão de verossimilhança positiva de 8,7 (IC 95% de 7,3-10,4) e probabilidade pós-teste para lesões benignas de 95% (IC 95% de 94%-96%). Exame anatomopatológico de congelação comparando resultados de malignidade *vs* lesões benignas possui uma razão de verossimilhança positiva, no somatório geral de todos estudos inclusos, de 303 (IC 95% de 101-605) com aumento da probabilidade pós-teste para malignidade de 98% (IC 95% de 97%-99%). Já na comparação de

tumores com malignidade limítrofe (*borderline*) *vs.* lesões benignas ovarianas foi encontrado no somatório geral de todos os estudos, uma razão de verossimilhança de 69 (IC 95% de 45-106) com aumento de probabilidade pós-teste para tumores com malignidade limítrofe de 79% (IC 95% de 71%-85%). Ao se comparar o resultado do somatório de todos os estudos com tumores com malignidade limítrofe (*bordeline*) *vs* tumores malignos, tem-se um valor de verossimilhança bem menor em relação aos anteriores, sendo de 18 (IC95% de 13-26) tendo uma probabilidade pós-teste somente de 51% (IC 95% de 42%-60%).

A revisão sistemática, com enfoque diagnóstico, da ultra-sonografia com Doppler, selecionou 12 estudos que incluíram 2.398 mulheres. O somatório de estudos para o cálculo de sensibilidade encontrou um valor de 0,87 (IC95% de 0,84-0,90); para especificidade esse valor foi de 0,92 (IC 95% de 0.87-0.90). A DOR para tumores ovarianos ou com malignidade limítrofe (*borderline*) *vs* lesões benignas foi de 125 (95% CI de 55-283). Foi realizado um sumário da curva ROC (SROC) devido à heterogeneidade encontrada na DOR. Entretanto, embora haja heterogeneidade na comparação dos tumores malignos ou malignidade limítrofe (*borderline*) *vs.* lesões benignas, a área sob a curva foi de 0,9573.

A revisão sistemática com enfoque diagnóstico do CA 125 selecionou 17 estudos que incluíram 2.374 mulheres. No cálculo que sumariza os resultados da sensibilidade de todos os estudos, encontrou-se um valor de 0,80 (IC95% de 0,76-0,82); na especificidade esse valor foi de 0,75 (IC 95% de 0,73-0,77). O valor da DOR para diagnóstico de câncer ovariano ou de lesões limítrofes *vs* tumorações benignas foi de 21,2 (IC 95% de 12-37). O sumário da curva ROC foi construído devido à heterogeneidade da DOR. Para diagnóstico de câncer ovariano ou de lesões limítrofes *vs* tumorações benignas, a área sob a curva (AUC) foi de 0,8877.

Conclusão

A revisão sistemática com enfoque terapêutico – laparoscopia *versus* laparotomia para câncer ovariano em estágio inicial (Ia, Ib,Ic) pela FIGO – não encontrou evidências que apóiem o uso da laparoscopia para o manejo inicial do câncer ovariano em seus estágios iniciais como rotina na prática clínica. Verifica-se a necessidade de estudos com boa qualidade metodológica comparando laparoscopia e laparotomia para o manejo cirúrgico inicial desses tipos de câncer.

Assim como a importância da classificação correta do desenho clínico dos estudos observacionais tanto pelos autores, bem como pelos revisores dos periódicos, visto que tivemos 2 artigos de coorte classificados erroneamente como sendo de casos e controles.

A revisão sistemática com enfoque diagnóstico do exame de anatomopatológico de congelação Conclui-se, assim, que a acurácia do exame de anatomopatológico de congelação é elevada no diagnóstico de tumores malignos e benignos, sendo baixa, porém, para as lesões com malignidade limítrofe.

A revisão sistemática, com enfoque diagnóstico, da ultra-sonografia com Doppler é um exame pré-teste importante na predição da natureza da tumoração ovariana – se maligna ou benigna.

A revisão sistemática com enfoque diagnóstico do CA 125 com nívell sérico ≥ 35 U/ml é importante, nas tumorações ovarianas, para predizer se a lesão ovariana é de natureza benigna ou maligna.

Em síntese, estes quatros estudos de revisão sistemática que abordam as tumorações ovarianas, mostraram níveis de evidência baixo para abordagem laparoscópica no câncer ovariano estágios inicias (Ia, Ib e Ic) pela FIGO. E que os três exames diagnósticos mais realizados na presença de tumoração ovariana são extremamente importantes para definir o diagnóstico. Definindo-se o diagnóstico, cabe adotar a melhor opção terapêutica.

Abstract

Introduction: In gynecological surgery, the treatment of the ovarian tumors is theme that raises countless controversies with relationship to the best approach if for laparoscopy or for laparotomy. However, in the cases of malignancy of ovarian tumor, the handling for that road becomes inadequate, due to the small incisions that hinder the operative cares extolled for those neoplasias. When an ovarian tumor is detected, it is necessary to establish its characteristics of benignancy or malignancy, and certain exams, in a certain way, guide the therapeutic conduct to be adopted one, in spite of they do not possess 100% of accuracy. Those exams are the ultrasonography with color Doppler, CA 125 assay and accuracy of frozen section the diagnosis of ovarian tumors. Due to the uncertainties that raise controversies regarding the relationship of the best initial approach for the operative handling of the ovarian tumor, if laparoscopy or laparotomy, we planned to accomplish studies of systematic review. The systematic review with therapeutic focus, compared laparoscopy and laparotomy, in the surgical approach of the ovarian cancer, in early stage (Ia Ib, Ic) by International Federation of Gynecology and Obstetrics (FIGO). Studies of systematic revision, with focus on diagnosis, evaluated the accuracy of the ultrasonography with color Doppler, of CA 125 and the frozen section.

Methods: The systematic review with therapeutic focus - laparoscopy versus laparotomy for ovarian cancer by FIGO stage I (Ia, Ib,Ic), was accomplished in a systematic way, from the years of 1990 to 2007, with the aid of the Group of Gynecological Cancer of Cochrane, in the following databases: Cochrane Central Register of Controlled Trials (CENTRAL) (The Cochrane Library Issue 2, 2007), MEDLINE (January 1990 to date), EMBASE (1990 to November 2007), LILACS (1990 to November 2007), BIOLOGICAL ABSTRACTS (1990 to November 2007) and Cancerlit (1990 to November 2007). We also searched our own publication archives, based on prospective handsearching of relevant journals up to November 2007. Reference lists of identified studies, gynaecological cancer handbooks and conference abstract were also scanned. Studies regarding patients with histologically

proven stage I ovarian cancer according to the International Federation of Gynaecology and Obstetrics (FIGO) was included in this review. Studies comparing laparoscopic surgery with laparotomy for early stage ovarian cancer were only available from 1990. It was anticipated that a very small number of randomised controlled trials (RCTs) were conducted studying the management of early stage ovarian cancer. Therefore, non-randomised comparative studies, cohort studies and case-controls studies, but not studies with historical controls, were also considered for this review. Data extractions were performed independently by five reviewers who assessed study quality and quality of extracted data. Extracted data included trial characteristics, characteristics of the study participants, interventions and outcomes. The quality of non randomised studies was assessed using appropriate quality evaluations tools from the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) and from the Newcastle-Ottawa tool for observational studies.

Statistical analysis was performed in accordance to the guidelines developed by the Cochrane Gynaecological Cancer Group. All trials were initially included in one analysis of surgical laparoscopy and laparotomy for early stage ovarian cancer. Statistical heterogeneity between the results of different studies was examined by checking the usual statistical test (Cochran's Q) where P values were obtained by comparing the statistic with a chi-square distribution. Care was taken in the interpretation of the chi-squared test, since it has low power in the (common) situation of a meta-analysis when trials had small sample size or when the sample is small. If there was no evidence of statistical heterogeneity ($P > 0.10$), a fixed effects model was used. If there was significant heterogeneity ($P < 0.10$), the possible clinical and methodological reasons for this was explored qualitatively and a random effects model was used. However, since clinical and methodological diversity always occur in a meta-analysis, statistical heterogeneity was inevitable. An alternative approach that quantifies the effect of heterogeneity is inconsistency (I2), which provides a measure of the degree of inconsistence in the studies' results with 95% uncertainty intervals. This describes the percentage of the variability in the effect estimates that is due to heterogeneity rather than sampling error (chance). A value of 0% indicates no observed heterogeneity and a value greater than 50% may be considered substantial heterogeneity. If it is not appropriate to pool the data because

of clinical or statistic heterogeneity, then a systematic review without a meta-analysis or a meta-analysis excluding outlying studies may be performed.

The systematic review, with focus on diagnosis to evaluate the accuracy of frozen section, was accomplished from the years of 1984 to 2003, in the same databases described in the first systematic review. Reference lists of all available primary studies were reviewed to identify additional relevant citation The medical subjects heading (MeSH) and text words for the terms "*ovarian neoplasm"* and "*frozen section"* were combined with the MeSH term *diagnosis* ("sensitivity and specificity"). This review focused on observational studies in which the results of the diagnostic test of interest were compared with the results of a reference standard. The final diagnoses of ovarian lesions were compared in four ways: (1) agreement as to benign *versus* borderline and/or malignant, (2) agreement as to malignant *versus* benign, (3) agreement as to borderline *versus* benign and (4) agreement as to borderline *versus* malignant. The reviewed studies were identified independently by four investigators. The extraction of the data includes to evaluate the characteristics clinics of each study. The methodological quality of each study was confronted through the methodological quality with reference to the *Oxford Centre for Evidence Based Medicine Level of Evidences Classification* and by "*Scoring of Study Quality*". To evaluate agreement between study eligibility and methodological quality assessments as well as agreement between frozen and paraffin section analyses the observed percentage agreement and κ coefficient for inter-rater reliability were calculated. For each study, we constructed 2 x 2 contingency tables in which all biopsies were classified as benign, borderline or malignant lesions. We calculated the true-positive rate (TPR; sensitivity), false-positive rate (FPR; 1 – specificity), likelihood ratios (LRs) and post-test probability for each study along with 95% confidence intervals (CIs). The association between sensitivity and specificity was calculated with *Spearman's correlation coefficient* for benign, borderline and malignant ovarian tumors. In this case, estimates of sensitivity and specificity can be calculated because there are two categories of results (negative or positive test outcome) and there is not variability of the diagnostic threshold. In addition, the pooled post-test probability estimates were calculated by multiplying the pre-test probability (prevalence) by the pooled likelihood ratios (LRs). The heterogeneity of the sensitivities and specificities from the studies was assessed using the Q_T

(Cochran) test for χ^2 distributions with N-1 degrees of freedom.

The systematic review, with focus on diagnosis to evaluate the accuracy of the ultrasonography with color Doppler, was accomplished for the years of 1990 to 2007, in the same databases described in the first systematic revision. Reference lists of all available primary studies were reviewed to identify additional relevant citation. The medical subjects heading (MeSH) and text words for the terms "*ovarian neoplasm"* and "*transvaginal ultrasound with color Doppler*" were combined with the MeSH term *diagnosis* ("sensitivity and specificity"). This review focused on observational studies evaluating clinically suspected adnexal masses, transvaginal probe ultrasonography with color Doppler in which the results of the diagnostic test of interest were compared with the results of a reference standard. Malignancy was suspected when the resistance index was ≤ 0.5. The final diagnoses of ovarian lesions were grouped and compared as malignant or borderline *versus* benign.

The systematic review, with focus diagnosis to evaluate the accuracy of CA 125, it was accomplished for the years of 1990 to 2007, in the same databases described in the first systematic revision. Reference lists of all available primary studies were reviewed to identify additional relevant citation. The medical subjects heading (MeSH) and text words for the terms "*ovarian neoplasm"* and "*CA 125*" were combined with the MeSH term *diagnosis* ("sensitivity and specificity"). This review focused on observational studies evaluating clinically suspected adnexal masses through the evaluation of CA 125 levels. Results of the diagnostic test of interest were compared with the results of a reference standard in all studies. The diagnostic test was CA 125 levels with a cutoff of 35 U/ml and the diagnostic reference was the result of histological analysis of standard paraffin-embedded sections. The final diagnoses of ovarian lesions were grouped and compared as malignant or borderline *versus* benign.

The two systematic reviews above (ultrasonography with color Doppler and CA 125) had their data extracted in an independent way for 4 reviewers that confronted the quality of the study and the quality of the extracted data. The extraction of the data includes to evaluate the clinic characteristics of each study. The methodological quality of each study was confronted through the methodological quality with reference to the *Oxford Centre for Evidence Based Medicine Level of*

Evidences Classification and by "*Scoring of Study Quality*". In the statistical analysis the agreement among the ultrasonography with color Doppler and CA 125 with the paraffin exam was calculated through the coefficient *Kappa*. For each study, 2 x 2 contingency tables were constructed, in which all biopsies were classified as normal or benign, borderline and malignant. The true-positive (TPR; sensitivity) and the false-positive rates (FPR; 1 – specificity) were calculated. The summary weighted sensitivity and specificity were calculated as the sum of sensitivities and specificities reported for each study, multiplied by the number of subjects in the study and divided by the total number of subjects in all studies. The 95% CIs for the mean weighted results were calculated using the exact method. The association between sensitivity and specificity for benign and borderline or malignant ovarian lesions was calculated using the *Spearman's correlation coefficient test*[13]. When there was no correlation, pooling sensitivities and specificities were calculated, since there were two categories of results (negative or positive test). In the case of correlation or heterogeneity between sensitivity and specificity, a summary receiver operating characteristic curve (SROC) was generated using data from all thresholds, by the Littenberg and Moses method. The diagnostic odds ratio (DOR) can relate to different combinations of sensitivity and specificity. The DOR describes the odds of the positive test results in participants with disease compared with the odds of positive test results in those without disease. A single diagnostic odds ratio corresponds to a set of sensitivities and specificities depicted by SROC. It can change according to the threshold and ROC curve used to define an abnormal examination resulting in the expected trade-off between sensitivity and specificity. Also, the area under the curve (AUC) can summarize the inherent capacity of a test for discriminating a diseased from a non-diseased subject. Perfect tests usually have AUCs close to 1 and poor tests usually have AUCs close to 0.5. The heterogeneity of the sensitivities and specificities of the studies were assessed using the Q_T (Cochran) test for χ^2 distributions with N -1 degrees of freedom.

Results: The systematic review with therapeutic focus - laparoscopy *versus* laparotomy for ovarian cancer of early stage ovarian cancer (stage I) by FIGO did not find any randomized study. Only, three observational studies were found, but with poor methodological quality. Therefore, it was not possible to perform a meta-analysis due to the difference among the studies. These three observational studies were classified by the authors as one being of a prospective cohort and the other two

studies as being case-controls. However, when analyzing the methodological drawing of the two studies classified as case-controls, they were, in fact, cohort studies. We cannot change the classification of the publication given by the authors and we maintained as case-controls, but we would like to point out that all three are, in fact, prospective cohort studies.

A quantitative systematic review was performed to estimate the diagnostic accuracy of frozen section in ovarian tumors selected fourteen primary studies which included 3659 women. For benign ovarian *vs.* borderline/or malignant tumor cases, occurrence of a positive frozen section result for benignity (pooled LR, 8.7; 95% CI, 7.3-10.4) and post-test probability for benignity diagnosis was 95% (95% CI, 94%-96%). A positive frozen section result for malignity *vs.* benign (pooled LR, 303; 95% CI, 101-605) increased the probability of ovarian cancer to 98% (95% CI, 97%-99%). In borderline *vs.* benign ovarian tumor cases, a positive frozen section result (pooled LR, 69; 95% CI, 45-106) increased the probability of borderline tumors to 79% (95% CI, 71%-85%). In borderline *vs.* malignant ovarian tumor cases, a positive frozen section result (pooled LR, 18; 95% CI, 13-26) increased the probability of borderline tumors to 51% (95% CI, 42%-60%).

A quantitative systematic review was performed to estimate the accuracy ultrasonography with color Doppler in the diagnosis of ovarian tumors selected twelve studies which included 2,398 women. The pooled sensitivity was 0.87 (IC95% 0.84-0.90); and the specificity was 0.92 (IC 95% 0.87-0.90). The DOR for ovarian cancer and borderline lesions *vs* benign lesions was 125 (95% CI, 55-283). SROC curves were constructed due to heterogeneity in the DOR. For malignant ovarian cancer and borderline *vs*. benign lesions the AUC was 0.9573.

A quantitative systematic review was performed to estimate the accuracy of CA 125 assay in the diagnosis of ovarian tumors selected seventeen studies which included 2,374 women. The pooled sensitivity was 0.80 (IC95% 0.76-0.82) and the specificity was 0.75 (IC 95% 0.73-0.77). The DOR for ovarian cancer and borderline lesions *vs* benign lesions was 21.2 (95% CI, 12-37). SROC curves were constructed due to heterogeneity in the DOR. For malignant ovarian tumors and borderline *vs*. benign lesions the AUC was 0.8877.

Conclusion: The systematic review with therapeutic focus - laparoscopy versus laparotomy for ovarian cancer by FIGO stage I (Ia, Ib,Ic) did not find evidence to support the use of laparoscopy for the management of early stage ovarian cancer as

routine clinical practice. Further studies with good quality are needed comparing laparoscopy surgery with laparotomy for the management of early stage ovarian cancer. As well as the importance of the correct classification of the study design for the authors, and for the reviewers of the papers, because we had two cohort studies incorrectly classified as being case-controls.

A quantitative systematic review performed to estimate the diagnostic accuracy of frozen section conclude that diagnostic accuracy rates for frozen section analysis is high for malignant and benign ovarian tumors, but for borderline tumors they remain relatively low.

A quantitative systematic review performed to estimate the accuracy ultrasonography with color Doppler is a useful pre-operative test for predicting the diagnosis of pelvic masses.

Also, the quantitative systematic review performed to estimate the accuracy of CA 125 level ≥ 35 U/ml concluded that this is a useful pre-operative test for predicting the benign or malignant nature of a pelvic mass.

In conclusion, these four studies of systematic reviews regarding the approach of ovarian tumors have shown low evidence levels laparoscopy in the early ovarian cancer ovarian (stage I by FIGO) and, also that the three diagnostic exams are really important in the presence of ovarian tumor, helping to define the diagnosis. Once the diagnostic is defined, it remains to adopt the better therapeutic option.

Lista de Quadro

Quadro 1 - Etapas da elaboração de uma revisão sistemática 25

Lista de Figura

Figura 1 - Sumário da curva ROC calculado a partir do método de Moses e Littenberger, referente à regressão linear mostrada acima 34

Apresentação

Este trabalho consiste na tese de doutorado apresentada ao Programa de Pós-Graduação em Epidemiologia da Universidade Federal do Rio Grande do Sul em 2008, sendo um estudo de revisão sistemática, com enfoque diagnóstico e terapêutico, das tumorações ovarianas com pressupostos de malignidade.

A revisão de literatura foi dividida em quatro partes: (1) revisão sistemática; (2) revisão sistemática com enfoque diagnóstico; (3) revisão sistemática de estudos observacionais com enfoque terapêutico; e (4) tumorações ovarianas.

Os resultados são apresentados em forma de quatro artigos. Três deles analisam, inicialmente, a acurácia diagnóstica do CA 125, da ultra-sonografia com Doppler colorido e do exame anatomopatológico de congelação das tumorações ovarianas com pressupostos de malignidade. O quarto artigo consiste de uma revisão sistemática com enfoque terapêutico, quanto à abordagem cirúrgica, se por laparoscopia ou por laparotomia, dos tumores ovarianos com pressupostos de malignidade nos estágios iniciais (Ia, Ib e Ic) pela Federação Internacional de Gincecologia e Obstetrícia (FIGO). As considerações finais discutem os principais achados dos artigos.

Documentos de apoio estão apresentados nos Anexos, incluindo o Projeto de Pesquisa (Anexo A).

1 INTRODUÇÃO

1 INTRODUÇÃO

Revisão sistemática realiza-se quando existe uma questão clínica que suscite dúvida, sendo seu objetivo principal ajudar na implantação de condutas validadas através da análise crítica dos estudos científicos (SACKETT & ROSENBERG, 1995; COOK, MULROW, HAYNES, 1997). Na área cirúrgica, existem diferentes estratégias para o tratamento de uma mesma doença, devendo-se avaliar os benefícios, os danos, os custos e a eficácia de cada procedimento escolhido (KREDER, 1999). Preferencialmente, a maior parte dessas informações deve proceder de revisões sistemáticas ou metanálises, a fim de validar a escolha da conduta cirúrgica com base nas melhores evidências científicas (REEVES, 1999; SAUERLAND. LEFERING, NEUGEBAUER, 1999).

Em cirurgia ginecológica, o tratamento dos tumores ovarianos é tema que suscita inúmeras controvérsias quanto à melhor abordagem – se por laparoscopia ou por laparotomia (CANIS *et al.*, 1994a; CHILDERS *et al.*, 1994; CHI & CURTIN, 1999), visto que os benefícios da cirurgia endoscópica são incontestáveis, como rápida recuperação, menos tempo de internação hospitalar e menor intensidade da dor no período pós-operatório (LORENZ *et al.*, 1999; TROIDL, 1999)

Entretanto, nos casos de neoplasia maligna de ovário, o manejo por essa via torna-se inadequado, em decorrência das pequenas incisões que dificultam os cuidados operatórios preconizados para essas neoplasias (BENEDET *et al.*, 2000). Além do que, o uso do gás carbônico, em cirurgia laparoscópica, propicia a disseminação e a implantação de células neoplásicas na cavidade abdominal (MÜLLER *et al.*, 1999; VOLZ *et al.*, 1999; SMIDT *et al.*, 2001; KÖHLER *et al.*, 2004). Assim, o inadequado diagnóstico e o tratamento de uma falsa tumoração benigna através da endoscopia são fatores de mau prognóstico para o câncer de ovário (WANG *et al.*, 1999).

No entanto, estudos de casos selecionados para as tumorações ovarianas malignas, em seus estágios iniciais (Ia,Ib e Ic) têm sido publicados com maior

freqüência nos últimos anos, dando a entender que a laparoscopia para as tumorações malignas iniciais (Ia, Ib e Ic), pela FIGO, seja o tratamento de escolha (CURTIN, 1994; CHI & CURTIN, 1999; DOTTINO, LEVINE, RIPLEY, 1999). Contudo permanece a controvérsia sobre ser a laparoscopia uma boa escolha para as tumorações ovarianas malignas em estágios iniciais (VERGOTE & TRIMBOS, 2003; VERGOTE & AMANT, 2004).

Quando uma tumoração ovariana é detectada, faz-se necessário estabelecer suas características de benignidade ou malignidade, sendo que determinados exames, de certa forma, guiam a conduta terapêutica a ser adotada, apesar de não possuírem 100% de acurácia. Esses exames são:

- A ultra-sonografia transvaginal que usa um sistema de escore baseado nos bordos internos da tumoração, na presença de septações, no tipo de ecogenecidade e no volume do ovário (SASSONE *et al.*, 1991). Ovários com volume superior a 20 cm^3 na idade reprodutiva e acima de 10 cm^3 em mulheres no período de pós-menopausa indicam necessidade de investigação (VAN, Jr. *et al.*, 2000).
- A ultra-sonografia com Doppler colorido que permite avaliar o fluxo sangüíneo nas tumorações através do cálculo do índice de resistência (IR) e do índice de pulsatilidade (PI), com pontos de corte para malignidade menor que 0,5 e 1 respectivamente (BROWN *et al.*, 1998).
- O marcador tumoral CA 125 com níveis suspeitos para malignidade quando superiores a 35 U/ml (MAGGINO *et al.*, 1987; MAGGINO *et al.*, 1994).
- Durante o ato cirúrgico, impõe-se, conforme o tipo de tumoração, o exame anatomopatológico de congelação que orienta o médico na melhor conduta terapêutica a ser seguida.

Devido às incertezas que suscitam controvérsias quanto à melhor abordagem inicial para o manejo operatório das tumorações ovarianas – se por laparoscopia ou por laparotomia –, planejamos realizar estudos de revisão sistemática para avaliar a acurácia diagnóstica: do CA 125, da ecografia transvaginal com Doppler colorido e do exame anatomopatológico de congelação durante o transoperatório, por serem os principais exames para auxiliar na escolha da melhor conduta cirúrgica a ser seguida nas tumorações ovarianas. Concomitantemente procedemos ao estudo de revisão sistemática comparando laparoscopia e laparotomia, na abordagem cirúrgica do câncer ovariano, em seus estágios iniciais (Ia, Ib e Ic), segundo a Federação Internacional de Ginecologia e Obstetrícia (FIGO).

2 REVISÃO DA LITERATURA

2 REVISÃO DA LITERATURA

2.1 Revisão sistemática

A revisão sistemática (RS) constitui método moderno para avaliação de um conjunto de dados simultâneos e permite ao pesquisador distinguir um tratamento eficiente daquele que não o é, resolver controvérsias em condutas e determinar terapêuticas que devam ser implementadas (EGGER, SMITH, PHILLIPS, 1997; MULROW, COOK, DAVIDOFF, 1997; WALTER & JADAD, 1999). Os protocolos de condutas (*guidelines*) estão preferencialmente baseados em estudos de revisões sistemáticas (COOK, MULROW, HAYNES, 1997), haja vista integrarem, de forma eficiente e racional, informações e dados que auxiliam em tomadas de decisões na área da saúde (MULROW, 1994; COOK, MULROW, HAYNES, 1997; MULROW, COOK, DAVIDOFF, 1997).

A RS envolve aplicação de estratégias científicas para sua elaboração, utilizando métodos específicos sistemáticos para identificar, selecionar e avaliar pesquisas relevantes, bem como coletar e analisar dados de estudos incluídos na revisão, tendo como objetivo principal limitar os vieses, devendo responder a uma pergunta específica que suscite controvérsia (OXMAN, COOK, GUYATT, 1994; COOK, MULROW, HAYNES, 1997; MULROW, COOK, DAVIDOFF, 1997). Para tanto, reúne, de forma organizada e crítica, grande quantidade de resultados de pesquisas clínicas (COOK, MULROW, HAYNES, 1997). Segundo esses mesmos autores, quando os resultados de estudos primários são sumarizados, mas não estatisticamente combinados, a revisão é chamada de “revisão sistemática qualitativa” (EGGER, SMITH, SCHNEIDER, 2001). Já a que utiliza métodos estatísticos para sumarizar seus resultados é denominada de “revisão sistemática quantitativa” ou metanálise (EGGER & SMITH, 1995; EGGER, SMITH, PHILLIPS, 1997, EGGER & SMITH, 1997; EGGER, EBRAHIM, SMITH, 2002).

A RS deve colocar de forma clara a questão a ser pesquisada e definir os critérios de inclusão e exclusão de estudos primários, o processo usado para identificar os estudos primários e os métodos utilizados para aferir a qualidade metodológica dos estudos selecionados, assim como os métodos empregados para sumarizar os resultados dos estudos primários nos quais a conclusão da revisão sistemática está baseada (EGGER & SMITH, 1995; COOK, MULROW, HAYNES, 1997; EGGER, SMITH, PHILLIPS, 1997; GREENHALGHb, 1997b).

No preparo de uma RS, não existe uma única fonte de busca de trabalhos, havendo necessidade de pesquisarem-se estudos relevantes em bancos de dados eletrônicos (*Medline*, *Embase, Lilacs*, *Cochrane Controlled Trials Database*), verificar as referências bibliográficas das publicações relevantes, solicitar estudos a colegas e pesquisar manualmente revistas e anais de congressos, sendo que para cada fonte utilizada deve-se detalhar o método adotado (OXMAN, COOK, GUYATT, 1994; COOK, MULROW, HAYNES, 1997; EGGER, SMITH, PHILLIPS, 1997; MULROW, COOK, DAVIDOFF, 1997). Um fator importante é a maneira com que é realizada a coleta de dados, sendo necessário que todas as variáveis sejam observadas nos estudos e resumidas, assim como as características dos métodos utilizados, dos participantes e dos desfechos clínicos, pois eles é que permitirão ou não a comparação entre os estudos selecionados (MULROW, 1994; EGGER & SMITH, 1995; COOK, MULROW, HAYNES, 1997; EGGER, SMITH, PHILLIPS, 1997; GREENHALGH, 1997b; MULROW, COOK, DAVIDOFF, 1997; JUNI *et al.*, 1999).

Nos trabalhos de RS com enfoque terapêutico, devem-se escolher os estudos randomizados, para evitar-se o viés de seleção, pois esse tipo de delineamento tipicamente mensura e compara diferentes eventos que estão presentes ou ausentes após os participantes receberam determinado tipo de intervenção (COOK, MULROW, HAYNES, 1997; EGGER, SMITH, PHILLIPS, 1997; JUNI *et al.*, 1999). O ideal é que a todos os participantes seja dada a mesma oportunidade de alocação tanto no grupo experimental quanto no controle, em outras palavras, o verdadeiro estudo randomizado não pode ser influenciado pelo pesquisador ou pelos participantes (COOK, MULROW, HAYNES, 1997; EGGER, SMITH, PHILLIPS, 1997; JUNI *et al.*, 1999).

Outras formas de viés em estudos randomizados ocorrem quando os resultados dos estudos são distorcidos, o que pode decorrer da pessoa que pratica a intervenção, da interpretação de quem recebe a intervenção ou do investigador que analisa e descreve o desfecho, sendo que a melhor maneira de evitar-se esse tipo de erro é praticar o cegamento dos pesquisadores e dos pacientes. Há, ainda, os vieses que ocorrem por transcrição errônea dos dados ou por violação do protocolo de pesquisa; pelo delineamento inadequado do estudo; pela disseminação equivocada dos resultados dos estudos em outros países (devido a problemas de interpretação ou tradução) e também pelo viés da publicação, isto é, conforme o resultado do manuscrito (direcionamento do achado), o editor aceita ou rejeita o estudo (OXMAN, COOK, GUYATT, 1994; COOK, MULROW, HAYNES, 1997; GREENHALGH, 1997b; EGGER, SCHNEIDER, DAVEY, 1998; EGGER, SMITH, STERNE, 2001; EGGER, EBRAHIM, SMITH, 2002).

Os estudos randomizados e as RS devem apresentar transparência na maneira de elaboração, condução e análise dos dados (MOHER, SCHULZ, ALTMAN, 2001). Por esse motivo, um grupo de epidemiologistas, estatísticos e editores da área biomédica propôs normas para elaboração de estudos de RS e de metanálise – *Quality of Reporting of Meta-analysis* (QUOROM), constituídas de 21 itens (*checklist*) que devem ser obedecidos na análise crítica de uma RS (MOHER *et al.*, 2000). Inicialmente, uma RS deve apresentar um fluxograma informando o número de artigos randomizados identificados bem como o número de artigos excluídos e as razões da exclusão. Os autores devem explicitar os critérios utilizados para análise da qualidade dos estudos, os bancos de dados utilizados para a pesquisa, os vieses encontrados e se houve restrição a idioma; e na discussão, comentar se os resultados obtidos podem ter sido influenciados por cada um dos vieses encontrados (COOK, MULROW, HAYNES, 1997; EGGER & SMITH, 1997; EGGER, SMITH, PHILLIPS, 1997; EGGER, SCHNEIDER, DAVEY, 1998; EGGER, EBRAHIM, SMITH, 2002).

Os estudos de revisão sistemática são diretamente influenciados pela qualidade dos trabalhos incluídos. Entretanto, a metodologia para análise da qualidade dos estudos incorporados em uma RS tem gerado muitos debates. Os aspectos relevantes da metodologia de cada um deles devem ser avaliados individualmente, e suas influências no tamanho de efeito analisadas (JUNI *et al.*, 1999). Deve-se verificar como foi feita a alocação, se havia cegamento na avaliação do evento, qual

o tempo de seguimento dos pacientes e qual o número de exclusões após ter sido realizada a alocação (com especificação dos motivos). Deve-se também avaliar se o paciente aderiu ao tratamento, se não ocorreu quebra do protocolo com violação dos critérios de elegibilidade e se a análise estatística baseou-se em "intenção de tratar" (*intention-to-treat*) (JUNI *et al.*, 1999). O uso de escores, para sumarizar os resultados, em escala de qualidade, é problemático. Um estudo de metarregressão para verificar se havia diferença entre as escalas utilizadas concluíu que o tamanho do efeito diminuía ou aumentava conforme a escala adotada. Nas revisões sistemáticas elaboradas na *Cochrane Library*, a validade da qualidade metodológica (interna e externa) e a forma de alocação dos estudos são categorizadas em: adequada (A), confusa (B) e inadequada (C) (Anexo B). Podem utilizar-se, também na avaliação da qualidade metodológica dos artigos, os níveis de evidência do *Oxford Centre Evidence Based Medicine*, produzido pelo Centro de Medicina Baseada em Evidências da Universidade de Oxford (MOHER *et al.*, 2000; PHILLIPS *et al.*, 2007) (Anexo C).

No *Medline*, os estudos randomizados encontram-se indexados como *randomized controlled trial.pt* ou *controlled clinical trial.pt* e, no *Embase,* como *randomized* controlled trial; utilizam-se ainda as seguintes palavras-chave nos dois bancos: *random*, *crossover* e *placebo*. Além dessas expressões deve-se empregar também *Medical Subject Headings* (*MeSH*) para localizar artigos relevantes a serem pesquisados (LEFEBVRE & CLARKE, 2001). Na análise da qualidade dos estudos randomizados, encontrados nos diferentes bancos de dados, é preciso observar se foram seguidas as normas para a condução desse tipo de trabalho – *Consolidated Standards of Reporting Trials* (CONSORT), constituídas de 22 itens (*checklist*) que devem ser obedecidos no que diz respeito a título, resumo, introdução, metodologia, resultados e discussão (MOHER, SCHULZ, ALTMAN, 2001).

Os bancos de dados do *Medline* e do *Embase* apresentam algumas características particulares que os diferenciam. O *Medline* contém mais de 10 milhões de artigos de revistas e a cada ano são acrescentadas mais 400.000 novas referências, abrange 3.900 revistas em 40 idiomas diferentes, tem suas estratégias de busca baseadas em palavras-chave específicas indexadas (*MeSH*), apresenta 76% dos artigos indexados com resumo em inglês, registra artigos desde 1966 até o presente, disponibiliza o acesso aos resumos através de CD-ROM e da internet (*PubMed*) sem custos e possui 52% das revistas publicadas nos Estados Unidos. Já o *Embase* contém oito milhões dos artigos de revistas e a cada ano são acrescentadas

mais 415.000 novas referências, abrange 4.000 revistas de 70 países, possui resumo em inglês em 80% dos artigos indexados, tem suas estratégias de busca usando palavras específicas (*thesaurus* – *EMTREE*), pode ser acessado através de CD e da internet e abrange 33% das revistas publicadas nos Estados Unidos, porém permite acesso somente por assinatura (LEFEBVRE & CLARKE, 2001).

Estas são algumas das vantagens da RS (GREENHALGH, 1997a; GREENHALGH, 1997b):

a) a utilização de metodologia científica que limita viés e identifica e rejeita estudos clínicos;

b) as conclusões são mais precisas e possuem maior acurácia devido aos métodos utilizados;

c) as conclusões dos estudos podem ser assimiladas rapidamente por outros pesquisadores e pelos responsáveis pelas políticas de saúde pública;

d) o tempo entre a descoberta do estudo e a efetiva implementação da estratégia diagnóstica e terapêutica pode ser reduzido;

e) os resultados de diferentes estudos podem ser formalmente comparados, permitindo generalizações quando neles houver consistência (homogeneidade);

f) as razões de heterogeneidade (estudos inconsistentes) podem ser identificadas e, a partir delas, novas hipóteses podem ser geradas, definindo novas áreas em que os ensaios clínicos fazem-se necessários;

g) a possibilidade de realizar-se metanálise de estudos com boa validade interna, aumentando-se com isso a precisão dos resultados e a redução do intervalo de confiança.

Foi elaborado um guia para decidir se uma RS pode ser considerada válida, sendo uma das primeiras orientações verificar se a revisão combina dados randomizados com não randomizados (OXMAN, COOK, GUYATT, 1994). Caso isso não ocorra, pode-se considerar uma revisão com conclusões limitadas e sujeita a viés. Após, deve-se verificar se, na revisão, encontram-se todos os artigos relevantes e sua validade interna e se a mesma especifica a magnitude e a precisão do efeito do tratamento. Deve-se também questionar se os resultados podem ser aplicados ao nosso paciente e se o tratamento pode ser executado por outra equipe e qual os reais benefícios e malefícios do tratamento que está sendo pesquisado.

Embora a RS seja realizada obedecendo a critérios rígidos, para minimizar vieses e maximizar precisão, no final o pesquisador pode concluir que os resultados dos estudos randomizados incluídos na revisão não podem ser combinados devido à sua heterogeneidade, o que pode ocorrer por várias razões (critérios de elegibilidade, intervenção, maneira de mensuração do desfecho, quantidade dos dados avaliados e qualidade metodológica) (COOK, MULROW, HAYNES, 1997; EGGER, SMITH, PHILLIPS, 1997; EGGER & SMITH, 1997; EGGER, SCHNEIDER, DAVEY, 1998; WALTER & JADAD, 1999).

Um estudo de RS deve ser cuidadosamente planejado, assim como qualquer outro projeto, com todos os detalhes descritos e com a formulação da questão da pesquisa preparada anteriormente ao início da RS (COOK, MULROW, HAYNES, 1997; EGGER, SMITH, PHILLIPS, 1997; EGGER & SMITH, 1997; EGGER, SCHNEIDER, DAVEY, 1998; WALTER & JADAD, 1999). Devem-se descrever os critérios de elegibilidade dos trabalhos que serão incluídos e a forma de avaliação da qualidade metodológica dos mesmos (Quadro 1). Na interpretação dos resultados, os revisores devem considerar os benefícios e os danos das intervenções de forma relativa e absoluta, analisando os aspectos econômicos e as implicações para estudos futuros (COOK, MULROW, HAYNES, 1997; EGGER, SMITH, PHILLIPS, 1997; EGGER & SMITH, 1997; EGGER, SCHNEIDER, DAVEY, 1998; WALTER & JADAD, 1999).

Quadro 1 - Etapas da elaboração de uma revisão sistemática

1 - Formular a questão da revisão (assunto que gera incerteza na conduta)
2 - Definição dos critérios de inclusão e exclusão:
• participantes
• intervenção e comparação
• desfechos
• delineamento do estudo e qualidade metodológica
3 - Estratégias de busca:
• The Cochrane Controlled Trials Register (CCTR)
• pesquisa em bancos de dados (Medline, Embase, Lilacs) informatizados e de estudos não registrados no CCTR
• verificação da lista de referências dos artigos
• pesquisa manual de revistas na área do estudo
• contato pessoal com pesquisadores na área do estudo
4 - Seleção de estudos:
• critérios de elegibilidade conferidos por mais de um observador
• desenvolver estratégias para resolver as discordâncias
• informar os artigos excluídos e as razões
5 - Avaliação da qualidade metodológica dos estudos:
• avaliação deve ser realizada por mais de um pesquisador
• usar uma lista simples (checklist), não adotar escala de qualidade
• sempre avaliar a maneira de alocação e cegamento dos dois grupos
• considerar o cegamento dos observadores para os autores, instituição e revistas
6 - Extração dos dados:
• desenhar um piloto para extração dos dados
• considerar extração dos dados por mais de um observador
• considerar cegamento dos observadores para os autores, instituição e revistas
7 - Apresentação e análise dos resultados:
• deve-se tabular os resultados dos estudos de forma individual
• examinar os gráficos (forest plot)
• explorar os causas possíveis de heterogeneidade
• considerar na metanálise todos os estudos ou subgrupos de cada estudo
• realizar a análise de sensibilidade examinando o gráfico do funil (funnel plots)
• disponibilizar uma lista com os estudos excluídos, caso o leitor tenha interesse
8 - Interpretação dos resultados:
• considerar as limitações, incluindo as publicações, e relatar os vieses
• considerar a força das evidências
• considerar a aplicabilidade
• considerar o número necessário para tratar (benefício e malefício)
• considerar as implicações econômicas
• considerar as implicações para estudos futuros

Fonte: EGGER & SMITH (2001)

* os itens de 1 até 7 devem estar descritos no protocolo

2.1.1 Revisão sistemática com enfoque diagnóstico

2.1.1.1 Histórico e definição

Testes diagnósticos são usados rotineiramente na medicina para rastreamento, diagnóstico e monitorização da progressão de uma doença. Informações diagnósticas são obtidas de diversas maneiras, incluindo história clínica, exame físico, análises bioquímicas, imagens e avaliações histológicas (SACKETT, 1969; SACKETT, 1978; SACKETT, 2002).

Estudos de RS com enfoque diagnóstico podem contribuir para avaliar o impacto de determinados testes diagnósticos e são essenciais para ajudar a escolher testes com ótima acurácia e que proporcionem apropriada interpretação dos resultados (DEEKS, 2001a).

RSs com enfoque diagnóstico são realizadas pela mesma razão com que são feitas as revisões com enfoque de intervenção: para produzir estimativas baseadas nas melhores evidências no assunto e para avaliar a variação dos achados entre os artigos (IRWIG *et al.*, 1994; IRWIG *et al.*, 1995; VAMVAKAS, 1998). Envolvem os mesmos estágios de uma RS com enfoque de intervenção, isto é, deve-se ter uma questão previamente definida e realizar estratégias de busca com as palavras-chave, sendo ampla a escolha dos bancos de dados (PUBMED, EMBASE, LILACS, CANCERLIT, CINAHL e Cochrane Library). Da mesma forma, deve-se avaliar os artigos em sua qualidade metodológica para elegibilidade, como também fazer uma lista de dados importantes extraídos de cada artigo (*checklist*), após escolher a melhor maneira para efetuar à síntese dos dados a serem analisados (DEEKS, 2001b).

Os estudos de acurácia diagnóstica diferem, em seu desenho clínico, dos estudos de intervenção, o que significa que se deve ter outra metodologia para avaliação da qualidade e do potencial viés de cada estudo. Com este intuito foi criado o STARD (*Standards for Reporting of Diagnostic Accuracy*) um *cheklist* com 25 itens que auxiliam na avaliação metodológica dos artigos com enfoque diagnóstico (BOSSUYT *et al.*, 2003). Com a mesma intenção, existem mais duas publicações, extremamente importantes para quem trabalha com RS com enfoque diagnóstico, que proporcionam ao pesquisador respostas mais precisas na avaliação

dos artigos com enfoque diagnóstico tendo como finalidade avaliar a inclusão do artigo ou não na RS (LIJMER *et al.*, 1999; BOSSUYT *et al.*, 2003) (Anexo D).

Uma RS com enfoque diagnóstico deve excluir estudos que não possuem os critérios de qualidade metodológica e que são mais suscetíveis a vieses, ou então incluir esses estudos e explorar as características e as diferenças entre os mesmos (DEEKS, 2001b). O poder estatístico é raramente discutido em estudos com enfoque diagnóstico, por não comparem dois grupos e não possuírem formalmente uma hipótese diagnóstica. Entretanto, o aumento do tamanho da amostra no somatório geral dos resultados aumenta a precisão da estimativa do teste diagnóstico (DEEKS, 2001b).

Estudos de teste de acurácia comparam os resultados entre dois grupos, os pacientes com a doença e os sem a doença, sendo que estes podem apresentar teste diagnóstico positivo ou negativo, sugerindo ter a doença ou não. O resultado do teste diagnóstico é comparado com exame considerado de referência para determinada doença. Tem-se, então, a mensuração da sensibilidade, da especificidade e dos resultados falso-positivos e falso-negativos. A partir destes dados podem-se calcular a razão de verossimilhança positiva e negativa, o valor preditivo positivo (probabilidade pós-teste positiva) e o valor preditivo negativo, bem como *diagnostic odds ratio* (DOR) (GO, 2001).

2.1.1.2 Estratégias de busca de artigos com enfoque diagnóstico

A identificação dos estudos para RS envolve a busca eletrônica dos bancos de dados, a busca manual em revistas não indexadas e a escolha de referências encontradas nos artigos solicitados. Normalmente são utilizados os seguintes termos de busca para RS com enfoque diagnóstico (DEEKS, 2001b):

Medical Subject Headings (MeSH terms)

- *explode “sensitivity and specificity”/;*
- *explode “mass screening”;*
- *“predictive-value-of teste”;*
- *“ROC-curve”;*
- *specificit*;*

- *false negative*;*
- *accuracy;*
- *screeniny;*
- *sensitivity;*
- *predictive value*;*
- *likelihood ratio*;*
- *diagnosis**

As palavras-chave acima são associadas às palavras-chave do que se deseja pesquisar, como, por exemplo, tumoração ovariana (*sub-headings*):

- Ovarian cysts
- ovar*[tw] AND tumo*[tw]
- ovar*[tw] AND cancer [tw];
- Adnexal diseases [mh];
- Ovarian neoplasms [mh];
- Pelvic*[tw] AND tumo* [tw];
- Pelvic*[tw] AND masses*[tw];
- “CA 125”;
- “Serum CA125”;
- frozen section;
- "color"[mh] AND Doppler [all fields];
- colour” [tw] AND Doppler [all fields];
- "ultrasonography"[Tw] AND Doppler [all fields];
- "color" [mh] AND transvaginal[All Fields];
- flow[All Fields] AND Doppler[All Fields]).

No Anexo A, encontram-se as estratégias de busca utilizadas para os três artigos com enfoque diagnóstico que foram utilizadas no PUBMED e no EMBASE, CINAHL.

2.1.1.3 Seleção dos estudos

Estudos ideais para inclusão de uma RS com enfoque diagnóstico são aqueles em que os pacientes, recrutados de uma população ampla, sejam selecionados de maneira consecutiva e cegados tanto o pesquisador quanto os pacientes. Entretanto, geralmente encontram-se estudos não cegados, retrospectivos, com pacientes oriundos de uma população restrita (LIJMER *et al.*, 1999). É de extrema importância que as amostras dos estudos sejam selecionadas de grupos de saúde com aspectos semelhantes (DEEKS, 2001a).

Todos os estudos devem ter o mesmo teste diagnóstico como teste de referência (padrão-ouro), sendo a referência padrão em todos os estudos selecionados para a RS com enfoque diagnóstico. É imprescindível que os dois testes, o da pesquisa e o de referência, sejam mensurados de maneira independentes (DEEKSb, 2001a).

Outro aspecto importante quando se inclui um estudo é que ele deva ter descrição clara do teste de referência e do teste experimental, que é o enfoque do estudo de acurácia diagnóstica. Devem estar descritos os desfechos positivos e negativos, as características demográficas da população em estudo, as co-morbidades e o histórico de cada paciente (DEEKS, 2001a).

2.1.1.4 Metanálise nos estudos de acurácia diagnóstica

A metanálise nos estudos de acurácia diagnóstica deve ser considerada somente quando os estudos tenham sido recrutados de população clinicamente semelhante, usando o mesmo teste diagnóstico comparando-o com o teste considerado padrão-ouro e preferencialmente que os estudos sejam prospectivos e cegados. Embora, muitas vezes, estes critérios sejam obedecidos na seleção dos artigos, ocorre heterogeneidade dentro da metanálise, porque nem todos os artigos apresentam alto nível de evidência científica (LIJMER *et al.*, 1999; RICHARDSON *et al.*, 1999; DEEKS, 2001a). Conforme o grau de heterogeneidade encontrado na metanálise e conforme os testes estatísticos realizados na investigação da heterogeneidade, muitas vezes torna-se impossível sumarizar o resultado da RS em

dados únicos, por exemplo, para: sensibilidade, especificidade e DOR (DEEKS, 2001a).

A metanálise passa por dois tipos de processo para realizar os somatórios dos resultados dos estudos de teste de acurácia. Inicialmente deve-se elaborar em cada estudo uma tabela 2 x 2 e depois realizar, através de sistema de computação, com o *software* especifico Meta-DiSc® (version Beta 1.1.1), o sumário estatístico de todos os estudos incluídos na revisão sistemática (DEEKS, 2001a; ZAMORA *et al.*, 2006).

A escolha do método estatístico para combinar os estudos depende da heterogeneidade observada entre os estudos incluídos na metanálise. O grau de heterogeneidade pode ser observado pelos gráficos que sumarizam a sensibilidade e a especificidade a partir dos resultados encontrados em cada estudo e através do sumário da curva ROC *(Receiver operator characteristic).* Divergências entre os estudos são esperadas, mas, conforme o tipo de seleção dos pacientes e o desenho clínico de cada estudo, a variabilidade ou heterogeneidade pode assumir valores importantes e deve se investigada através de cálculos estatísticos específicos (DEEKS, 2001a). Essa variabilidade leva à introdução da variação no limiar do teste diagnóstico. Os estudos então incluídos na RS podem usar diferentes limiares para definir se o resultado do teste é positivo ou negativo.

2.1.1.4.1 Somatório da sensibilidade e especificidade na metanálise diagnóstica

Um método simples é combinar estudos de acurácia diagnóstica e computar o peso das médias da sensibilidade e da especificidade de cada estudo. Entretanto, esse método só deve ser utilizado se não houver variabilidade no limiar do teste diagnóstico. Quando existe heterogeneidade, ela pode ser investigada através da curva ROC. A homogeneidade da sensibilidade e especificidade pode ser avaliada pelo teste Q_T (Cochran) distribuição do χ^2 com 1 grau de liberdade. E o cálculo da correlação entre sensibilidade e especificidade é realizado através do teste de correlação de *Spearman*. Quando não existe correlação, o somatório das sensibilidades e das especificidades é calculado da seguinte forma, tendo como

resultado dois tipos de categoria (positiva e negativa) (DEEKS, 2001b):

$$\text{sensiblidade } pi = \frac{Yi}{ni}\text{; especificidade } pi = \frac{Yi;}{ni}$$

Para todos os estudos se calcula a sensibilidade e a especificidade, para, após, sumarizar o resultado usando aproximação com inverso da variância. A estimativa total para sensibilidade e especificidade é calculada de acordo com a seguinte fórmula (DEEKS, 2001a):

$$p = \frac{\Sigma Yi}{\Sigma ni}$$

Onde ΣYi é o somatório dos verdadeiros positivos (sensibilidade) ou somatório dos verdadeiros negativos (especificidade), e Σni é o somatório de doentes (sensibilidade) ou não doentes (especificidade) (DEEKS, 2001a).

O erro padrão é computado através da fórmula (DEEKS, 2001a):

$$SE(p) = \sqrt{\frac{p(1-p)}{\Sigma n_1}}$$

Se os cálculos acima mostram homogeneidade e ausência de correlação entre sensibilidade e especificidade, usam-se modelos de efeitos fixos para encontrar o sumário da sensibilidade e da especificidade bem como da DOR (DEEKS, 2001a).

Entretanto, quando existir heterogeneidade pelo teste Q_T (Cochran) distribuição do χ^2 com 1 grau de liberdade, mas a correlação de *Spearman* for negativa entre especificidade e sensibilidade (DEEKS, 2001a), os cálculos sumarizados da sensibilidade, especificidade e da DOR podem ser apresentados realizando-se os cálculos através do modelo de afeitos aleatórios (DEEKS, 2001a). Neste caso, o cálculo para sumarizar os dados da sensibilidade e da especificidade é feito, por exemplo, utilizando-se o valor da sensibilidade de cada estudo e multiplicando-se o valor pelo número de sujeitos de cada estudo; após realiza-se o somatório dos valores acima descritos e divide-se pelo número total dos sujeitos de todos os

estudos. A mesma fórmula utiliza-se para cálculo da especificidade (SMITH-BINDMAN *et al.*, 1998):

$$\Sigma[(\text{sensibilidade}_{i)x}\, n_i) \,/\, (n_1+n_2+n_3+\ldots\ldots\, n_i).$$

2.1.1.4.2 DOR (*diagnostic Odds Ratio*)

A DOR é uma conveniente mensuração quando se combinam estudos de uma RS. Resume acurácia diagnóstica descrevendo muito bem como o teste sumariza os dados da doença e de quem não tem a doença (sensibilidade e especificidade). Entretanto, é difícil aplicar-se diretamente na prática clínica. Uma maneira simples de sumarizar-se a DOR é através da curva ROC que corresponde à sensibilidade e à especificidade encontradas na metanálise com enfoque diagnóstico (DEEKS, 2001a; DEEKS, 2001a e b). O valor da DOR igual a 1 indica que o teste não tem poder discriminatório, sendo que quanto maior for o valor da DOR maior será o grau de relevância do teste diagnóstico.

$$\text{DOR} = \frac{\text{sensiblidade}}{\text{1-Sensibilidade}} \Big/ \frac{\text{1-especificidade}}{\text{especificidade}}$$

Outra maneira simples de ser calculada a DOR é com o uso de tabela 2 x 2. No caso de alguma casela conter valor 0, adiciona-se nesta nela 0,5 e em todas as demais também realiza-se essa adição, para evitar problemas no cálculo final (DEEKS, 2001a; DEEKS, 2001b; DEEKS, ALTMAN, BRADBURN, 2001).

$$\text{DOR} = \frac{\text{verdadeiros positivos x verdadeiros negativos}}{\text{falso-positivos x falso-negativos}} = \frac{\text{A x D}}{\text{B x C}}$$

Quando se observa heterogeneidade entre os estudos, observa-se concomitante variação no limiar do teste diagnóstico, e com isso as estimativas de sumário de especificidade e sensibilidade são superestimadas. Nesta situação, o apropriado sumário da metanálise, não é somente um simples ponto na curva ROC, mas o sumário da curva ROC (SROC) que deve ser analisada como um todo, sendo um balanço final entre a sensibilidade e a especificidade. Com a curva ROC pode-se também estimar a área sob a curva, conhecida como AUC (*area under the curve*) que pode sumarizar a capacidade do teste discriminatório para doença. Testes perfeitos usualmente possuem AUCs próximos ao valor 1, e testes considerados não satisfatórios possuem valor igual ou inferior a 0,5. A área sob a curva pode sumarizar a capacidade inerente de um teste diagnóstico ou de um biomarcador para discriminar doença de não doença (TOSTESON & BEGG, 1988; LITTENBERG & MOSES, 1993; MIDGETTE, STUKEL, LITTENBERG, 1993; DEEKS, ALTMAN, BRADBURN, 2001).

A curva ROC assimétrica ocorre quando a DOR muda com o limiar diagnóstico da sensibilidade e da especificidade. Para estimar-se o sumário da curva ROC usa-se um modelo linear de Littenberg e Moses com a seguinte fórmula (LITTENBERG & MOSES, 1993; SUTTON *et al.*, 2000; DEEKS, ALTMAN, BRADBURN, 2001):

D = $a+\beta S$
D = *ln* DOR
S = logiti verdadeiros positivos + logiti de falso-positivos
A = *ln* OR (intercepto)
β = examina quanto a OR depende do valor da variação do limiar

Caso o coeficiente de regressão fique próximo do zero, a acurácia de cada estudo primário pode ser sumarizada por OR, dada pelo intercepto. Pode-se avaliar também de outra forma, caso o valor de *p* do intercepto β seja menor que 0,05, ou com valor muito próximo a este, o que sugere que a DOR muda com o limiar, ou seja, não se pode ter um valor único que represente aquele teste diagnóstico. Entretanto, se o valor de *p* do intercepto β for maior que 0,1 isso sugere que a DOR não modifica seu limiar, mesmo havendo heterogeneidade entre os estudos, e pode-se ter um valor único para DOR (LITTENBERG & MOSES, 1993; SUTTON *et al.*, 2000; DEEKS, ALTMAN, BRADBURN, 2001). Abaixo indica-se o cálculo da regressão e a curva ROC correspondente a esse cálculo.

Método de Moses e Littenberger

Parâmetro	Intercepto	Erro padão	T	P
a	3,059	0,324	9,4	0,00
β	-0,203	0,165	1,233	0,23

Observa-se que o coeficiente de regressão em β está próxima de zero e tem seu valor de $p > 0.1$

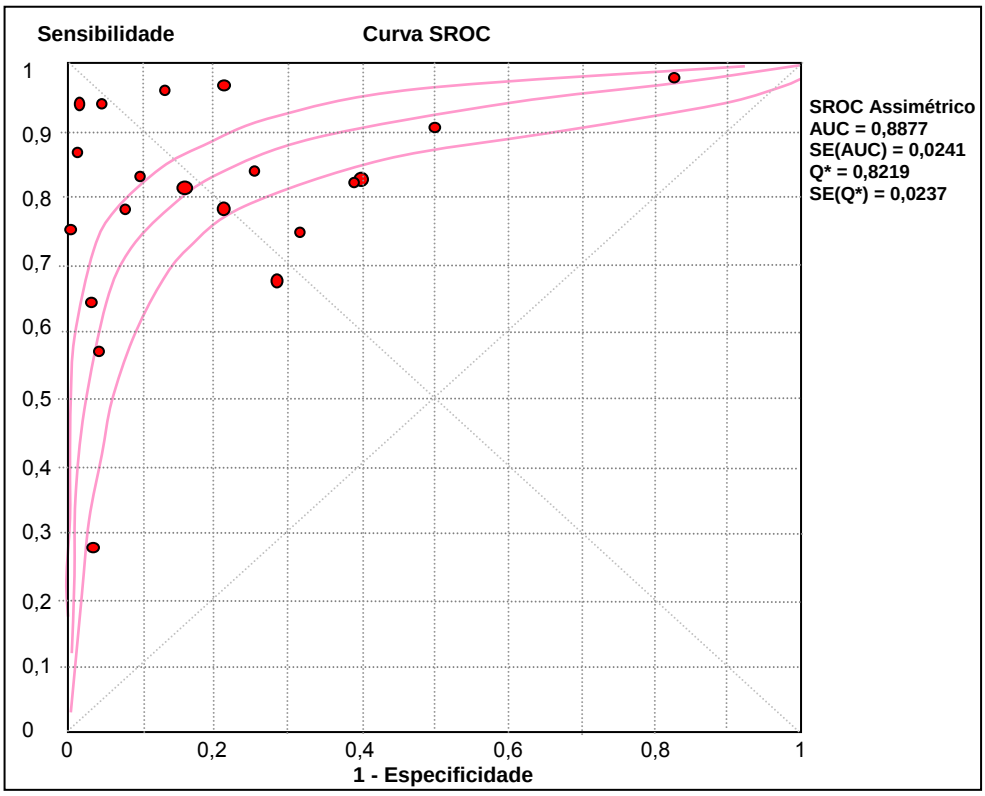


Figura 1 - Sumário da curva ROC calculado a partir do método de Moses e Littenberger, referente à regressão linear mostrada acima, com intevalo de confiança de 95%.

2.1.2 Revisão sistemática de estudos observacionais com enfoque terapêutico (qualitativa)

RS e metanálises de estudos observacionais são mais comuns que as advindas de estudos randomizados. Cerca de 40% de metanálises são provenientes de estudos de coorte e de casos e controles (EGGER, SMITH, STERNE, 2001). Entretanto, confundimento e viés freqüentemente distorcem os achados dos estudos observacionais. Existe um grande risco de uma metanálise realizada a partir de estudos observacionais produzir resultados precisos, mas na realidade ilegítimos (EGGER & SMITH, 1995; EGGER, SCHNEIDER, DAVEY, 1998).

É importante ressaltar que a maneira de realizar uma metanálise de estudos observacionais obedece aos mesmos critérios de elaboração da RS com enfoque de intervenção de estudos randomizados. Os passos da elaboração encontram-se resumidos no Quadro 1, seção 2.1. Na tentativa de se produzirem RSs a partir de estudos observacionais com boa qualidade metodológica, foi criado um instrumento para auxiliar o pesquisador intitulado MOOSE (*Meta-analysis of Observational Studies in Epidemiology*), que possui 35 itens que o pesquisador deve seguir na elaboração da RS. Entretanto, deve-se ter bem claro que, muitas vezes, a metanálise poderá não ser realizada porque os dados estatísticos não combinam ou passam a ser inapropriados pelos fatores de confundimento não corrigidos nos estudos, tornando os dados estatísticos sumarizados enviesados e, desta forma, impossível de serem demonstrados (EGGER, SMITH, SCHNEIDER, 2001). Realiza-se então a RS qualitativa, mostrando e discutindo os resultados sem apresentar a metanálise. (Anexos E).

Com o intuito de diminuir os dados enviesados advindos de uma RS observacional foram criados instrumentos para auxiliar na avaliação da qualidade metodológica de cada um dos artigos inclusos na revisão. Esses instrumentos são o *Newcastle-Ottawa Scale* (NOS), para estudos de coorte e de casos e controles e, mais recentemente, o STROBE (*Strengthening the Reporting of Observational Studies in Epidemiology*), que possui 22 itens a serem considerados dentro de cada um dos artigos para se avaliar a qualidade metodológica, tanto para estudos de coorte como para estudos de casos e controles (VON *et al.*, 2007; WELLS *et al.*, 2007) (Anexos F e Anexo G).

Estudos não randomizados incluídos como casos e controles e estudos de coorte, podem ter sua qualidade avaliada através do NOS, que foi um instrumento elaborado pelas Universidades de Newcastle e Ottawa justamente para auxiliar na qualidade metodológica de cada artigo para inclusão ou não em um estudo de metanálise. Esse instrumento usa um sistema de julgamento chamado *star system*, considerando três tipos aspectos: seleção dos estudos, comparabilidade dos grupos e avaliação do desfecho de interesse (WELLS *et al.*, 2007) (Anexo F).

Já o STROBE é mais profundo na avaliação da qualidade metodológica dos artigos, principalmente na parte estatística, avaliando se houve ou não correção dos fatores de confundimento e se ocorreu ou não interação entre as variáveis e o desfecho (VON *et al.*, 2007) (Anexo G).

Devido ao viés implícito de uma RS de estudos observacionais, a discussão deve relatar o viés e os fatores de confundimento ou interação encontrados em cada artigo e como ocorreu a avaliação dos artigos através dos instrumentos (NOS, STROBE). A conclusão de uma revisão sistemática de um estudo observacional deve fazer considerações sobre os resultados observados e verificar se eles podem sofrer generalização (STROUP *et al.,* 2000).

2.2 Tumores ovarianos malignos

2.2.1 Dados epidemiológicos

A neoplasia maligna de ovário é responsável por 4% de todos os cânceres na mulher, sendo a segunda causa de morte por câncer ginecológico e a quarta entre todos os tipos de câncer feminino, com estimativa de 23.400 novos casos e 13.900 mortes nos países do leste Europeu e nos Estados Unidos para 2001 (VOLZ, KÖSTER, SCHAEFF, 1997; VOLZ *et al.*, 1998; VOLZ *et al.*, 1999; EKERHOVD *et al.*, 2001). Sua incidência aumenta com a idade, haja vista que, entre 15 e 39 anos, é de 3,2 casos/100.000 (YANCIK, 1993; OSMERS *et al.*, 1998; NELSON, EKBOM, GERDIN, 1999), e acima de 50 anos o índice é de 43,6/100.000 (YANCIK, 1993), atingindo pico máximo, entre 70 e 84 anos, de 54/100.000 (YANCIK, 1993). Mulheres acima de 50 anos, com tumoração ovariana, possuem risco relativo (RR) de 8,6 com IC 95% (6,3 - 11) para câncer ovariano (KOONINGS *et al.*, 1989).

No Brasil, entretanto, não há uma definição exata da real situação epidemiológica das neoplasias malignas de ovário. Dados publicados pelo Instituto Nacional de Câncer em 2007 as deixam entre as de baixa incidência, ficando além do 11º lugar dentre aquelas que acometem as mulheres (PASPALICCHIO *et al.*, 2000; INCA, 2007).

A grande problemática nas tumorações malignas de ovário é a maneira insidiosa com que a doença se instala (YANCIK, 1993). Os sintomas são difusos e inespecíficos nos estágios iniciais, haja vista que em dois terços dos casos o diagnóstico é realizado nos estágios III e IV – segundo classificação da *International Federation of Gynecology and Obstetrics* (FIGO) (YANCIK, 1993; BENEDET *et al.*,

2000) (Anexo H). Nesses estágios mais avançados, o índice de sobrevida é de 23% em cinco anos, em contraste com o de 90% quando a detecção é realizada no estágio I (CRAYFORD *et al.*, 2000). Já em casos de neoplasias com baixo potencial de malignidade, evidenciam-se índices de sobrevida de 88,2% e 76,5% em cinco e 10 anos respectivamente, muito provavelmente pelo fato de o diagnóstico ser feito em 59,4% dos casos no estágio I (FIGO) (ELTABBAKH *et al.*, 1999).

Na grande maioria das mulheres com cistos ovarianos benignos, ou com baixo potencial de malignidade e/ou malignos, não há como identificar o fator de risco, pois a patogênese da doença é indefinida. Menarca precoce (antes de 13 anos), nuliparidade, câncer de colo e/ou endometrial, história de rubéola na infância, infertilidade, tratamento para infertilidade, terapia de reposição hormonal, uso de talco em região genital, síndrome dos ovários policísticos e obesidade constituem fatores de risco menores (EDMONDSON, 2001).

O fator de risco maior para desenvolvimento das tumorações malignas de ovário é a predisposição genética, visto que 10% dos tumores epiteliais de ovário são devidos à mutação do gene BRCA 1 localizado no cromossomo 17q21 e responsável por cerca de 90% dos cânceres de ovário hereditário (BERCHUCK *et al.*, 1999).

2.2.2 Classificação histológica dos tumores de ovário

O ovário é uma estrutura complexa do ponto de vista embriológico, fisiológico e histológico (BARBER, 1984). Nas últimas décadas, inúmeros estudos têm sido conduzidos buscando evidências de que a neoplasia benigna de ovário possa adquirir potencial para transformar-se em lesão maligna (SCULLY, 2000), haja vista terem sido identificadas lesões benignas, limítrofes e malignas na mesma peça cirúrgica e/ou presença de áreas com displasia na superfície epitelial ovariana ou em cistos de inclusão, mostrando que essas alterações são resultado da expansão clonal em diferentes etapas (SCULLY, 1995; SCULLY, 2000), podendo ter uma velocidade de transformação de até 15 anos de evolução (SCULLY, 1995).

A classificação dos tumores de ovário preconizada pela Organização Mundial da Saúde é feita de acordo com a origem histológica do tecido (SCULLY, 1999; BENEDET *et al.*, 2000) (Anexo I). Observa-se que 90% das tumorações malignas de ovário são de origem epitelial (YANCIK, 1993; SCULLY, 1999; BENEDET *et al.*, 2000), sendo 50% a 60% dos tumores epiteliais benignos e 25% malignos (SCULLY, 1999; BENEDET *et al.*, 2000). **As tumorações ovarianas de origem epitelial** podem ser subdividas em (SCULLY, 1999; BENEDET *et al.*, 2000):

- Tumores epiteliais serosos de ovário (30% de todos os tumores ovarianos), sendo 75% benignos, 13% com malignidade limítrofe e 25% malignos.
- Tumores epiteliais mucinosos de ovário, que respondem por 25-30% dos tumores epiteliais. Aproximadamente 80% são benignos, 10% possuem baixo potencial de malignidade e 10% são malignos.
- Tumores epiteliais endometrióides de ovário, predominantemente malignos, respondem por cerca de 15% -20% dos tumores epiteliais, sendo em 30% bilaterais.
- Tumor epitelial de células claras são malignos são mais raros e estão associados a endometriose ovariana.
- Tumores epiteliais de células transicionais de ovário, formados por células transicionais do sistema urogenital, podem ser de baixo potencial de malignidade.
- Tumores epiteliais de células escamosas de ovário, mais raros, têm como origem células do epitélio germinativo.
- Tumores epiteliais mistos são compostos pela combinação de um ou mais dos seis tipos cima.
- Carcinomas indiferenciados epiteliais de ovário, extremamente malignos, acometem mulheres jovens abaixo de 23 anos, e 90% das pacientes morrem em 1 ano.

Tumores das células da granulosa e da teca do ovário podem ser classificados em dois subtipos principais: juvenis e adultos. São de baixo potencial de malignidade e podem ser subdivididos em (SCULLY, 1999; BENEDET *et al.*, 2000):

- Fibromas ou fibrotecomas, relativamente comuns, correspondem a cerca de 4% de todos os tumores de ovário. Tumores com mais de 6 cm associam-se a ascite em 40% dos casos, havendo associação também com hidrotórax.

- Tumores estromais de ovário da célula de Leydig Setoli, extremamente raros, correspondem 0,2% dos casos de neoplasia ovariana.

Tumores de células germinativas correspomdem a 15-20% de todas as neoplasias ovarianas e em 10% dos casos são bilaterais. Cerca de 97% deles são benignos. Podem ser subdivididos em (SCULLY, 1999; BENEDET *et al.*, 2000):

- Teratomas podem ser maduros e imaturos. Quando imaturos, podem estar associados a malignidade (raro). O tipo maduro benigno acomete mulheres com idade média de 32 anos.
- Disgerminomas respondem por cerca de 2% de todos os tumores malignos de ovário, são freqüentemente sólidos e 75% das pacientes acometidas encontram-se entre 10 e 30 anos de idade.

2.2.3 Diagnóstico das tumorações ovarianas

O grande problema, no diagnóstico das neoplasias malignas ovarianas, é a forma insidiosa de instalação da doença, haja vista que os casos potencialmente curáveis (Anexo H) não ocasionam sintomas em sua grande maioria (VAN Jr. *et al.*, 2000). Desta forma, em pacientes com história e/ou exame físico sugestivos de tumoração anexial, há necessidade de investigação através de ultra-sonografia (pélvica e transvaginal) e, conforme as características morfológicas encontradas, devem-se associar estudo ultra-sonográfico com Doppler colorido e determinação sérica do marcador tumoral CA 125 (KURJAK *et al.*, 1989; DePRIEST *et al.*, 1994; DUFFY *et al.*, 2005). Caso os exames acima sugiram tratar-se de uma neoplasia maligna ovariana, é de suma importância a realização de outros exames de imagem para a investigação da tumoração pélvica, como tomografia computadorizada, ressonância magnética, urografia excretória, cistoscopia e colonoscopia (CURTIN, 1994).

2.2.3.1 Ultra-sonografia transvaginal e com Doppler colorido

A ultra-sonografia transvaginal possibilita uma melhor avaliação da tumoração ovariana pelo fato de seus transdutores possuírem maior freqüência (5-10 MHz) (GRANBERG, 1991; CURTIN, 1994). A imagem ultra-sonográfica do ovário foi descrita como uma estrutura ovóide que apresenta ecogenecidade uniforme (CURTIN, 1994; VAN Jr. *et al.*, 2000). Seu volume é calculado a partir dos valores máximos dos diâmetros ântero-posterior, transverso e longitudinal multiplicados entre si [D1xD2xD3] e por um valor constante de 0,523. Volumes uterinos iguais ou superiores a 20 cm^3 ou 10 cm^3, respectivamente, na menacma e na pós-menopausa é sugestivo de lesão ovariana a ser investigada (VAN Jr. *et al.*, 2000).

Foram propostos critérios morfológicos ultra-sonográficos para distinguir tumorações ovarianas benignas de malignas, que são: volume, ecogenecidade, característica interna e espessamento da parede e dos septos da tumoração, os quais possibilitam criar um escore tumoral (SASSONE *et al.*, 1991; VAN Jr. *et al.*, 2000). Escore morfológico superior a cinco é sugestivo de malignidade da lesão (DePRIEST *et al.*, 1994).

Para diminuir os resultados falso-positivos desses escores, há necessidade da associação com ultra-sonografia com Doppler colorido, com a finalidade de verificar a velocidade média do fluxo de sangue nos vasos durante a diástole, pois o processo de angiogênese é intenso no crescimento tumoral maligno. A ultra-sonografia com Doppler mostra-se um método mais fidedigno na avaliação das tumorações ovarianas com pressupostos de malignidade (BOURNE *et al.*, 1989; KURJAK *et al.*, 1989; BOURNE, 1991; BROWN *et al.*, 1998).

O exame permite calcular o índice de pulsatilidade (IP) e o índice de resistência (IR) dos vasos da tumoração ovariana, para diagnóstico diferencial entre lesão benigna e maligna. Considera-se como ponto de corte para o IP valor inferior a um e, para o IR, valor inferior a 0,5 para as tumorações ovarianas com pressupostos de malignidade (BOURNE *et al.*, 1989; KURJAK *et al.*, 1989; BOURNE, 1991; KURJAK *et al.*, 1992; KURJAK & PREDANIC, 1992; WEINER *et al.*, 1992; KURJAK *et al.*, 1993a; KURJAK *et al.*, 1993b; BROWN *et al.*, 1998; KURTZ *et al.*, 1999).

Resultados falso-positivos podem ocorrer em processos benignos, como cistos ovarianos luteínicos ou processos inflamatórios ovarianos que podem ter IR igual ou

inferior a 0,5 (BOURNE, 1991; WU *et al.*, 1994). Outra causa comum de resultados falso-positivos para malignidade, no exame com Doppler, é a presença de endometriose ovariana ou cisto dermóide, pois seus vasos possuem fluxo com baixa impedância (BOURNE, 1991; WU *et al.*, 1994; KURJAK & KUPESIC, 1995). Desta forma, foi proposto que, na suspeita de endometriose ovariana, fosse utilizado um critério de associação entre dados clínicos (idade, história clinica, dismenorréia e infertilidade) e dosagem sérica do marcador tumoral CA 125 (KURJAK & KUPESIC, 1995).

2.2.3.2 Dosagem de marcador CA 125

O marcador tumoral CA 125 é uma glicoproteína de alto peso molecular que se torna um determinante antigênico ao reconhecer o anticorpo monoclonal *cancer 125* (OC125), expresso na superfície de alguns dos tipos de tumores ovarianos (KABAWAT *et al.*, 1983)

O marcador CA 125 é o mais utilizado na investigação diagnóstica das tumorações ovarianas. Encontra-se elevado em 86% das pacientes com neoplasia maligna ovariana, principalmente as de origem epitelial (KABAWAT *et al.*, 1983; BRIOSCHI *et al.*, 1987; VERGOTE, BORMER, ABELER, 1987). Na avaliação da dosagem sérica do CA 125 considera-se como ponto de corte, para sugerir alteração, valor máximo de 35 U/ml (KABAWAT *et al.*, 1983; EINHORN *et al.*, 1986; BRIOSCHI *et al.*, 1987; MAGGINO *et al.*, 1987; EINHORN *et al.*, 1989; SOPER *et al.*, 1990; MAGGINO *et al.*, 1994; TWICKLER *et al.*, 1999; DUFFY *et al.*, 2005).

Entretanto, a dosagem sérica do CA 125 é limitada por fatores como: tipo histológico da neoplasia, tamanho tumoral e estágio da doença e pela natureza da tumoração, se benigna ou com malignidade limítrofe ou maligna (JACOBS & BAST, 1989).

Em síntese, ultra-sonografia transvaginal com Doppler colorido associada a dosagem sérica do CA 125 e a pós-menopausa possuem grande poder discriminatório para sugerir a natureza biológica da tumoração – se benigna ou maligna (KURJAK & KUPESIC, 1995; ASLAM *et al.*, 2000). Tais exames, portanto, são exames extremamente importantes para orientação na tomada de decisão da abordagem cirúrgica – se por laparoscopia ou laparotomia (TWICKLER *et al.*, 1999).

2.2.4 Manejo clínico e cirúrgico das tumorações ovarianas

O manejo clínico ou cirúrgico das tumorações ovarianas gera inúmeras controvérsias, uma vez que, estabelecido o diagnóstico, é importante definir qual a natureza biológica da lesão – se benigna ou maligna – e qual a conduta mais adequada, se expectante ou cirúrgica. E, no caso de ser invasiva, decidir se a remoção da lesão deva ser por laparoscopia ou laparotomia (CURTIN, 1994; TWICKLER *et al.*, 1999; DUFFY *et al.*, 2005).

A conduta expectante somente é tomada na suspeita de cistos funcionais de ovário decorrentes das alterações hormonais das gonadotrofinas e da síntese de esteróides (WESTHOFF & BERAL, 1984; CURTIN, 1994). Conforme o tamanho, esses tipos de lesão regridem após supressão hormonal com uso de anticoncepcionais em um período de três meses (SPANOS, 1972; WESTHOFF & BERAL, 1984; VESSEY *et al.*, 1987).

2.2.4.1 Laparoscopia *versus* laparotomia nas tumorações ovarianas malignas estágios iniciais pela FIGO (Ia, Ib e Ic)

Os benefícios cirúrgicos endoscópicos são incontestáveis do ponto de vista estético e de trauma cirúrgico, em virtude da diminuição da resposta inflamatória. O método, em síntese, mantém uma maior integridade física com menor supressão imunológica (TROIDL, 1999). A laparoscopia cirúrgica nas tumorações ovarianas benignas reduz seqüelas de aderências, proporciona uma rápida recuperação, menos tempo de internação hospitalar e retorno mais rápido às atividades habituais (CANIS *et al.*, 1994a; CANIS *et al.*, 1994b; CANIS *et al.*, 1997; CANIS *et al.*, 2002).

O principal problema da técnica são os casos com resultados falso-negativos para malignidade após a inspeção e/ou exame anatomopatológico de congelação, pois a ruptura capsular freqüentemente ocorre durante a cirurgia endoscópica, ao se procurar reduzir o diâmetro da tumoração e possibilitar a retirada da peça mediante pequenas incisões (CANIS *et al.*, 1994a; CANIS *et al.*, 1994b; CHILDERS, 1994; CHILDERS, NASSERI, SURWIT, 1996; CANIS *et al.*, 1997; CANIS *et al.*, 2002). No caso de a tumoração ser maligna, ocorre extravasamento de células neoplásicas

para o interior da cavidade abdominal e concomitante mudança dos estágios Ia ou Ib para Ic da FIGO (Anexo H), com possível modificação no prognóstico em razão do aumento de recidiva, com redução no índice de sobrevida em cinco anos (CUESTA *et al.*, 1994; Einhorn *et al.*, 2007).

Contudo, de acordo com alguns autores, a ruptura capsular que ocorre durante o procedimento cirúrgico no estágio I do câncer de ovário, pela classificação da FIGO, não modifica o prognóstico, haja vista que os fatores prognósticos principais considerados após estudos com análise multivariada são a radicalidade cirúrgica, o grau de diferenciação histológica, o estadiamento clínico, o volume residual da tumoração após a primeira intervenção e a presença de aderências densas (DEMBO *et al.*, 1990; SEVELDA *et al.*, 1990; SÖVALL, NILSSON, EINHORN, 1994; VOLZ, KÖSTER, SCHAEFF, 1997; VOLZ *et al.*, 1998; VOLZ *et al.*, 1999; DEMBO *et al.*, 2007; GHEZZI *et al.*, 2007).

Na tentativa de evitar diagnóstico errôneo de tumoração benigna quando maligna são propostos alguns cuidados transoperatórios durante a laparoscopia, tais como: lavado peritoneal; cuidadosa inspeção de toda a cavidade abdominal; na presença de lesão suspeita, realização de biópsias e conversão do procedimento para laparotomia; conforme a idade, características da ultra-sonografia e do Doppler colorido e o nível sérico do CA 125, realização, durante o transoperatório, de exame anatomopatológico de congelação (SELTZER, 1993). Esse exame permite um rápido diagnóstico, durante o transoperatório, do tipo histológico da tumoração e auxilia o cirurgião na tomada de decisão (LERMAN & PITCOCK, 1972). Entretanto, deve-se conhecer a limitação do método, por haver possibilidade de erro. Recente metanálise de acurácia diagnóstica do anatomopatológico de congelação mostrou que, para os tumores limítrofes, a probabilidade de diagnóstico correto era de 51% (MEDEIROS *et al.*, 2005).

Contudo, diversos autores preconizam a abordagem laparoscópica para o tratamento do câncer ovariano, com realização da histerectomia total e anexectomia bilateral associada à linfaadenectomia pélvica e paraórtica e omentectomia. Essas publicações demonstraram que a cirurgia laparoscópica é factível em casos de tumoração maligna ovariana e que as complicações são mínimas desde que executadas por médicos que dominem bem a técnica. Porém, na análise das evidências dessas publicações, verificam-se estudos com baixo grau de recomendação, por serem relatados de casos, séries de casos ou provindos de coorte retrospectiva

(KADAR, 1995; KIU-KWONG, FANG-PING, SHUENN-DYH, 1995; POMEL *et al.*, 1995; KADAR, 1997; VOLZ, KÖSTER, SCHAEFF, 1997; LÉCURU & TAURELLE, 1998; VOLZ *et al.*, 1998; DONATTINO *et al.*, 1999; DOTTINO, LEVINE, RIPLEY, 1999; VOLZ *et al.*, 1999; MANOLITSAS & FOLWER, 2001; LÉCURU *et al.*, 2004; TOZZI *et al.*, 2004; DOTTINO *et al.*, 2005; HUA, JIM, XU, 2005; TOZZI & SCHNEIDER, 2005; QUERLEU *et al.*, 2006a; QUERLEU *et al.*, 2006b; GHEZZI *et al.*, 2007) (Anexo C).

A eficácia do uso da laparoscopia cirúrgica no tratamento do câncer ovariano nos estágios iniciais não está bem estabelecida, havendo necessidade de ensaios clínicos randomizados para avaliarem-se o tempo cirúrgico, os custos, os índices de complicações e o tempo de sobrevida após o procedimento cirúrgico endoscópico (LÉCURU & TAURELLE, 1998; CHI & CURTIN, 1999; MANOLITSAS & FOLWER, 2001; LÉCURU *et al.*, 2004; HUA, JIM, XU, 2005).

Na cirurgia endoscópica, o gás dióxido de carbono (CO_2) é o mais utilizado para provocar o pneumoperitônio, não sendo, porém, aconselhável para a abordagem cirúrgica de neoplasias malignas, pois o CO_2 facilita a ocorrência de metástases nos portais dos trocartes, por proporcionar ativação de enzimas celulares que levam à mitose, promovendo aumento do fator de crescimento tumoral, com significância estatística ($p < 0,05$) se comparado com grupo controle de outros gases (GREENE, 1995; ELEFTHERIADIS & KOTZAMPASSI, 1996). Além do que, a pressão de CO_2 constante por mais de 60 minutos ocasiona isquemia intestinal com produção de radicais livres e predispõe a ocorrência de translocação das bactérias do lúmen do intestino para os linfonodos mesentéricos, hepáticos e esplênicos (ELEFTHERIADIS & KOTZAMPASSI, 1996). E estudos experimentais evidenciaram que, após 3 horas de pressão contínua de 14 mmHg de CO_2 na cavidade abdominal, há aumento de liberação de endotoxinas circulantes do interior do intestino para a circulação, havendo diminuição da fagocitose (VOLZ *et al.*, 1998; VOLZ & KÖSTER, 1999; VOLZ *et al.*, 1999). Além disso, CO_2 altera a integridade da membrana mesotelial por ocasionar acidose grave com dano mecânico e químico do mesotélio, o que predispõe à implantação de células neoplásicas na matriz celular lesada (VOLZ, KÖSTER, SCHAEFF, 1997; VOLZ *et al.*, 1998; VOLZ *et al.*, 1999).

2.2.5 Síntese da revisão de literatura

Em face de tantas particularidades, verifica-se que a decisão sobre o tipo de abordagem para a resolução cirúrgica dos tumores ovarianos malignos nos estágios iniciais (Ia, Ib e Ic) pela FIGO – se por laparoscopia ou se por laparotomia – é tomada em condições de incerteza devido a inúmeras publicações que se colocam a favor do procedimento laparoscópico e de outras que são terminantemente contra essa abordagem em caso de tumoração maligna.

Verifica-se também que a escolha da abordagem cirúrgica das tumorações ovarianas baseia-se em três testes diagnósticos reconhecidos como os principais para a tomada de decisão: a ultra-sonografia com Doppler colorido, dosagem sérica do CA 125 e exame anatomopatológico de congelação. Devido à importância dos testes diagnósticos na tomada de decisão, decidiu-se elaborar estudos de RS para avaliar a acurácia diagnóstica de cada um destes testes diagnósticos.

Da mesma forma, com existe controvérsia quanto à abordagem cirúrgica – se por laparoscopia ou por laparotomia nas tumorações malignas estágios iniciais (Ia, Ib e Ic) pela FIGO – decidiu-se efetuar um estudo de RS com enfoque de intervenção para auxiliar na tomada de decisão.

3 REFERÊNCIAS BIBLIOGRÁFICAS

3 REFERÊNCIAS BIBLIOGRÁFICAS

Aslam N, Banerjee S, Carr J V, Savvas M, Hooper R, Jurkovic D. Prospective evaluation of logistic regression models for the diagnosis of ovarian cancer. Obstet Gynecol 2000; (96): 75-80.

Barber H R. Ovarian cancer: diagnosis and management. Am J Obstet Gynecol 1984; (150): 910-916.

Benedet J L, Bender H, Jones H, III, Ngan H Y, Pecorelli S. FIGO staging classifications and clinical practice guidelines in the management of gynecologic cancers. FIGO Committee on Gynecologic Oncology. Int J Gynaecol Obstet 2000; (70): 209-262.

Berchuck A, Schildkraut J M, Marks J R, Futreal P A. Managing hereditary ovarian cancer risk. Cancer 1999; (86): 2517-2524.

Bossuyt P M, Reitsma J B, Bruns D E, Gatsonis C A, Glasziou P P, Irwig L M, Lijmer J G, Moher D, Rennie D, de Vet H C. Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD Initiative. Ann Intern Med 2003; (138): 40-44.

Bourne T, Campbell S, Steer C, Whitehead M I, Collins W P. Transvaginal colour flow imaging: a possible new screening technique for ovarian cancer. BMJ 1989; (299): 1367-1370.

Bourne T H. Transvaginal color Doppler in gynecology. Ultrasound Obstet Gynecol 1991; (1): 359-373.

Brioschi P A, Irion O, Bischof P, Bader M, Forni M, Krauer F. Serum CA 125 in epithelial ovarian cancer. A longitudinal study. Br J Obstet Gynaecol 1987; (94): 196-201.

Brown D L, Doubilet P M, Miller F H, Frates M C, Laing F C, DiSalvo D N, Benson C B, Lerner M H. Benign and malignant ovarian masses: selection of the most discriminating gray-scale and Doppler sonographic features. Radiology 1998; (208): 103-110.

Canis M, Mage G, Pouly J L, Walttiez A, Manhes H, Bruhat M A. Laparoscopic diagnosis of adnexal cystic masses:a 12 year experience with long term follow up. Obstet Gynecol 1994a; (83): 707-712.

Canis M, Watiez A, Mage G, Pouly J L, Raiga J, Goff B M B, Manhes H, Bruhat M A. Laparoscopic management of adnexal masses. Baillieres Best Pract Res Clin Obstet Gynaecol 1994b; (8): 723-734.

Canis M, Pouly J L, Walttiez A, Mage G, Manhes H, Bruhat R S. Laparoscopic management of adnexal masses suspicious at ultrasound. Obstet Gynecol 1997; (89): 679-683.

Canis M, Rabischong B, Houlle C B, Jarson K, Safi A W A, *et al.* Laparoscopic management of adnexal masses: a gold standart? Curr Opin Obstet Gynecol 2002; (14): 423-428.

Chi D S, Curtin J P. Gynecologic cancer and laparoscopy. Obstet Gynecol Clin North Am 1999; (26): 201-215.

Childers J M. Operative laparoscopy in gynaecological oncology. Baillieres Best Pract Res Clin Obstet Gynaecol 1994; (8): 831-847.

Childers J M, Aqua K A, Surwit E A, Hallum A V, Hatch K. Abdominal-Wall Tumor implantation after laparoscopy for malignant conditions. Obstet Gynecol 1994; (84): 765-769.

Childers J M, Nasseri A, Surwit E A. Laparoscopic management of suspious adnexal masses. Am J Obstet Gynecol 1996; (175): 1451-1459.

Cook D J, Mulrow C D, Haynes R B. Systematic reviews: synthesis of best evidence for clinical decisions. Ann Intern Med 1997; (126): 376-380.

Crayford T J, Campbell S, Bourne T H, Rawson H J, Collins W P. Benign ovarian cysts and ovarian cancer: a cohort study with implications for screening. Lancet 2000; (355): 1060-1063.

Cuesta S R, Goff B A, Fuller A F, Nikrui N, Eichhorn J H, Rice L W. Prognostic importance of intraoperative rupture of malignant ovarian epithelial neoplasm. Obstet Gynecol 1994; (84): 1-7.

Curtin J P. Management of the adnexal mass. Gynecol Oncol 1994; (55): S42-S46.

Deeks J J. Systematic reviews of evaluations of diagnostic and screening tests. In: Systematic reviews in health care - Meta-analysis in context. (Eds.Egger M, Smit DG, Altman DG). London: BMJ Publishing Group, 2001a; 248-282.

Deeks J J. Systematic reviews in health care: Systematic reviews of evaluations of diagnostic and screening tests. BMJ 2001b; (323): 157-162.

Deeks J J, Altman D G, Bradburn M J. Statistical methods for examining heterogeneity and combining results from several studies in meta-analysis. In: Systematic Reviews in Health Care - Meta-analysis in context. (Eds. Egger M, Smit D G, Altman D G). London: BMJ Publishing Group, 2001; 2nd: 285-312.

Dembo A J, Davy M, Stenwig A E, Berle E J, Bush R S, Kjorstad K. Prognostic factors in patients with stage I epithelial ovarian cancer. Obstet Gynecol 1990; (75): 263-273.

Dembo A J, Davy M, Stenwing A, Berle E J, Bush R S, Kjorstad K. Prognostic factors in patients with stage I epithelial ovarian cancer. Obstet Gynecol 2007; (75): 263-73D.

DePriest P D, Varner E, Powell J, Fried A, Puls L, Higgins R, Shenson D, Kryscio R, Hunter J E, Andrews S J,. The efficacy of a sonographic morphology index in identifying ovarian cancer: a multi-institutional investigation. Gynecol Oncol 1994; (55): 174-178.

Donattino P R, Tobias D H, Beddoe A M, Golden A L, Cohen C J. Laparoscopic lymphadenectomy for gynecologic malignancies. Gynecol Oncol 1999; (73): 383-388.

Dottino P, Levine D A, Ripley D C C. Laparoscopic Managment of adnexal Masses in premenopausal and Postmenopausal women. Obstet Gynecol 1999; (93): 223-227.

Dottino P, Tobias D H, Beddoe A M, Golden A L, Cohen C J. Laparoscopic managment of bordeline ovarian tumors: results of a French multicenter study. Ann Oncol 2005; (16): 403-410.

Duffy M J, Bonfrer J M, Kulpas J, Rustin G J S, Soletormos G, Torre G C, Tuxen M K, Zwirner M. CA 125 in ovarian cancer: European Group on tumor markes guidelines for cliical use. Int J Gynecol Cancer 2005; (15): 679-691.

Edmondson R J M J. The epidemiology of ovarian cancer. Int J Gynecol Cancer 2001; (11): 423-429.

Egger M, Smith G D. Misleading meta-analysis. BMJ 1995; (310): 752-754.

Egger M, Smith G D. Meta-Analysis. Potentials and promise. BMJ 1997; (315): 1371-1374.

Egger M, Smith G D, Phillips A N. Meta-analysis: principles and procedures. BMJ 1997; (315): 1533-1537.

Egger M, Schneider M, Davey S G. Spurious precision? Meta-analysis of observational studies. BMJ 1998; (316): 140-144.

Egger M, Smith G D, Schneider M. Sytematic reviews of observational studies. In: Systematic reviews in Health Care - Meta-analysis in context. (Eds. Egger M, Smith G D, Schneider M). London: BMJ Publishing, 2001; 2nd: 211-227.

Egger M, Smith G D, Sterne J A. Uses and abuses of meta-analysis. Clin Med 2001; (1): 478-484.

Egger M, Ebrahim S, Smith G D. Where now for meta-analysis? Int J Epidemiol 2002; (31): 1-5.

Einhorn N, Bast R C, Jr., Knapp R C, Tjernberg B, Zurawski V R, Jr. Preoperative evaluation of serum CA 125 levels in patients with primary epithelial ovarian cancer. Obstet Gynecol 1986; (67): 414-416.

Einhorn N, Knapp R C, Bast R C, Zurawski V R, Jr. CA 125 assay used in conjunction with CA 15-3 and TAG-72 assays for discrimination between malignant and non-malignant diseases of the ovary. Acta Oncol 1989; (28): 655-657.

Einhorn N, Bast R, Knapp R, Nilsson B, Zurawski V, Sjövall K. Long term follo-up of the Stockholm screening study on ovarian cancer. Gynecol Oncol 2007; (79): 466-470.

Ekerhovd E, Wienerroith H, Staudach A, Granberg S. Preoperative assessment of unilocular adnexal cysts by transvaginal ultrasonography: a comparison between ultrasonographic morphologic imaging and histopathologic diagnosis. Am J Obstet Gynecol 2001; (184): 48-54.

Eleftheriadis E, Kotzampassi K P M H M S K. Gut Ischemia, oxidative stress, and bacterial translocation in elevated abdominal pressure in rats. World J Surg 1996; (20): 11-16.

Eltabbakh G H, Natarajan N, Piver M S, Mettlin C J. Epidemiologic differences between women with borderline ovarian tumors and women with epithelial ovarian cancer. Gynecol Oncol 1999; (74): 103-107.

Ghezzi F, Cromi A U S, Bergamini V, Tomera S, Franchi M, Bolis P. Laparoscopy versus laparotomy for the surgical management of apparent early stage ovarian cancer. Gynecol Oncol 2007; (105): 409-413.

Go A S. Refinando a probabilidade:introdução à solicitação de exames complementares. In: Medicina baseada em evidências. (Eds. Friedland DJ, Go A S, Davoren J B, Shlipak M G, Bent S W, Subak L L, Mendelson T). Rio de Janeiro: Editora Guanabara Koogan, 2001; 1 ed: 12-31.

Granberg S. Ultrasound in the diagnosis and treatment of ovarian tumors. Acta Obstet Gynecol Scand 1991; (70): 385-386.

Greene F L. Principles of cancer biology in relayion to minimal acess surgical techniques. Semin Laparosc Surg 1995; (2): 155-157.

Greenhalgh T. How to read a paper. Papers that report diagnostic or screening tests. BMJ 1997a; (315): 540-543.

Greenhalgh T. Papers that summarise other papers (systematic reviews and meta-analyses). BMJ 1997b; (315): 672-675.

Hua K Q, Jim F M, Xu H Z Z L J F Y. [Evaluation of laparoscopic surgery in the early stage malignant tumor of ovary with lower risk]. Zhonghua Yi Xue Za Zhi 2005; (85): 169-172.

INCA. Incidênia de câncer no Brasil em mulheres. Internet. 2007. Ref Type: Electronic Citation

Irwig L, Tosteson A N, Gatsonis C, Lau J, Colditz G, Chalmers T C, Mosteller F. Guidelines for meta-analyses evaluating diagnostic tests. Ann Intern Med 1994; (120): 667-676.

Irwig L, Macaskill P, Glasziou P, Fahey M. Meta-analytic methods for diagnostic test accuracy. J Clin Epidemiol 1995; (48): 119-130.

Jacobs I, Bast R C, Jr. The CA 125 tumour-associated antigen: a review of the literature. Hum Reprod 1989; (4): 1-12.

Juni P, Witschi A, Bloch R, Egger M. The hazards of scoring the quality of clinical trials for meta-analysis. JAMA 1999; (282): 1054-1060.

Kabawat S E, Bast R C, Welch W R, Knapp R C, Colvin R B. Immunopathologic characterization of a monoclonal antibody that recognizes common surface antigens of human ovarian tumors of serous, endometrioid, and clear cell types. Am J Clin Pathol 1983; (79): 98-104.

Kadar N. Laparoscopic surgery for gynaecological malignancies in women age 65 years or more. Gynaecological Endoscopy 1995; (4): 173-181.

Kadar N. Laparoscopy managment of gynecological malignancies. Curr Opin Obstet Gynecol 1997; (9): 247-255.

Kiu-Kwong C, Fang-Ping C, Shuenn-Dyh. Laparoscopic surgical procedures for early ovarian cancer. Acta Obstet Gynecol Scand 1995; (74): 391-401.

Köhler C, Klemm P, Schau A, Possover M, Krause N, Tozzi R, Schneider A. Introduction of transperitoneal lymphaadenectomy in a gynecologic oncology center: analysis of 650 laparoscopic pelvic and/or paraaortic transperitoneal lymphadenectomies. Gynecol Oncol 2004; (95): 52-61.

Koonings P P, Campbell K, Mishell D R, Jr., Grimes D A. Relative frequency of primary ovarian neoplasms: a 10-year review. Obstet Gynecol 1989; (74): 921-926.

Kreder H J. Evidence-based surgical practice: what is it and do we need it? World J Surg 1999; (23): 1232-1235.

Kurjak A, Zalud I, Jurkovic D, Alfirevic Z, Miljan M. Transvaginal color Doppler for the assessment of pelvic circulation. Acta Obstet Gynecol Scand 1989; (68): 131-135.

Kurjak A, Predanic M. New scoring system for prediction of ovarian malignancy based on transvaginal color Doppler sonography. J Ultrasound Med 1992; (11): 631-638.

Kurjak A, Schulman H, Sosic A, Zalud I, Shalan H. Transvaginal ultrasound, color flow, and Doppler waveform of the postmenopausal adnexal mass. Obstet Gynecol 1992; (80): 917-921.

Kurjak A, Shalan H, Kupesic S, Predanic M, Zalud I, Breyer B, Jukic S. Transvaginal color Doppler sonography in the assessment of pelvic tumor vascularity. Ultrasound Obstet Gynecol 1993a; (3): 137-154.

Kurjak A, Shalan H, Matijevic R, Predanic M, Kupesic-Urek S. Stage I ovarian cancer by transvaginal color Doppler sonography: a report of 18 cases. Ultrasound Obstet Gynecol 1993b; (3): 195-198.

Kurjak A, Kupesic S. Transvaginal color Doppler and pelvic tumor vascularity: lessons learned and future challenges. Ultrasound Obstet Gynecol 1995; (6): 145-159.

Kurtz A B, Tsimikas J V, Tempany C M, Hamper U M, Arger P H, Bree R L, Wechsler R J, Francis I R, Kuhlman J E, Siegelman E S, Mitchell D G, Silverman S G, Brown D L, Sheth S, Coleman B G, Ellis J H, Kurman R J, Caudry D J, McNeil B J. Diagnosis and staging of ovarian cancer: comparative values of Doppler and conventional US, CT, and MR imaging correlated with surgery and histopathologic analysis--report of the Radiology Diagnostic Oncology Group. Radiology 1999; (212): 19-27.

Lécuru F, Taurelle F. Transperitoneal laparoscopic pelvic lypmphadenectomy for gynecologic malignancies. Sur Endosc 1998; (12): 97-100.

Lécuru F, Desfeux P, Camatte S, Bissery A, Robin F, Blanc B, Querleu D. Stage I ovarian cancer:comparison of laparoscopy and laparotomy on stanging and survival. Eur J Gynaec Oncol 2004;571-576.

Lefebvre C, Clarke M J. Identifying randomised trials. In: Systematic reviews in health care - Meta-analysis in context. (Eds. Egger M, Smith G D, Altman D G). London: BMJ Pushishing Group, 2001; 69-86.

Lerman R I, Pitcock J A. Frozen section experience in 3,249 specimens. Surg Gynecol Obstet 1972; (135): 930-932.

Lijmer J G, Mol B W, Heisterkamp S, Bonsel G J, Prins M H, van der Meulen J H, Bossuyt P M. Empirical evidence of design-related bias in studies of diagnostic tests. JAMA 1999; (282): 1061-1066.

Littenberg B, Moses L E. Estimating diagnostic accuracy from multiple conflicting reports: a new meta-analytic method. Med Decis Making 1993; (13): 313-321.

Lorenz W, Troidl H, Solomkin J S, Nies C, Sitter H, Koller M, Krack W, Roizen M F. Second step: testing-outcome measurements. World J Surg 1999; (23): 768-780.

Maggino T, Sopracordevole F, Matarese M, Di P C, Tambuscio G. CA-125 serum level in the diagnosis of pelvic masses: comparison with other methods. Eur J Gynaecol Oncol 1987; (8): 590-595.

Maggino T, Gadducci A, D'Addario V, Pecorelli S, Lissoni A, Stella M, Romagnolo C, Federghini M, Zucca S, Trio D,. Prospective multicenter study on CA 125 in postmenopausal pelvic masses. Gynecol Oncol 1994; (54): 117-123.

Manolitsas T P, Folwer J M. Role of laparoscopy in the management of the adnexal mass and staging of gynecologic cancers. Clin Obstet Gynecol 2001; (44): 495-521.

Medeiros L R, Rosa D D, Edelweiss M I, Stein A T, Bozzetti M C, Zelmanowicz A, Pohlmann P R, Meurer L, Carballo M T. Accuracy of frozen-section analysis in the diagnosis of ovarian tumors: a systematic quantitative review. Int J Gynecol Cancer 2005; (15): 192-202.

Midgette A S, Stukel T A, Littenberg B. A meta-analytic method for summarizing diagnostic test performances: receiver-operating-characteristic-summary point estimates. Med Decis Making 1993; (13): 253-257.

Moher D, Cook D J, Eastwood S, Olkin I, Rennie D, Stroup D F. Improving the Quality of Reports of Meta-Analyses of Randomised Controlled Trials: The QUOROM Statement. Onkologie 2000; (23): 597-602.

Moher D, Schulz K F, Altman D. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. JAMA 2001; (285): 1987-1991.

Müller J M, Schwenk W, Jacobi C A, Böhm B. Endoscopy surgery: fit for malignancy? World J Surg 1999; (23): 808-815.

Mulrow C D. Rationale for systematic reviews. BMJ 1994; (309): 597-599.

Mulrow C D, Cook D J, Davidoff F. Systematic reviews: critical links in the great chain of evidence. Ann Intern Med 1997; (126): 389-391.

Nelson L, Ekbom A, Gerdin E. Ovarian cancer in young women in Sweden, 1989-1991. Gynecol Oncol 1999; (74): 472-476.

Osmers R G, Osmers M, von M B, Wagner B, Kuhn W. Evaluation of ovarian tumors in postmenopausal women by transvaginal sonography. Eur J Obstet Gynecol Reprod Biol 1998; (77): 81-88.

Oxman A D, Cook D J, Guyatt G H. Users' guides to the medical literature. VI. How to use an overview. Evidence-Based Medicine Working Group. JAMA 1994; (272): 1367-1371.

Paspalicchio J C, Fristachi C E., Castanho P R O L, Kue C M, Piatto S, Bacarat F F. Epidemiologia do câncer de ovário no Brasil. Revista da Sociedade Brasileira de Cancerologia 11. 2000.
Ref Type: Electronic Citation

Phillips B, Ball C, Sackett D *et al.* Oxford Centre for evidence-based Medicine Level of evidence Grades of Recomendations (may 2001). Oxford Centre for evidence-based Medicine. 2007.
Ref Type: Electronic Citation

Pomel C, Provencher D D J, Gauthier P, Le Bouedec G, Drouuin P, Audet-Lapointe P, *et al.* Laparoscopic staging of early ovarian cancer. Gynecol Oncol 1995; (58): 301-306.

Querleu D, Leblanc E, Cartron G N F F G, Martel P. Audit of preoperative and early complications lymph node dissection in 1000 gynecologic cancer patients. Am J Obstet Gynecol 2006a; (195): 1287-1292.

Querleu D, Leblanc E, Ferron G, Narducci F. Laparoscopy surgery in gynaecological oncology. EJSO 2006b; (32): 853-858.

Reeves B. Health-technology assessment in surgery. Lancet 1999; (353 Suppl 1): SI3-SI5.

Richardson W S, Wilson M C, Guyatt G H, Cook D J, Nishikawa J. Users' guides to the medical literature: XV. How to use an article about disease probability for differential diagnosis. Evidence-Based Medicine Working Group. JAMA 1999; (281): 1214-1219.

Sackett D L. Clinical epidemiology. Am J Epidemiol 1969; (89): 125-128.

Sackett D L. Clinical diagnosis and the clinical laboratory. Clin Invest Med 1978; (1): 37-43.

Sackett D L, Rosenberg W M. On the need for evidence-based medicine. J Public Health Med 1995; (17): 330-334.

Sackett D L. Clinical epidemiology. what, who, and whither. J Clin Epidemiol 2002; (55): 1161-1166.

Sassone A M, Timor-Tritsch I E, Artner A, Westhoff C, Warren W B. Transvaginal sonographic characterization of ovarian disease: evaluation of a new scoring system to predict ovarian malignancy. Obstet Gynecol 1991; (78): 70-76.

Sauerland S, Lefering R, Neugebauer E A. The pros and cons of evidence-based surgery. Langenbecks Arch Surg 1999; (384): 423-431.

Scully R E. Early de novo ovarian cancer and cancer developing in benign ovarial lesions. Int J Gynecol Cancer 1995; (49): S9-S15.

Scully R E. Histological typing of ovarian tumours - World Health Organization International histological classification of tumors. Springer-Verlag, Berlin 1999.

Scully R E. Influence of origin of ovarian cancer on efficacy of screening. Lancet 2000; (355): 1028-1029.

Seltzer V. Laparoscopic surgery for ovarian lesions: potential pitfalls. Clin Obstet Gynecol 1993; (36): 402-412.

Sevelda P, Vavra N, Schemper M, Salzer H. Prognostic factors for survival in stage I epithelial ovarin carcinoma. Cancer 1990; (65): 2349-2352.

Smidt V J, Singh D M, Hurteau J A, Hurd W W. Effect of carbon dioxide on human ovarian carcinoma cell growth. Am J Obstet Gynecol 2001; (185): 1314-1317.

Smith-Bindman R, Kerlikowske K, Feldstein V A, Subak L, Scheidler J, Segal M, Brand R, Grady D. Endovaginal ultrasound to exclude endometrial cancer and other endometrial abnormalities. JAMA 1998; (280): 1510-1517.

Soper J T, Hunter V J, Daly L, Tanner M, Creasman W T, Bast R C, Jr. Preoperative serum tumor-associated antigen levels in women with pelvic masses. Obstet Gynecol 1990; (75): 249-254.

Sövall K, Nilsson B, Einhorn N. Different types of rupture of the tumor capsule and the impact on survival in early ovarian carcinoma. Int J Gynecol Cancer 1994; (4): 333-336.

Spanos W J. Preoperative hormonal therapy of cystic adnexal masses. Trans Pac Coast Obstet Gynecol Soc 1972; (40): 111-116.

Stroup D F, Berlin J A, Morton S C, Olkin I, Williamson G D, Rennie D, Moher D, Becker B J, Sipe T A, Thacker S B. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. JAMA 2000; (283): 2008-2012.

Sutton A J, Abrams K R, Jones D R, Sheldon T A, Song F. Meta-analysis of different types of data. In: Methods for Meta-Analysis in Medical Research. (Eds. Sutton A J, Abrams K R, Jones D R, Sheldon T A, Song F). Chichester: John Wiley & Sons, 2000; 1: 205-228.

Tosteson A N, Begg C B. A general regression methodology for ROC curve estimation. Med Decis Making 1988; (8): 204-215.

Tozzi R, Köhler C, Ferrara A, Schneider A. Laparoscopic treatment of early ovarian cancer: surgical and survival outcomes. Gynecol Oncol 2004; (93): 199-203.

Tozzi R, Schneider A. Laparoscopy treatment of early ovarian cancer. Curr Opin Obstet Gynecol 2005.

Troidl H. Endoscopic surgery: innovation versus evaluation-introduction. World J Surg 1999; (23): 743-744.

Twickler D M, Forte T B, Santos-Ramos R, McIntire D, Harris P, Scott D. The Ovarian Tumor Index predicts risk for malignancy. Cancer 1999; (86): 2280-2290.

Vamvakas E C. Meta-analyses of studies of the diagnostic accuracy of laboratory tests: a review of the concepts and methods. Arch Pathol Lab Med 1998; (122): 675-686.

van N J, Jr., DePriest P D, Reedy M B, Gallion H H, Ueland F R, Pavlik E J, Kryscio R J. The efficacy of transvaginal sonographic screening in asymptomatic women at risk for ovarian cancer. Gynecol Oncol 2000; (77): 350-356.

Vergote I B, Bormer O P, Abeler V M. Evaluation of serum CA 125 levels in the monitoring of ovarian cancer. Am J Obstet Gynecol 1987; (157): 88-92.

Vergote I, Trimbos JB. Treatment of patients with early epithelial ovarian cancer. Curr Opin Oncol 2003; (15): 452-455.

Vergote I B, Amant F. Early ovarian cancer--time for a rethink on stage? Gynecol Oncol 2004; (94): 607-608.

Vessey M, Metcalfe A, Wells C, McPherson K, Westhoff C, Yeates D. Ovarian neoplasms, functional ovarian cysts, and oral contraceptives. Br Med J (Clin Res Ed) 1987; (294): 1518-1520.

Volz J, Köster S, Schaeff B. Laparoscopy managment of gynaecological malignancies: time to hesite. Gynaecological Endoscopy 1997; (6): 145-146.

Volz J, Koster S, Schaeff B, Paolucci V. Laparoscopic surgery: the effects of insufflation gas on tumor-induced lethality in nude mice. Am J Obstet Gynecol 1998; (178): 793-795.

Volz J, Köster S. Laparoscopy:to inflate or lift. Cancer 1999; (86): 749-750.

Volz J, Koster S, Spacek Z, Paweletz N. The influence of pneumoperitoneum used in laparoscopic surgery on an intraabdominal tumor growth. Cancer 1999; (86): 770-774.

von E E, Altman D G, Egger M, Pocock S J, Gotzsche P C, Vandenbroucke J P. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet 2007; (370): 1453-1457.

Walter S D, Jadad A R. Meta-analysis of screening data: a survey of the literature. Stat Med 1999; (18): 3409-3424.

Wang P H, Yuan C C, Lin G, Ng H T, Chao H T. Risk factors contributing to early occurrence of port site metastases of laparoscopic surgery for malignancy. Gynecol Oncol 1999; (72): 38-44.

Weiner Z, Thaler I, Beck D, Rottem S, Deutsch M, Brandes J M. Differentiating malignant from benign ovarian tumors with transvaginal color flow imaging. Obstet Gynecol 1992; (79): 159-162.

Wells B, Shea B, O'coneell, Peterson J, Welch V, Losos M, Tugwell P. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analysis. Internet. 2007.
Ref Type: Electronic Citation

Westhoff C L, Beral V. Patterns of ovarian cyst hospital discharge rates in England and Wales, 1962-79. Br Med J (Clin Res Ed) 1984; (289): 1348-1349.

Wu C C, Lee C N, Chen T M, Shyu M K, Hsieh C Y, Chen H Y, Hsieh F J. Incremental angiogenesis assessed by color Doppler ultrasound in the tumorigenesis of ovarian neoplasms. Cancer 1994; (73): 1251-1256.

Yancik R. Ovarian cancer. Age contrasts in incidence, histology, disease stage at diagnosis, and mortality. Cancer 1993; (71): 517-523.

Zamora J, Abraira V, Muriel A, Khan K S, Coomarasamy A. Meta-DiSc: a software for meta-analysis of test accuracy data. In: BMC Medical Research Methodology, 2006; 31.

4 OBJETIVOS

4 OBJETIVOS

4.1 Objetivo geral

Realizar estudos de revisão sistemática com enfoque diagnóstico e terapêutico das tumorações ovarianas com pressupostos de malignidade em seus estágios iniciais (Ia, Ib e Ic) pela FIGO, tendo como finalidade principal que seus resultados auxiliem na tomada de decisão ou sugeriram a realização de novos estudos embasados em melhores níveis de evidência.

4.2 Objetivos específicos

4.2.1 Realizar revisão sistemática com enfoque de intervenção cirúrgica, das tumorações ovarianas quanto à abordagem por laparoscopia ou por laparotomia.

4.2.2 Realizar revisão sistemática com enfoque diagnóstico avaliando a acurácia diagnóstica do exame anatomopatológico de congelação nas tumorações ovarianas.

4.2.3 Realizar revisão sistemática com enfoque diagnóstico avaliando a acurácia diagnóstica da ultra-sonografia com Doppler colorido nas tumorações ovarianas.

4.2.4 Realizar revisão sistemática com enfoque diagnóstico avaliando a acurácia diagnóstica do marcador tumoral CA 125 nas tumorações ovarianas.

5 ARTIGOS

5.1 Artigo 1

LAPAROSCOPY VERSUS LAPAROTOMY FOR FIGO STAGE I OVARIAN CANCER

Submetida revisão sistemática completa, data da publicação 16 abril de 2008

***The Cochrane Library,* Issue 4, 2007. Oxford: Update Software. (Protocol)**

Issue 2, 2008 (Complete Systematic Review)
Submission to MSDG deadline 30 November 2007
Copy-edit support deadline 30 Jan 2008
Module submission deadline 21 Feb 2008
Publication date 16 Apr 2008

Laparoscopy Versus Laparotomy for Figo Stage I Ovarian Cancer

Author: Lidia Rosi Medeiros, MD, MSc[1]
Address: José de Alencar 1244 apt 1009, Porto Alegre, RS, Brasil, CEP 90880-480
lidia.rosi@terra.com.br
Co-authors:

Mary Clarisse Bozzetti, MD, PhD[1,2,4]
Daniela Dornelles Rosa, MD[2]
Maria Ines da Rosa[1,9]
Maria Isabel Edelweiss, MD, PhD[2,5]
Airton Tetelbon Stein, MD, PhD[2,3,6]
Alice Zelmanowicz, MD, PhD[7]
Anaelena de Bragança, PhD [8]
Roselaine Ruviaro Zanini, PhD[8]

Affiliations of all authors:

[1] Postgraduate Program in Epidemiology at Federal University of Rio Grande do Sul, Porto Alegre Brazil
[2] Postgraduate Program in Medicine: Medical Sciences at Federal University of Rio Grande do Sul, Porto Alegre, Brazil
[3] Public Health Postgraduate Course at Universidade Luterana do Brasil
[4] Department of Social Medicine, Faculty of Medicine at Federal University of Rio Grande do Sul, Porto Alegre Brazil
[5] Pathology Unit of Hospital de Clinicas de Porto Alegre, Porto Alegre, Brazil
[6] Postgraduate Program in Medical Science at Porto Alegre Federal Faculty of Medical Science, Porto Alegre, Brazil
[7] Cancer Prevention Center, Complexo Hospitalar Santa Casa, Porto Alegre, Brazil
[8] Department of Statistics, Federal University of Santa Maria, Santa Maria, Brazil
[9] Medical School at University of Extremo Sul Catarinense, Criciúma, Brazil

Sinopse

Discussões e controvérsias têm ocorrido entre cirurgiões endoscópicos e oncologistas quanto ao tratamento laparoscópico dos tumores ovarianos malignos em seus estágios iniciais. Esta revisão sistemática não encontrou evidências que suportem o uso da laparoscopia para o manejo terapêutico inicial desses tumores como rotina na prática clínica.

RESUMO

Fundamentação teórica

Nos últimos 10 anos, a laparoscopia tem se tornado um procedimento comum no manejo cirúrgico das tumorações ovarianas malignas em seus estágios iniciais. Entretanto, existem incertezas quanto ao valor dessa intervenção nestes tipos de casos. Esta revisão sistemática foi realizada justamente para avaliar as evidências dos benefícios e o risco do procedimento endoscópico em pacientes com câncer de ovário em seus estágios iniciais comparando-o com a laparotomia.

Objectivo

O objetivo desta revisão sistemática foi avaliar o impacto da laparoscopia no manejo cirúrgico em pacientes com estágio I de câncer de ovário (Ia, Ib e Ic) pela FIGO quando comparado com a laparotomia.

Estratégias de busca

A pesquisa foi realizada de maneira sistemática, entre os anos de 1990 a 2007, com o auxílio do Grupo de Câncer Ginecológico da Cochrane, nos seguintes bancos de dados: Cochrane Central de Registro de Estudos Controlados (CENTRAL), MEDLINE (janeiro de 1990 a novembro de 2007), EMBASE (janeiro de 1990 a novembro de 2007), LILACS (janeiro de 1990 a Novembro 2007), BIOLOGICAL ABSTRACTS (janeiro de 1990 a novembro 2007) and CANCERLIT (janeiro de 1990 a novembro 2007). Foram procuradas também publicações através de busca manual, em periódicos médicos relevantes até novembro de 2007 e listas de referência de artigos bem como resumos de conferências médicas.

Critérios de seleção

Foram incluídos estudos relativos a pacientes, com comprovação histológica de câncer ovário em seus estágios iniciais de acordo com a Federação Interenacional de Ginecologia e Obstetrícia (FIGO). Estudos comparando laparoscopia e laparotomia para manejo clínico dessas começaram a ser realizados a partir de 1990, presumindo-se que devem existir muito poucos estudos randomizados com esse tipo de enfoque. Por tal razão, foram também considerados estudos de coorte e de casos e controles comparando as duas técnicas, tendo sido excluídos somente estudos de séries de casos e retrospectivos.

Extração de dados e análise

Os dados foram extraídos de forma independente por cinco revisores (LRM, DDR, MIR, MCB e MIE) que aferiam a qualidade do estudo e a qualidade dos dados extraídos. A extração dos dados inclui avaliar as características clínicas de cada estudo, como tipo de participantes, intervervenção e desfechos. A qualidade dos estudos não randomizados foi aquilatado por dois tipos de intrumentos de avaliação metodológica: o STROBE (*Strengthening the Reporting of Observational Studies in Epidemiology*) e o NOS (*Newcastle-Ottawa*) para estudos com enfoque observacional.

Resultados principais

Não foi encontrado nenhum estudo randomizado. Somente, três estudos observacionais, mas com pobre qualidade metodológica. Portanto, tornou-se impossível realizar a metanálise devido a diferença entre os estudos.

Conclusão dos revisores

Esta revisão sistemática não encontrou evidências que apóiem o uso da laparoscopia para o manejo inicial do câncer ovariano em seus estágios iniciais como rotina na prática clínica. Verifica-se a necessidade de estudos com boa qualidade metodológica comparando laparoscopia e laparotomia para o manejo cirúrgico inicial desses tipos de câncer. Embora haja necessidade de estudos randomizados controlados, estudos observacionais com alta qualidade metodológica podem proporcionar evidências seguras.

Synopsis

Controversial discussion has arisen among endoscopists and oncologists about the laparoscopic management of early stage ovarian tumours. This systematic review found no evidence to support the use of laparoscopic for the management of early stage ovarian cancer as routine in clinical practice.

ABSTRACT

Background

Over the last ten years laparoscopy has become an increasingly common approach for the surgical removal of early stage ovarian tumours. There remains uncertainty about the value of this intervention. This review has been undertaken to assess the available evidence for the benefits and harms of laparoscopic surgery for the management of early stage ovarian cancer compared to laparotomy.

Objectives

The objective of this review was to evaluate the impact of laparoscopy in the surgical treatment of FIGO stage I ovarian cancer (stages Ia, Ib and Ic) when compared with laparotomy.

Search strategy

Trials were identified by searching the Cochrane Gynaecological Cancer Group trials register, Cochrane Central Register of Controlled Trials (CENTRAL) (The Cochrane Library Issue 2, 2007), MEDLINE (January 1990 to date), EMBASE (1990 to November 2007), LILACS (1990 to November 2007), BIOLOGICAL ABSTRACTS (1990 to November 2007) and Cancerlit (1990 to November 2007). We also searched our own publication archives, based on prospective handsearching of relevant journals from November 2007. Reference lists of identified studies, gynaecological cancer handbooks and conference abstract were also scanned.

Selection criteria

Studies regarding patients with histologically proven stage I ovarian cancer according to the International Federation of Gynaecology and Obstetrics (FIGO) was included in this review.

Studies comparing laparoscopic surgery with laparotomy for early stage ovarian cancer were only available from 1990.

It was anticipated that a very small number of randomised controlled trials (RCTs) were conducted studying the management of early stage ovarian cancer. Therefore, non-randomised comparative studies, cohort studies and case-controls studies, but not studies with historical controls, were also considered for this review.

Data collection & analysis

Data extractions were performed independently by five reviewers (LRM, DDR, MIR, MCB and MIE) who assessed study quality and quality of extracted data. Extracted data included trial characteristics, characteristics of the study participants, interventions and outcomes. The quality of non randomised studies was assessed using appropriate quality evaluations tools from the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) and from the Newcastle-Ottawa tool for observational studies.

Main results

We did not find any randomised controlled trials. Three observational studies were found, but their quality was poor. It was not possible to performed a meta-analysis due to the difference among the studies.

Reviewers' conclusions

This review has found no evidence to support the use of laparoscopy for the management of early stage ovarian cancer as routine clinical practice. Further studies with good quality are needed comparing laparoscopy surgery with laparotomy for the management of early stage ovarian cancer. Although there is a need of RCTs, observational studies of higher quality could provide us with more reliable evidence.

Background

Malignant ovarian neoplasms are responsible for four per cent of all cancer affecting women and are the second most common cause of death from gynaecological cancer and the fourth most common cause of death from all types of cancer affecting women (Yancik, 1993). Diagnosis of early stage ovarian cancer (limited to the ovaries) is rare and is mainly made by accidental discovery at the time

of routine ultrasonography or during laparoscopy. The incidence of managing an unexpected ovarian cancer by laparoscopy is 6.5 in 1000 women with an adnexal mass (Wenzl *et al.*, 1996).

Most cancers of the ovary are of the epithelial types. The most common histologic subtype is serous, which comprise 40% to 70% of all types; endometrioid tumours are the second most common, (20% to 25% of all cases). Mucinous epithelial tumours are rarer, comprising 5% to 20% of all cases (Kosary, 1994). Borderline ovarian tumours constitute approximately 5.9% of primary epithelial ovarian cancers (Medeiros *et al.*, 2005).

The diagnosis of borderline ovarian tumours is more difficult due to variations in the histopathologic criteria used among different countries for the differential diagnosis between borderline and malignant lesions (Burger *et al.*, 2000, Medeiros *et al.*, 2005). Stromal and germ cell tumours make up 1.1% to 1.7% of all cases of malignant ovarian tumours (Medeiros *et al.*, 2005).

The prognosis of all ovarian tumours is independently affected by the following factors: stage of cancer at diagnosis, histological subtype, tumoral grading and the volume of residual disease after surgery (Benedet *et al.*, 2000, Medeiros *et al.*, 2005). Current standard treatment for patients with early stage ovarian cancer is a laparotomy with a longitudinal median incision to allow the required surgical staging (Benedet *et al.*, 2000, Hand *et al.*, 1993, Kosary, 1994). The primary tumour, if limited to the ovary, must be examined to look for capsular rupture (Benedet *et al.*, 2000). There is evidence that overall survival rate can be high when the transformed cells are confined the ovaries (Crayford *et al.*, 2000).

For patients with borderline tumours with an obviously limited disease (stage Ia) and normal examination of the opposite ovary conservative therapy can be proposed when there is a desire to maintain fertility (Benedet *et al.*, 2000, Vinatier *et al.*, 1996). For all other patients, the proposed surgical treatment includes total hysterectomy and bilateral salpingo-oophorectomy and all sites of tumour must be removed (Benedet *et al.*, 2000, Vinatier *et al.*, 1996). Further, the omentum, pelvic and para-aortic lymph nodes should be removed for histological examination in order to obtain an accurate staging (Benedet *et al.*, 2000, Vinatier *et al.*, 1996).

Recently two parallel randomised clinical trials (RCTs), the International Collaborative Ovarian Neoplasm 1 (ICON1) and the Adjuvant Chemotherapy in Ovarian Neoplasm (ACTION) in early-stage ovarian cancer compared platinum-

based adjuvant chemotherapy with observation following surgery. They showed that adjuvant chemotherapy may provide further benefits for women with stage I ovarian cancer (Trimbos *et al.*, 2003). ICON1 reported an improvement in overall survival of 8% and in recurrence-free survival of 11% in patients treated with adjuvant platinum-based chemotherapy compared with observation only (Trimbos *et al.*, 2003).

However, ACTION also showed that adjuvant chemotherapy significantly improved the overall and the disease-free survival only in inadequately staged patients (Trimbos and van der Burg, 2004), though this was a post hoc sub group analysis. In addition, a systematic review led by Elit *et al.* found similar results, especially when patients were not submitted lymphadenectomy as part of the surgical staging (Elit *et al.*, 2004). Therefore in the patients who had undergone optimal surgical staging, adjuvant chemotherapy may have had no effect on the prognosis (Trimbos and van der Burg, 2004, Vergote and Trimbos, 2003). Many physicians that the best policy for the treatment of patients with early stage ovarian cancer is to make efforts to achieve optimal surgical staging and to save adjuvant chemotherapy for those patients in whom optimal staging is not feasible (Trimbos and van der Burg, 2004). However, there are no randomised trials addressing optimal staging or surgery.

Laparoscopy has been restricted to patients with pre-operative evidence of a benign diagnosis (Lehner *et al.*, 1998, Vergote, 2004) The inappropriate treatment of a malignant condition by endoscopy is associated with worse prognosis(Lehner R *et al.*, 1998). Rupture of an ovarian malignant tumour should be avoided at the time of surgery for an early stage ovarian cancer (Lehner *et al.*, 1998, Vergote, 2004). Some endoscopic procedures are performed using CO2 laser techniques, and this is considered by some authors to increase the risk of activating cell enzymes which may lead to mitosis and an increase in the production of tumour growth factors. If the duration of laparoscopic surgery is prolonged there may also occur mechanical or chemical damage of the mesothelium which, in some cases of malignancy may be inadvertently treated as a benign lesion, increasing the risks of metastases in the abdominal cavity (Greene, 1995, Volz and Köster, 1999). However, reports addressing the selective use of laparoscopic techniques in the management of malignant gynaecologic disease have been published with increasing frequency (Chi and Curtin, 1999, Dottino *et al.*, 1999, Kadar, 1997b), although it still remains controversial whether laparoscopy is a good choice for the management early stage

ovarian cancer (Vergote and Amant, 2004).

It is not yet established whether laparoscopy is as good as or better than the conventional surgical approach for the treatment of ovarian tumours which are assumed to be malignant. Given the limited evidence from randomised trials in this area of surgery, and the concerns that have arisen over quality, an objective analysis of the literature requires evaluation of both randomised and non-randomised studies. We performed a systematic review to compare laparoscopy with laparotomy as surgical approaches for the treatment of early stage ovarian cancer. The conclusions of this study may help to implement management protocols validated by good levels of evidence, highlighting the needing of further research.

OBJECTIVES

The objective of this review was to evaluate the impact of laparoscopy in the surgical treatment for FIGO stage I ovarian cancer (stages Ia, Ib, and Ic) when compared with laparotomy.

The following issues were addressed in this review:

(1) Is laparoscopy (intervention group) effective in improving overall survival compared with laparotomy (control group) in patients with FIGO stage I ovarian cancer?

(2) Is laparoscopy (intervention group) effective in reducing progression-free survival compared with laparotomy (control group) in patients with FIGO stage I ovarian cancer?

(3) Does primary laparoscopy result in less surgical complications than laparotomy (control group) in patients with FIGO stage I ovarian cancer?

(4) Does primary laparoscopy (intervention group) result in more local recurrence (port site) than laparotomy (control group) in midline incision in patients with FIGO stage I ovarian cancer?

(5) Does primary laparoscopy (intervention group) result in more distant recurrence than laparotomy (control group) in patients with FIGO Stage I ovarian cancer?

(6) Does primary laparoscopy (intervention group) result more tumour spillage at the time of surgery than laparotomy (control group) in patients with FIGO stage I ovarian cancer?

Criteria for considering studies for this review

Types of studies

Inclusion criteria

Studies regarding patients with histologically proven stage I ovarian cancer according to the International Federation of Gynaecology and Obstetrics (FIGO) were included in this review.

Studies comparing laparoscopic surgery with laparotomy for early stage ovarian cancer were only available from 1990.

It was anticipated that a very small number of randomised controlled trials (RCTs) have been conducted analyzing patients with early stage ovarian cancer. Therefore, non-randomised comparative studies, cohort studies and case-control studies, but not studies with historical controls, were also considered in this review.

Histological sub grouping for malignant ovarian tumours were considered whenever possible (Scully, 1999):

(1) Surface epithelial-stromal tumours:

(a) serous type (borderline and malignant)

(b) mucinous type (borderline and malignant)

(c) endometrial type

(2) Germ cell tumours:

(a) teratoma (immature and monodermal types)

(b) dysgerminoma

(c) yolk sac tumour

(d) embryonal carcinoma

(e) carcinoid tumours

(3) Sex cord-stromal tumours:

(a) granulosa-stromal cell tumours

(b) sertoli-stromal cell tumours (androblastoma)

(c) sex cord tumour with annular tubules

(d) gynandroblastoma

(e) unclassified sex cord-stromal tumour

(f) steroid (lipid) cell tumour

Exclusion criteria

All studies regarding patients with early stage ovarian cancer who desired to remain fertile, treated by conservative surgery (unilateral salpingo-oophorectomy).

All studies where ovarian cancer was inadequately staged.

Types of participants

Patients with early stage ovarian cancer was included, i.e. patients with disease confined to the ovaries, no lymph node involvement or distant metastases.

The International Federation of Gynaecology and Obstetrics (FIGO) distinguishes patients with stage I ovarian cancer as follows (Scully, 1999):

- Stage Ia: unilateral tumours
- Stage Ib: bilateral tumours
- Stage Ic: identifies tumour spillage, tumour capsular penetration, positive peritoneal cytology

No lymph node involvement or distant metastases

Whenever possible the results were stratified by: histological subgroups of ovarian cancer.

Types of interventions

In this review two surgical approaches used for the management of FIGO stage I ovarian cancer were compared: laparoscopy (intervention group) and laparotomy (control group).

Types of outcome measures

Primary outcomes

(1) Survival at five years.

(2) Progression-free survival at five years.

Secondary outcomes

(1) Tumour spillage at the time of surgery.

(2) Local recurrence: laparoscopy (porte site) and laparotomy (midline incision).

(3) Distant recurrence.

(4) Surgical outcome:

(a) *Surgical complications (immediate and delayed):*

(i) injury (to the bladder, urether, vascular, small bowel and colon injuries);

(ii) presence/complication of adhesions;

(iii) fever;

(iv) intestinal obstruction;

(v haematoma;

(vi) infection;

(vii) rate of conversion to laparotomy.

(b) *Systemic complications:*

(i) chest infection;

(ii) deep venous thrombosis;

(iii) pulmonary embolism;

(iv) cardiac failure;

(v) cardiac ischemias;

(vi) cerebrovascular accident

(c) Operative time.

(d) Recovery from surgery: length of hospital stay and re-admission rates.

Search strategy for identification of studies

Searches were conducted to identify all published and unpublished RCTs and non RCTs comparing laparoscopy and laparotomy for early stage ovarian cancer. The search strategy were identify studies in all languages and, when necessary, non English language papers will be translated so that they could be fully assessed for potential inclusion in the review. Trials were identified by searching the Cochrane Gynaecological Cancer Group trials register, Cochrane Central Register of Controlled Trials (CENTRAL) (The Cochrane Library Issue 2, 2007), MEDLINE (January 1990 to date), EMBASE (1990 to November 2007), LILACS (1990 to November 2007), BIOLOGICAL ABSTRACTS (1990 to November 2007) and Cancerlit (1990 to November 2007).

MEDLINE was searched using the following strategies:

1. Randomized controlled trial. pt.
2. Controlled clinical trial.pt
3. Randomizes controlled trials/
4. random allocation/
5. double -blind method/
6. single-blind method/
7. or/1-6
8. clinical trial.pt
9. exp clinical trials/

10. (clin$ adj25 trial$).ti,ab,sh.
11. ((singl$ or doubl$ or tripl$ or trebl$) adj25 (blind$ or masks$)).ti,ab,sh.
12. placebos/
13. placebo$.ti,ab,sh
14. random$.ti,ab,sh.
15. Research design/
16. or/8-15
17. (animal not human).sh
18. 16 not 17
19. comparative study.sh
20. exp evaluation studies
21. follow up studies.sh
22. prospective studies
23. (control$ or prospectiv$).mp or volunter$.ti.ab.
24. exp cohort studies/
25. cohort.tw
26. exp longitudinal studies/
27. (cohort adj5 (stud$ or trial$)).tw
28. (prospectiv$ adj5 (stud$ or trial$)).tw
29. (longitudinal adj5 (stud$ or trials)).tw
30. or/18-29
31. exp Ovarian Neoplasms/
32. (ovar$ adj5 tumo?r).tw
33. (ovar$ adj5 neoplas$).tw
34. (ovar$ adj5 cancer$).tw
35. (ovar$ adj5 carcino$).tw
36. exp Adnexal Diseases/
37. exp Ovarian Cancer/
38. or/31-37
39. exp "early ovarian cancer"/
40. exp "early ovarian neoplasm"/
41. "stage I ovarian cancer"/
42. or/ 39-41
43. 38 and 42

44. exp SURGERY/
45. surg$.tw.
46. laparo$.tw.
47. exp Surgical procedures, Operative/
48. or/44-47
49. 43 and 48
50. 30 and 49

EMBASE was searched using the following strategies:

1. Controlled study/or Randomized Controlled trial/
2. double blind procedure/
3. single blind procedure/
4. crossover procedure/
5. drug comparison/
6. placebo/
7. random$.ti,ab,hw,tn,mf.
8. latin square.ti,ab,hw,tn,mf.
9. crossover.ti,ab,hw,tn,mf.
10. cross-over.ti,ab,hw,tn,mf.
11. placebo$.ti,ab,hw,tn,mf.
12. ((doubl$ or singl$ or tripl$ or trebl$) adj5 (blind$ or mask$)).ti,ab,hw,tn,mf.
13. (comparative adj5 trial$).ti,ab,hw,tn,mf.
14. (clinical adj5 trial$).ti,ab,hw,tn,mf.
15. or/ 1-14
16. nonhuman/
17. (animal not human)/
18. or/16-17
19. 15 not 18
20. comparative study.ti,ab,hw,tn,mf.
21. follow up studies.ti,ab,hw,tn,mf.
22. prospective studies.ti,ab,hw,tn,mf.
23. (control$ or prospectiv$).mp or volunteer$.ti.ab.
24. cohort studies/
25. cohort.ti,ab,hw,tn,mf.
26. longitudinal studies.ti,ab,hw,tn,mf.

27. (cohort adj5 trial$).ti,ab,hw,tn,mf.
28. (prospectiv$ adj5 trial$).ab,hw,tn,mf.
29. (longitudinal adj5 trials).ti,ab,hw,tn,mf.
30. or/19-29
31. exp Ovarian Neoplasms/
32. (ovar$ adj5 tumo?r).tw
33. (ovar$ adj5 neoplas$).tw
34. (ovar$ adj5 cancer$).tw
35. (ovar$ adj5 carcino$).tw
36. exp Adnexal Diseases/
37. exp Ovarian Cancer/
38. or/ 31-37
39. exp "early ovarian cancer"/
40. exp "early ovarian neoplasm"/
41. "stage I ovarian cancer"/
42. or/39-41
43. 38 and 42
44. exp SURGERY/
45. surg$.tw.
46. laparo$.tw.
47. exp Surgical Technique
48. or/43-46
49. 43 and 48
50. 30 and 49

CENTRAL on the current issue of The Cochrane Library, the National Research Register (NRR) and Clinical Trials register were also searched in all fields using the following words: ovarian cancer, laparotomy, laparoscopy, ovarian surgery. The citation list of relevant publications, abstracts of scientific meetings and list of included studies were checked through hand searching and experts in the field contacted to identify further reports trials. The results of handsearching of the following conferences were searched:

Gynecologic Oncology

International Journal of Gynaecological Cancer

British Journal of Cancer

British Cancer Research Meeting

Annual Meeting of the International Gynaecologic Cancer Society

Annual Meeting of the American Society of Gynecologic Oncologist

Annual Meeting of the European Society of Medical Oncology (ESMO)

Annual Meeting of the American Society of Clinical Oncology (ASCO)

METHODS OF THE REVIEW

Selection of studies

All eligible studies were assessed for their methodological quality and relevance to the review objectives. Study selection was undertaken by the reviewers. No effort was made to blind the reviewers for names of authors, institutions and journals. The reason for this is that all reviewers were very familiar with the literature on early stage ovarian cancer treatment. As it is known to us that no RCTs have been published, we decide to incorporate other types of studies in this review, i.e. cohort studies and case-control studies, but not studies with historical controls.

Assessment of methodological quality of included studies

The quality of allocation concealment for RCTs was graded as either adequate (A), unclear (B), or inadequate (C), following the detailed descriptions of these categories provided by the Cochrane Gynaecological Cancer Group.

All assessments of the quality of trial and data extraction were performed independently by five reviewers (LRM, DDR, MIR, MCB,MIE) using forms from Cochrane guidelines for selections bias and allocation concealment). The quality score was assessed by the allocation concealment method according Cochrane Collaboration Handbook (Deeks *et al*, 2005).

- Grade A: Adequate concealment
- Grade B: Uncertain concealment
- Grade C: Clearly inadequate concealment

We excluded trial with inadequate concealment and trials with quasi-randomised designs. All studies were assessed with the aid of a critical review form.

One for case-control studies and one for cohort studies (Table 2; Table 3, Table 4).

The critical review forms were filled out independently by the reviewers to assess whether the studies meet the inclusion criteria. Extracted data included trial characteristics, characteristics of the study participants, interventions and outcomes (Table 1). The quality of non randomised studies were assessed using appropriate quality evaluations tools by STROBE (von *et al.*, 2007) and Newcastle-Ottawa tool for observational studies (Wells *et al.*, 2007). A "star system" has been developed in which a study was judged based on three broad perspectives: the selection of the study groups; the comparability of the groups; and the ascertainment of the exposure and outcome of interest for case control and cohort studies. The goal of this project was to develop an instrument providing an easy and convenient tool for quality assessment of non randomised studies to be included in a systematic review (Wells *et al.*, 2007).

Differences were resolved by discussion. When paper contained insufficient information to make a decision about eligibility or when additional information was required we contacted the author/ principal investigator asking for further information.

Statistical analysis

Statistical analysis was performed in accordance to the guidelines developed by the Cochrane Gynaecological Cancer Group. All trials were initially included in one analysis of surgical laparoscopy and laparotomy for early stage ovarian cancer. Statistical heterogeneity between the results of different studies was examined by checking the usual statistic test (Cochran's Q) where P values were obtained by comparing the statistic with a chi-square distribution. Care was taken in the interpretation of the chi-squared test, since it has low power in the (common) situation of a meta-analysis when trials had small sample size or were few in number. If there was no evidence of statistical heterogeneity ($P > 0.10$), a fixed effects model was used. If there was significant heterogeneity ($P < 0.10$), the possible clinical and methodological reasons for this was explored qualitatively and a random effects model was used (Deeks, 2004, Wells *et al.*, 2007).

However, since clinical and methodological diversity always occur in a meta-analysis, statistical heterogeneity was inevitable. An alternative approach that quantifies the effect of heterogeneity is inconsistency (I2), which provides a measure of the degree of inconsistence in the studies' results with 95% uncertainty intervals (Higgins *et al.*, 2003, Wells *et al.*, 2007). This describes the percentage of the

variability in the effect estimates that is due to heterogeneity rather than sampling error (chance). A value of 0% indicates no observed heterogeneity and a value greater than 50% may be considered substantial heterogeneity. If it is inappropriate to pool the data because of clinical or statistic heterogeneity then a systematic review without a meta-analysis or a meta-analysis excluding outlying studies may be performed.

If sufficient trials of adequate quality are available and their populations are clinically similar, meta-analyses of primary and secondary end-points may carried out. For meta-analyses of the time-to-event outcomes, the most appropriate statistic is the hazard ratio (HR) and when this is provided in a trial report, it should be used. When the HR is not provided, then it can be estimated indirectly from other summary statistics when possible (Parmar *et al.*, 1998, Wells *et al.*, 2007). When this estimative is not possible, the odds ratio may be calculated and interpreted with caution, bearing in mind the possibility of mortality/morbidity and hence the odds ratio, which is influenced by length of follow-up. For meta-analyses of dichotomous outcomes, relative risks (RR) will be calculated with 95% confidence intervals (CIs) and combined for meta-analysis with RevMan software.

Continuous data were combined for meta-analysis. We used means and standard deviations to derive a weighted mean difference (WMD) with 95% CIs using a fixed effect model. As a general rule, a fixed effect model were used for calculations of summary estimates and their 95% CIs unless there was significant heterogeneity in which case results were confirmed using a random effects statistical model. Difficulties were encountered when reporting continuous outcomes, (for example blood loss during surgery), since the data was skewed and authors correctly presented their data as median with range. Whenever possible original data were obtained from the authors but posttreatment means and standard deviations were not always available of calculated. When only medians and ranges were available, the median was regarded as being identical to the mean and a crude estimate of the standard deviation was calculated from the range ((range X 0.95)/4). This method is not ideal for skewed data and is likely to result in an over-estimation of the SD but it was planned a priori to perform a sensitivity analysis with and without inclusion of these trials in the meta-analysis. Without exception, the distribution of data on menstrual blood loss was skewed and the ideal and most appropriate statistical analysis was nonparametric rather than inclusion in a meta-analysis assessing a

weighted mean difference.

Whenever possible, subgroup analyses were performed to compare the study results according to type of intervention, histological subtypes, study design and quality of report (adequate versus unclear allocation concealment for RCTs).

Description of studies

The initial search identified 706 citations, of which 663 were excluded and 43 were retrieved for detailed examination. Only three published trials met the inclusion criteria, one cohort (Tozzi *et al.*, 2004) and two case and controls studies (Ghezzi *et al.,* 2007, Hua *et al.*, 2005) (Figure 1). We did not find RCTs.

EXCLUSION AND REASONS

Please see the Table of Characteristics of Excluded studies (Anexo J). Thirty nine studies were excluded often for more than on reason. The most common causes for exclusion were design other than a randomised controlled trial, a narrative review, a series of cases or a retrospective studies.

1. Settings

The three included studies were of single-centre design, conducted, in Italy (Tozzi *et al.*, 2004, Ghezzi *et al.*, 2007) and China (Hua *et al.,* 2005). Hua *et al.* was translate from Chinese to Portuguese (Anexo L).

2. Designs

One case-control study compared ten consecutive patients submitted to laparoscopy for early stage ovarian cancer with eleven patients with the same diagnosis who underwent laparotomy (Hua *et al.,* 2005). Another case-control study compared 15 patients with early stage ovarian cancer submitted to laparoscopy with another group of 19 patients with the same diagnosis submitted to laparotomy (Ghezzi *et al.*, 2007).

We found one prospective cohort study with 42 patients eligible to enter the study with ovarian cancer FIGO stage IA- IB and follow up around 46.4 months (SD 16.25; range 2-72 months), initially submitted to laparoscopy (Tozzi, 2004). However, 18 patients were excluded due to tumour rupture (n = 5) not explained if occurred during or before surgery, presence of peritoneal implants (n = 3), ovarian surface

invasion (n = 4), or microscopic invasion at frozen section analysis (n = 2), tumour size large than 11 x 8 cm, which is the largest diameter of the endobag (n = 4). In these, 18 cases laparoscopy was converted to laparotomy.

3. Participants

All women included in the trials had malignant ovarian tumours and underwent a preliminary workup, including ultrasonography, CA 125, and colour Doppler ultrasonography followed by surgery (laparoscopy or laparotomy). In the cohort study (Tozzi *et al.,* 2004) 24 patients were submitted to laparoscopy, the median age was 36.8 years (SD 13.5, range 19-76) and histological results were as follows: 20.8% serous, 12.5%, respectively, were mucinous, and dysgerminoma, 8,3% were endometrioid, and 4,1%, respectively, clear cell, yolk sack tumour, teratoma and granuloma cells. Histological grading was G1, G2 and G3, in 50%, 33% and 16,6%, respectively. Tumour stage was 50% IA, 20.8% IB and 29.2% IC (Tozzi *et al.*, 2004).

Thirty-four patients with apparent early stage ovarian cancer was submitted to surgery. In the case-control study (Ghezzi *et al.*, 2007) 15 patients undergoing a comprehensive laparoscopic stating were compared with 19 patients that were submitted to laparotomy. Age in the laparoscopy group: 55 years (SD 13.5; range 13-70). Age in the laparotomy group: 61 years (SD 0.58; range 44-70), the body mass index 23.8 (SD 4.2) in the laparoscopy group and 25.8 (SD 3.1) in the laparotomy group. The histopathologic study of the surgical specimens in the laparoscopic group showed: 7 serous cystadenomas, 3 mucous cystadenomas, 3 endometrioid tumours, 1 dysgerminoma and 1 carcinosarcoma. Tumour was - 53.3% G2 and 46.6% G3. In the laparotomy group there were: 14 serous cystadenomas, 2 mucous cystadenomas, 1 endometrioid tumour, 1 cell carcinoma in a mature teratoma. Tumour grading was 5.2% G1, 26.3 G2 and 68.4% G3. Final stage in the laparoscopy group: Ia (n = 5); Ic (n = 6), IIIa (n = 2), IIIc (n = 2); final stage in the laparotomic group: Ia (n = 8); Ic (n = 5), IIIa (n = 3), IIIc (n = 3). Controls were selected from consecutive women who underwent laparotomy for an apparent early stage ovarian cancer between 1997 to 2003, and who met the same criteria for eligibility as the laparoscopy group. Patients were operated in all cases by the same surgeons, with extensive training and experience both in gynaecologic oncology and in advances laparoscopic procedures. All patients received a single dose of prophylactic antibiotic 1 h prior to the intervention (ampicillin/sulbactam 1.5 g

intravenously) as well as anti-thrombotic prophylaxis with heparin (Ghezzi *et al*., 2007).

In the case-control study (Hua *et al.*, 2005) 10 patients with early stage ovarian cancer underwent laparoscopic surgery and 11 patients with the same diagnosis underwent laparotomy. Age in the laparotomy group:42 (SD 6). Age in laparoscopy group: 40 years (SD 8). In the laparoscopy group nine were epithelial tumor and one was stromal; in the laparotomy group nine were epithelial and two cases were stromal, and all cases had tumour grading G3 (Table 1).

4. Interventions

In the cohort study patients in the laparoscopy group were submitted to bilateral salpingo-oophorectomy with laparoscopic assisted vaginal hysterectomy, pelvic lymphadenectomy, infrarenal para-aortic lymphadenectomy, complete resection of the infundibulopelvic-pelvic ligament, appendectomy and partial omentectomy (Tozzi *et al.*, 2004). In the case-control studies, the patients in the laparoscopy group were submitted to bilateral salpingo-oophorectomy with laparoscopic assisted vaginal hysterectomy, pelvic lymphadenectomy, infrarenal para-aortic lymphadenectomy, complete resection of the infundibulopelvic-pelvic ligament, appendectomy and partial omentectomy; in the laparotomy group the patients were submitted the same procedures (Ghezzi *et al.*, 2007, Hua *et al.*, 2005). Frozen section analysis was performed in all included studies (Table 1).

5. Outcomes

In the cohort study: survival at five years, progression free survival at five years, intraoperative complications, blood transfusions, operative time for completeness of staging and primary surgery, number of pelvic lymph nodes resected, presence of trocar site metastasis (Tozzi *et al.*, 2004).

In the case-control study: survival, progression free survival at five years, operative time, intraoperative blood loss, intraoperative complications, number of pelvic lymph nodes resected (Ghezzi *et al.*, 2007, Hua *et al.,* 2005).

Methodological quality of included studies

Inter-rater overall agreement for quality assessment was good (Cohen's kappa = 0.78). Initial disagreements were solved through discussion in all cases. The quality

of non-randomised studies was assessed using appropriate quality evaluations tools by Newcastle-Ottawa for observational studies (Wells *et al.*, 2007) (Table 2 and Table 3) and by Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) (von *et al.*, 2007) (Table 4).

1. Cases-control studies (Ghezzi *et al.*, 2007, Hua *et al.*, 2005) (Newcastle-Ottawa):

Selection

(a) Is the case definition adequate? Yes,with independent validation=yes*

(b) Representative of the cases: consecutives series of cases=yes*

(c) Selection controls: hospital control with the same disease

(d) Definition of controls: history of disease with the same diagnosis that in the laparoscopy group (treatment)=yes*

Comparability

(a) controls for early stage ovarian cancer treated by laparoscopy or laparotomy the most important factor was selected (survival) = Yes* for Ghezzi *et al.,* 2007, and No for Hua *et al.*, 2005

(b) controls for any additional factor (surgery complications, operative time, blood transfusion, number of pelvic lymph nodes resected, trocar site metastasis, recurrence) = yes*

Exposure

(a) Assessment of outcome: secure record (surgical records) = yes*

(b) Same method of ascertainment for cases and controls = yes*

(c) Non-response rate: rate different and no designation

We found six stars in the two case-control studies. The studies have the good quality by Newcastle-Ottawa (Table 2).

2. Cohort study (Tozzi *et al.*, 2004) (Newcastle-Ottawa):

Selection

(a) Representativity of the exposed cohort: the population represented the average patients with early stage ovarian cancer describe in the community = yes*.

(b) Draw from the same community as the exposed cohort = yes*

(c) Ascertainment of exposure: surgical records = yes*

(d) Demonstration that the outcome of interest was not present at the start of the study = yes*

Comparability

(a) controls for early stage ovarian cancer treated by laparoscopy or laparotomy - the most important factor was select (survival) = yes*

(b) controls for any additional factor (surgery complications, operative time, blood transfusion, number of pelvic lymph nodes resected, trocar metastasis, recurrence) = yes*

Outcomes

(a) Assessment of outcome: record linkage = yes*

(b) Was the follow up long enough for the occurrence of outcomes? yes (46,4 months) = yes*

(c) Adequacy of follow up of cohorts

Complete follow up - all subjects accounted for = yes*

We found nine stars in the cohort study, the corresponding to an excellent quality by Newcastle-Ottawa (Table 3).

Evaluation of quality by STROBE (Strengthening the Reporting of Observational studies in Epidemiology) (von *et al.*, 2007) of three included studies. From a checklist of 22 items, all three studies showed problems with 7 items: title, variables, bias, statistical methods, participants, main results and other analysis (Table 4).

The studies (Ghezzi *et al.*, 2007, Hua *et al.*, 2005, Tozzi *et al.*, 2004) did not give information about potential confounders and effects modifiers (item 7) and did not describe potential sources of bias (item 9). In the statistical methods there was no description for the control of confounding factors and, the subgroups and interactions were not described in the sensitivity analysis (item 12). In the results there was no consideration for the use of a flow diagram (item 13). In the main results unadjusted estimates were not given and, when applicable, confounder-adjusted estimate and their precision were not describe (e.g., 95% confidence intervals) (item 16). There was no report on subgroup analysis, interactions and sensitivity analysis (item 17). When describing the limitations of the studies was no description of sources,

directions and magnitude of potential bias. Conclusion: there were problems in important items in the STROBE, for considering a study of good quality.

RESULTS

A meta-analysis was not possible due to difference among studies and because STROBE has a low quality for evaluations of important outcomes. Therefore, we performed a qualitative systematic review. We used three selected observational studies (two case-control and one cohort study) (Ghezzi *et al.*, 2007, Hua *et al.*, 2005, Tozzi *et al.*, 2004). These three studies met the inclusion criteria with a total 97 patients with early stage ovarian cancer.

COHORT STUDY (Tozzi *et al.*, 2004). Forty-two patients were eligible for the study and were submitted to laparoscopy. In, 18 of these patients there were conversion to laparotomy due to tumour rupture in 5 case, peritoneal implants in 3 cases, ovarian surface invasion in 4 cases, and microscopic invasion at frozen section analysis was in 2 cases, and large tumour size in 4 cases.

Primary outcomes

(1) *Survival at five years*

In (Tozzi *et al.*, 2004) the survival for 24 patients at maximum follow up 72 months (median 46.6, SD 16.25; range 2-72 months) and overall survival was 100%.

(2) *Progression free-survival at five years*

In Tozzi *et al.*, 2004 2 patients (8.3%) had tumour recurrence. One patient, primarily treated with surgery and chemotherapy for an epithelial ovarian cancer FIGO IB G3, had a pelvic recurrence and underwent secondary surgery with debulking and bowel resection followed by second-line chemotherapy. The second patients with a diagnosis of epithelial ovarian cancer FIGO IA G3 received six cycles of platinum and paclitaxel because of positive peritoneal biopsies at second-look laparoscopy. Disease-free survival was 91.6% in 24 cases.

Secondary outcomes

(1) Tumour spillage at the time of surgery: from forty-two patients eligible patients tumour spillage occurred in 5 (11.9).

(2) Local recurrence: for the laparoscopy (porte-site) and laparotomy (midline

incision) groups: until 2004 there was no trocar site metastasis in 24 cases. In 15 out of 24 patients (62.5%) a second-look laparoscopy was performed, and local recurrence occurred in one case.

(3) Distance recurrence: until 2004 there were no distant recurrences, there were two pelvic recurrences out of 24 cases.

(4) Surgical outcome

(a) *Complications (immediate and delayed):*

(i) Injury (bladder, urether, vascular, small bowel and colon injuries): It did not occur.

(ii) Presence /complications and adhesions: not describe.

(iii) Fever: not describe.

(iv) Intestinal obstruction: did not occur.

(v) Haematoma: did not occur.

(vi) Infection: not describe.

(vii) Convertion to laparotomy: from 42 patients initially eligible to laparoscopy, 18 were submitted to a laparotomy (42.8%).

(vii) Systemic complications: one patient developed chylous ascites with spontaneous evacuation of the lymphatic fluid through abdominal drainage, the patients was discharged 12 days after surgery. There were no cases of chest infection, deep venous thrombosis, pulmonary embolism, cardiac failure, cardiac ischemias or cerebrovascular accident.

(b) *Operative time:* mean operative time for all patients was 176 min (SD 48.45; range:102-306 min); it a took a mean of 166 min (SD 20.9 range 118-206 min) for complete stating and 182 (SD 39.18; range 141-306) for the primary treatment of patients who underwent LAVH (laparoscopy assisted vaginal hysterectomy) with contralateral salpingo-oophorectomy.

(c) *Recovery from surgery*:

(i) length of hospital stay: The overall mean length of hospital stay was 7 days (SD 1.66; range 5-12). Adjuvant treatment was proposed in five cases.

(ii) re-admission rates: not described.

(d) *Mean number of bilateral pelvic lymph nodes*:19,6 (SD 7.1; range 5-35).

CASE-CONTROL (Hua *et al.,* 2005). Ten patients with early stage ovarian cancer were submitted to laparoscopic total hysterectomy, pelvic lymph nodes dissection, bilateral adnexectomy, high ligation of the ovarian aortic and vein, omentectomy, and additional appendectomy. Eleven patients with the same diagnosis underwent the same procedure by laparotomy.

Primary outcomes

(1) Survival at five years: not reported

(2) Progression free-survival at five years: not reported

Secondary outcomes

(1) Tumour spillage at time of surgery:

Laparoscopy: yes, all cases by vaginal puncture ovarian tumour

Laparotomy: not reported.

(2) Local recurrence: laparoscopy (porte site) and laparotomy (midline incision):

Laparoscopy:did not report.

Laparotomy:did not report.

(3) Distance recurrence:

Laparoscopy: did not report.

Laparotomy: did not report.

(4) Surgical outcome:

(a) *Complications (immediate and delayed):*

(i) Injury (bladder, urether, vascular, small bowel and colon injuries):

Laparoscopy: The right obturator nerve was injured and was sutured.

Laparotomy: did not occur.

(ii) Presence /complications and adhesions:

Laparoscopy: right obturador nerve was injured.

Laparotomy: one case of urinary retention, one case of chylous ascites.

(iii) Fever:

Laparoscopy: did not report.

Laparotomy: not reported.

(iv) Intestinal obstruction:

Laparoscopy: did not occur.

Laparotomy: did not occur.

(v) Haematoma:

Laparoscopy:did not occur.

Laparotomy:did not occur.

(vi) Infection:

Laparoscopy:did not occur.

Laparotomy: one case of wound infection.

(vii) Conversention to laparotomy: did not occur.

(vii) Systemic complications:

Laparoscopy: did not occur.

Laparotomy: did not occur.

(viii) Blood loss

Laparoscopy: 280 ml (SD 156 ml),

Laparotomy: 346 ml (SD 170 ml). There were statistically significantly differences in blood loss between the two groups ($p < 0.05$).

(b) Operative time:

Laparoscopy: 298 min (SD 60 min)

Laparotomy: 182 min (SD 43 min). There were statistically significantly differences the two groups ($p < 0.05$)

(c) *recovery from surgery:*

(i) length of hospital stay:

Laparoscopy: did not report.

Laparotomy: did not reported.

(ii) re admissions rate:

Laparoscopy:did not reported.

Laparotomy:did not reported.

(d) *Mean number of bilateral pelvic lymph nodes*

Laparoscopy: 25 (SD 5)

Laparotomy: 27 (SD 7).There were no statistically significantly differences between the two groups ($p > 0.05$)

CASE-CONTROL (Ghezzi *et al.*, 2007). 15 patients with early stage ovarian cancer were submitted to laparoscopic total hysterectomy, pelvic lymph nodes dissection, bilateral adnexectomy, high ligation of the ovarian aortic and vein, omentectomy, and additional appendectomy. Nineteen patients with the same

diagnosis who underwent the same procedure by laparotomy served as a control group.

Primary outcomes

(1) Survival at five years: The laparoscopy group had at least 2 years of follow up with 100% survival. For the laparotomy group the survival was also 100%, but the follow up time had a median of 29 months (SD 18,5; range 14-92).

(2) Progression free-survival at five years: in the laparoscopy group there were no recurrences and in the laparotomy group there were 4 recurrences (7.1%).

Secondary outcomes

(1) Tumour spillage at the time of surgery:

Laparoscopy: in three cases.

Laparotomy: in two cases.

(2) Local recurrence: laparoscopy (porte site) and laparotomy (midline incision):

Laparoscopy:not reported.

Laparotomy:not reported.

(3) Distance recurrence:

Laparoscopy: not reported.

Laparotomy:not reported.

(4) Surgical outcome:

(a) *Complications (immediate and delayed):*

(i) Injury (bladder, urether, vascular, small bowel and colon injuries):

Laparoscopy:not reported.

Laparotomy: not reported.

(ii) Presence /complications and adhesions:

Laparoscopy:not reported

Laparotmy:not reported

(iii) Febrile morbidity:

Laparoscopy: not reported.

Laparotomy: not reported.

(iv) Intestinal obstruction:

Laparoscopy:not reported.

Laparotomy:not reported.

(v) Haematoma:

Laparoscopy: retroperitoneal haematoma requiring laparotomy and hypogastric arteries ligature occurred 7 h after surgery.

Laparotomy: there were no haematomas.

(vi) Infection:

Laparoscopy: not reported.

Laparotomy: 6 cases: urinary infection (n = 4) and wound infection (n = 2).

(vii) Convertion to laparotomy: occurred 7 h after surgery.

(vii) Systemic complications:

Laparoscopy: not reported.

Laparotomy: not reported.

(viii) Blood loss

Laparoscopy: 250 ml (SD 225; range 50-1000) in one patient who had a retroperitoneal haematoma and had to received six units of packed red blood cells plus 4 four units of fresh frozen blood.

Laparotomy: 400 ml (SD 201; range 150-1000). Not significant.

(b) *Operative time:*

Laparoscopy: 377 min (SD 47 min)

Laparotomy: 272 min (SD 81 min). There were statistically significance differences between two groups ($P < 0.05$).

(c) *recovery from surgery:*

(i) length of hospital stay:

Laparoscopy: 3 days (SD 2.3; range 2-12).

Laparotomy: 7 days (SD 2.3; range 4-14). There were statistically significance differences between two groups ($P < 0.001$).

(ii) re-admissions rate:

Laparoscopy: not reported.

Laparotomy: not reported.

(d) *Mean number of bilateral pelvic lymph nodes*

Laparoscopy: 25.2 (SD 9.3)

Laparotomy: 25.1(SD 5.8).There were no statistically significant differences between the two groups ($p > 0.05$).

We describe the differences between two cases e controls studies (Ghezzi *et al.*, 2007, Hua *et al.*, 2005) in table 5.

DISCUSSION

The main problem when trying to conduct a systematic review on surgical management by laparoscopy or laparotomy in patients with early stage ovarian cancer is that a rare disease. It is not realistic to expect a larger number of randomised controlled trials. We found only three observational studies with good quality by Newcastle – Ottawa tool (Wells *et al.*, 2007), although they had important problems in the STROBE checklist (von *et al.*, 2007). Good reporting reveals the strengths and weaknesses of a study and facilitates interpretations and applications of the results. In this systematic review a meta-analysis was not possible due to differences in the quality among studies. Therefore we performed a qualitative systematic review. Egger *et al.* showed that meta-analysis of observational data may produce precise but spurious results. The statistical combination of data should therefore not be an important component of systematic reviews of observational studies (Egger *et al.*, 2001a).

However, clinical decisions may still be made on the basis of evidence derived from non-randomised observational studies, such as cohort and case-control studies. Although observational studies may provide useful results, they are limited due to unrecognized confounding factors, which may distort results (bias). Concato *et al.,* showed that results of well-designed observational studies do not overestimate the magnitude of the effects of treatments systematically as compared to results from randomised controlled trials on the same topic (Concato *et al.,* 2000).

Controversy has arisen between endoscopists and oncologists about the laparoscopic management of early stage ovarian tumours. Kinderman *et al.* wrote that 39% of the stage Ia ovarian cancer may spread after endoscopic procedures, demonstrating implant and metastases, even in an early follow up phase. It was harmful for the majority of patients when the subsequent laparotomy indicated due to very early implants and metastases in the pelvis, in the abdominal cavity or in the laparoscopic trocar site was delayed for more than 8 days after the endoscopic procedure (Klidermann *et al.*, 1995). For Ramirez *et al.*, laparoscopic port-site metastases are potential complications of laparoscopy in patients with gynaecological cancer (Ramirez *et al.*, 2004).

Gleeson *et al.*, Childers *et al.,* Leminen *et al.* and Kadar *et al.*, reported cases of abdominal wall metastases from ovarian cancer after laparoscopy (Childers *et al.*,

1994, Gleeson *et al.*, 1993, Leminen and Mage, 1999). Romagnolo *et al.,* described tumour rupture or spilling during surgery, with a statiscally significant greater incidence in the group of patients treated by laparoscopy (34.6%) when compared to laparotomy (6.6%), p < 0.0001 (Romagnolo *et al.*, 2006). Dembo *et al.,* performed a multivariable analysis by Cox Regression for survival analysis and found the following prognostic factors: grade, adherences and ascites. In their analysis capsular rupture, stages Ia, Ib, size and age were not significant factors for survival (Dembo *et al.*, 1990). In addition Sjövall *et al.* did not find differences in survival between patients whose tumours had intact capsules and those in whom rupture occurred during surgery (78 and 85%, respectively) (Sjovall *et al.*, 1994). However, when the rupture occurred before surgery survival was only 59% (Sjovall *et al.*, 1994). Volz *et al.*, analyzed animal models microscopically and showed, that induction of a pneumoperitoneum caused diffuse damage to the entire mesothelial cell layer with exposure of the basal lamina and development of extensive mechanisms of repair. The exposure of the extracellular matrix proteins including laminin, fibronectin, and vitronectin to the tumour cell surface is a possible mechanism for increased tumour cell adherence. A second mechanism may be the promotion of intraperitoneal tumour cell growth by increased interleukin 1 production by the peritoneal macrophages, which are extensively involved in this unique repair mechanism (Volz *et al.*, 1999).

The guidelines for epithelial ovarian carcinoma FIGO stage I included both surgical and adjuvant therapeutic procedures (Sijmos *et al.,* 2007). Recently two parallel randomised clinical trials, the ICON1 and the ACTION trials showed in that adjuvant chemotherapy would provide further benefits for women with stage I ovarian cancer (Trimbos *et al.*, 2003). There are still no consensus on how to separate patients with surgical stage I disease who are at a higher risk of tumour recurrence and death from those with a low risk. Histological grade is considered one the most important prognostic factors in stage I epithelial ovarian cancer (Vergote *et al.*, 2001).Tumour rupture, capsular penetration and dense adhesions are generally believed to be associated with worse prognosis in these cases (Vergote *et al.*, 2001). Obermair *et al*, analysed 456 patients, with Grade 3 stage I ovarian cancer and found an, overall survival in five years of 87 % (95%CI 80.3-93.6); if CA 125 was higher than 30 U/ml overall survival in five years was 86% (95% CI 81.8-90.9) (Obermair *et al.*, 2007). In the staging ovarian cancer microscopic assessment of grade provide a better discrimination for the necessity of further interventions than blind biopsies.

Grade and ploidy may be surrogates for genetic instability, which may be the principal determinant of prognosis. With the publication of the ICON 1 and ACTION trials plus other evidences in the literature in last few years, tumoral grade achieved the power to determine adjuvant treatment in early stage ovarian cancer and should now be incorporate to stage for treatment decisions (Editorial, 2003).

In this systematic review, the cohort study showed of 100% with 2 recurrences, and a follow up ranging 2 to 72 months; 42.2% were eligible initially to laparoscopy but had conversion to laparotomy. The operative time was 176 min (SD 20.9) in the laparoscopy group (Tozzi *et al.*, 2004). In case-control studies, Hua *et al.,* showed an operative time for the laparoscopy group of 298 min (SD 60 min) and Ghezzi *et al.*, found an time of 377 min (SD 47 min). In vitro, the ovarian carcinomatous cells exposed to carbon dioxide for 3 hours had a 52% increase in growth by 4 days after exposure (Ghezzi *et al.*, 2007, Hua *et al.*, 2005). This increased cell growth had a linear relationship with the length of exposure to carbon dioxide when compared to now-exposed control cells (Smidt *et al.*, 2001). Three major pathways exist for the dissemination of ovarian malignancies: via bloodstream, via lymphatic channels, and spread through the abdomen and pelvis as a result of rupture of the ovarian capsule (Sugarbaker *et al.*, 1996). For Greene *et al.,* the mechanical effect of pneumoperitoneum and the probable result of the pressure may cause cellular dissemination. The effects of this mechanical dissemination in an already immunocompromised host sets up an ideal mechanism for growth that may be observed early in the postoperative evaluation of the violated abdominal wall (Greene, 1995).

On the other hand, there are a number of reports in the literature describing the use of operative laparoscopy in patients with early stage ovarian cancer (Childers *et al.*, 1994). Pelvic and para-aortic laparoscopic lymphadenectomy, appear to be feasible and adequate, although there may occur a mechanical effect caused by the pneumoperitoneum damage in the mesothelial cell layer. According to the FIGO, the prognoses of all ovarian tumours are independently affected by the following factors: stage of cancer at diagnosis, histological subtype grade and volume of residual disease after surgery (Benedet *et al.*, 2000). Therefore, the staging laparotomy is the most important part of the early management of ovarian tumours. Benedet *et al.*, showed that laparoscopy is more appropriate if the suspicion favours a benign diagnosis in a young woman with normal levels of tumoral (Benedet *et al.*, 2000).

Canis *et al.* showed that the incidence of spread of ovarian cancer after laparoscopy surgery is difficult to establish and the prognostic relevance of trocar site metastasis is not known (Canis *et al.*, 2001). The authors concluded that the laparoscopic management of ovarian cancer remains controversial and that; it should be performed only in prospective clinical trials (Canis *et al.*, 2001). Until the results of such studies become available, an immediate vertical midline laparotomy remains the gold standard if a malignant tumour is found (Canis *et al.,* 2001).

Reviewers' conclusions

Implications for practice

According to FIGO the primary surgery for patients with early stage ovarian cancer should be a vertical abdominal incision, with sampling from the peritoneal fluid and, the entire peritoneal surface of the abdominopelvic wall; inspection and palpation of the cavity, from the pelvis to the diaphragm, is recommended the search for tumoral implants (Benedet *et al.*, 2000). The abdominal organs should be inspected and the sizes of all lesions should be reported. Random biopsies of the pelvic peritoneum, abdominal peritoneum (including paracolic gutters) and bilateral para-aortic and pelvic nodes might be performed. The total hysterectomy plus bilateral salpingo-oophorectomy, omentectomy and systematic appendicectomy is controversial, although in the cases of mucinous tumours, 8% of the appendices are involved (Benedet *et al.*, 2000).

We did not find any good evidence for the recommendation of laparoscopy for the routine management of patients with early stage ovarian cancer. This review does not support the use of laparoscopy in the routine practice for the management early stage ovarian cancer.

Implications for research

There are very few trials in this files. Further trials should carefully address the methods of randomisation as blinding with is not possible in these kind of studies. Future research should include specific patient subgroups and include additional outcomes such as surgical efficacy, tumour recurrence, patient satisfaction, quality of life, costs, survival at 5 years and progression free-survival at five years. The follow up period should provide more information on recurrence, and on the potentially harmful effects of laparoscopy. For evaluation of costs it would be helpful if it were

reported separately for the preoperative, intraoperative and postoperative periods.

Survival data for patients with gynaecologic malignancies managed by laparoscopy are still lacking. It is imperative that the survival is not compromised by employing new surgical techniques. These and other important issues should be addressed by future trials before the role of laparoscopy in gynaecological oncology can be determined.

Acknowledgements

We would like to thank the members of the Cochrane Gynaecological Cancer Review Group based at the University Royal United Hospital (Wolfson Centre, Combe Park, Bath, UK), for their help, advice and support during the witting of this review: G Quinn, CJ Williams, G Quinn, C Jess C, A Oestmann A, J Kite.

Potential conflict of interest

None to disclose.

Reference List

Benedet J L, Bender H, Jones H, III, Ngan H Y, Pecorelli S. FIGO staging classifications and clinical practice guidelines in the management of gynecologic cancers. FIGO Committee on Gynecologic Oncology. Int J Gynaecol Obstet 2000; (70): 209-262.

Burger C W, Prinssen H M, Baak J P, Wagenaar N, Kenemans P. The management of borderline epithelial tumors of the ovary. Int J Gynecol Cancer 2000; (10): 181-197.

Canis M, Rabischong B, Botchorishvili R, Tamburro S, Watiez A, Mage G, Pouly JL, *et al.* Risk of spread of ovarian cancer afher laparoscopic surgery. Curr Opin Obstet Gynecol 2001; (13): 9-14.

Chi D S, Curtin J P. Gynecologic cancer and laparoscopy. Obstet Gynecol Clin North Am 1999; (26): 201-215.

Childers J M, Aqua K A, Surwit E A, Hallum A V, Hatch K D. Abdominal-wall tumor implantation after laparoscopy for malignant conditions. Obstet Gynecol 1994; (84): 765-769.

Concato J, Shah N, Horwitz R I. Randomized, controlled trials, observational studies, and the hierarchy of research designs. N Engl J Med 2000; (342): 1887-1892.

Crayford T J, Campbell S, Bourne T H, Rawson H J, Collins W P. Benign ovarian cysts and ovarian cancer: a cohort study with implications for screening. Lancet 2000; (355): 1060-1063.

Deeks JJ H J A D e. Analysing and presenting results. In: Cochrane Reviewers' Handbook 4.2.2 [updated March 2004]. (Ed.In: Alderson P GSHJe). Oxford: Cochrane Library, 2004.

Deeks JJ, Higgins, JPT, Altman DG, editors. Analysing and presenting results. In: Higgins JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions 4.2.5 [updated May 2005]; Section 6. http://www.cochrane.org/resources/handbook/hbook.htm (accessed 31st May 2005).

Dembo A J, Davy M, Stenwig A E, Berle E J, Bush R S, Kjorstad K. Prognostic factors in patients with stage I epithelial ovarian cancer. Obstet Gynecol 1990; (75): 263-273.

Dottino P, Levine DA, Ripley D C C. Laparoscopic Managment of adnexal Masses in premenopausal and Postmenopausal women. Obstet Gynecol 1999; (93): 223-7.

Editorial. Early ovarian cancer - time for a retthink on stage? Gynecol Oncol 2003; (90): 253-257.

Egger M, Smith GD, Schneider M. Sytematic reviews of observational studies. In: Systematic reviews in Health Care - Meta-analysis in context. (Eds.Egger M, Smith GD, Schneider M). London: BMJ Publishing, 2001a; 2nd: 211-227.

Egger M, Smith G D, Sterne J A. Uses and abuses of meta-analysis. Clin Med 2001b; (1): 478-484.

Elit L, Chambers A, Fyles A, Covens A, Carey M, Fung M F. Systematic review of adjuvant care for women with Stage I ovarian carcinoma. Cancer 2004; (101): 1926-1935.

Ghezzi F, Cromi A U S, Bergamini V, Tomera S, Franchi M, Bolis P. Laparoscopy versus laparotomy for the surgical management of apparent early stage ovarian cancer. Gynecol Oncol 2007; (105): 409-13.

Gleeson NC, Nicosia SV, Mark JE, Hoffman MS, Cavanagh D. Adominal wall metastases from ovarian cancer after laparoscopy. Am J Obstet Gynecol 1993; (169): 522-3.

Greene FL. Principles of cancer biology in relayion to minimal acess surgical techniques. Semin Laparosc Surg 1995; (2): 155-157.

Hand R, Fremgen A, Chmiel J S, Recant W, Berk R, Sylvester J, Sener S. Staging procedures, clinical management, and survival outcome for ovarian carcinoma. JAMA 1993; (269): 1119-1122.

Higgins J P, Thompson S G, Deeks J J, Altman D G. Measuring inconsistency in meta-analyses. BMJ 2003; (327): 557-560.

Hua KQ, Jim FM, Xu H Z Z L J F Y. [Evaluation of laparoscopic surgery in the early stage malignant tumor of ovary with lower risk]. Zhonghua Yi Xue Za Zhi 2005; (85): 169-72.

Kadar N. Laparoscopy managment of gynecological malignancies. Curr Opin Obstet Gynecol 1997a; (9): 247-255.

Kadar N. Port site recurrence following laparoscopic operations for gynaecological malignancies. Br J Obstet Gynaecol 1997b; (104): 1308-13.

Klidermann G, Massen V, Kuhn W. [Laparoscopic preliminary surgery of ovarian malignancies. Experiences from 127 German gynecologic]. Geburtsshife Frauenheilkd 1995; (55): 687-94.

Kosary C L. FIGO stage, histology, histologic grade, age and race as prognostic factors in determining survival for cancers of the female gynecological system: an analysis of 1973-87 SEER cases of cancers of the endometrium, cervix, ovary, vulva, and vagina. Semin Surg Oncol 1994; (10): 31-46.

Lehner R, Wenzl R, Heinzl H, Husslein P, Sevelda P. Influence of delayed statging laparotomy after laparoscopy removal of ovariam masses later found malignant. Obstet Gynecol 1998; (92): 967-71.

Leminen A, Mage G. Spread of ovarian cancer after laparoscopic surgery report of eight cases. Gynecol Oncol 1999; (75): 387-390.

Medeiros L R, Rosa D D, Edelweiss M I, Stein A T, Bozzetti M C, Zelmanowicz A, Pohlmann P R, Meurer L, Carballo M T. Accuracy of frozen-section analysis in the diagnosis of ovarian tumors: a systematic quantitative review. Int J Gynecol Cancer 2005; (15): 192-202.

Obermair A, Fuller A L-V E v G T, Vergote I, Eaton L F, *et al.* A new prognostic model for FIGO stage 1 epithelial ovarian cancer. Gynecol Oncol 2007; (104): 607-11.

Parmar M K, Torri V, Stewart L. Extracting summary statistics to perform meta-analyses of the published literature for survival endpoints. Stat Med 1998; (17): 2815-2834.

Ramirez P T, Frumovitz M, Wolf J K, Levenback C. Laparoscopic port-site metastases in patients with gynecological malignancies. Int J Gynecol Cancer 2004; (14): 1070-1077.

Romagnolo C, Gadducci A, Sartori E, Zola P, Maggino T. Management of bordeline ovarian tumors:results of an Italian mulcenter study. Gynecol Oncol 2006; (101): 255-260.

Scully RE. Histological typing of ovarian tumours - World Health Organization International histological classification of tumors. Springer-Verlag, Berlin 1999.

Sijmos EA, van Lankveld AL, Witteveen PO, Peeters PHM, Kooot VCM, van Leeuwen JS. Compliance to clinical guidelines for early-stage epithelial ovarian cancer in relation to patients outcome. Eur J Obstet Gynecol Reprod Biol 2007; (131): -203.

Sjovall K, Nilsson B, Einhorn N. Different types of rupture of the tumor capsule and the impact on survival in early ovarian carcinoma. Int J Gynecol Cancer 1994; (4): 333-336.

Smidt V J, Singh D M, Hurteau J A, Hurd W W. Effect of carbon dioxide on human ovarian carcinoma cell growth. Am J Obstet Gynecol 2001; (185): 1314-1317.

Sugarbaker TA, Chang D, Koslowe P S P. Pathobiology of peritoneal carcinomatosis from ovarian malignancy. Cancer Treat Res 1996; (81): 63-74.

Tozzi R, Köhler C, Ferrara A, Schneider A. Laparoscopic treatment of early ovarian cancer: surgical and survival outcomes. Gynecol Oncol 2004; (93): 199-203.

Trimbos J B, Parmar M, Vergote I, Guthrie D, Bolis G, Colombo N, Vermorken J B, Torri V, Mangioni C, Pecorelli S, Lissoni A, Swart A M. International Collaborative Ovarian Neoplasm trial 1 and Adjuvant ChemoTherapy In Ovarian Neoplasm trial: two parallel randomized phase III trials of adjuvant chemotherapy in patients with early-stage ovarian carcinoma. J Natl Cancer Inst 2003; (95): 105-112.

Trimbos J B, van der Burg M E. [Adjuvant chemotherapy in patients operated on for early ovarian carcinoma]. Ned Tijdschr Geneeskd 2004; (148): 874-878.

Vergote I, De B J, Fyles A, Bertelsen K, Einhorn N, Sevelda P, Gore M E, Kaern J, Verrelst H, Sjovall K, Timmerman D, Vandewalle J, Van G M, Trope C G. Prognostic importance of degree of differentiation and cyst rupture in stage I invasive epithelial ovarian carcinoma. Lancet 2001; (357): 176-182.

Vergote I, Trimbos JB. Treatment of patients with early epithelial ovarian cancer. Curr Opin Oncol 2003; (15): 452-55.

Vergote I. Role of surgery in ovarian cancer: an update. Acta Chir Belg 2004; (104): 246-256.

Vergote I B, Amant F. Early ovarian cancer-time for a rethink on stage? Gynecol Oncol 2004; (94): 607-608.

Vinatier D, Dufour P, Cosson M, Querleu D. Laparoscopy in gynaecological cancer. Surg Oncol 1996; (5): 211-220.

Volz J, Köster S. Laparoscopy:to inflate or lift. Cancer 1999; (86): 749-50.

Volz J, Koster S, Spacek Z, Paweletz N. The influence of pneumoperitoneum used in laparoscopic surgery on an intraabdominal tumor growth. Cancer 1999; (86): 770-774.

von E E, Altman D G, Egger M, Pocock S J, Gotzsche P C, Vandenbroucke J P. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet 2007; (370): 1453-1457.

Wells B, Shea B, O'coneell, Peterson J, Welch V, Losos M, Tugwell P. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analysis. Internet. 2007. Ref Type: Electronic Citation

Wenzl R, Lehner R, Husslein P, Sevelda P. Laparoscopic surgery in cases of ovarian malignancies: an Austria-wide survey. Gynecol Oncol 1996; (63): 57-61.

Yancik R. Ovarian cancer. Age contrasts in incidence, histology, disease stage at diagnosis, and mortality. Cancer 1993; (71): 517-523.

Table 1. Participants characteristics and scoring of Jadad 1996

Study	Methods	Participants	Interventions	Outcomes	Quality
Ghezzi *et al.,* 2007	Case-control study. They did not write about sample size calculation and power. The study was carried in 2003, in Departament of Obstetrics and Gynecology, University of Insubria, Del Ponto Hospital, Piazza Biroldim I, Vareses, Italy and in Departament of Obstetrics and Gynecology, University of Verona, Italy.	34 patients with apparent early ovarian cancer was submitted a surgery. 15 patients undergoing a comprehensive laparoscopic stating and were compared with 19 patients that were submitted a laparotomy. All women were submitted a preliminary workup, including ultrasonography, CA 125, as well as colour Doppler ultrasonography. Age (years) in the laparoscopy group: 55 years (SD 13.5; range 13-70). Age (years) in the laparotomy group: 61 (SD 0.58; range 44-70), the body mass index 23.8 (SD 4.2) in the laparoscopy group and 25.8 (SD 3.1) in the laparotomy group.	Laparoscopic and laparotomy for treatment early ovarian cancer. After the diagnosis of malignancy, multiple random peritoneal were performed. Bilateral lymphadenectomy was performed as previously describe and all patients, external iliac, internal iliac and obturator lymphonodes were removed. Common iliac and paraaortic lymphadenectomy were performed. Total infracolic omentectomy was than performed using scissors and bipolar coagulation. Appendicectomy were performed. Salpingo-oophorectomy and total laparoscopic hysterectomy were performed.	Postoperative complications were defined as adverse events occuring within 30 days of surgery as a result of the procedure. Febrile morbidity was defined two temperatures > 38, hospital stay, blood loss, blood transfusions, pelvic lymph nodes, paraaortic lymphnodes.	C
Study	**Methods**	**Participants**	**Interventions**	**Outcomes**	**Quality**
Hua *et al.,* 2005	Case-control study. The study was carried out between september 2002 to may 2004 in Departament of Gynecology, Fudan University, Gynecology and Obstetric hospital, Shangai 200011, China.	21 patients with early ovarian cancer were evaluated. 10 were submitted the laparoscopic operation and 11 were submitted the laparotomy. Mean age in laparoscopy group was 40 (SD8). Mean age in laparotomy group was 42 (SD6)	Laparoscopic in 10 patients with early ovarian cancer who underwent laparoscopic total hysterectomy, pelvic lymph nodes dissection, bilateral adnexetomy, ovarian aortic and vein high ligation, omentectomy, and additional appendicectomy Laparatomy in 11 patients with early ovarian cancer. who underwent the same procedure. Frozen section method during operation proved the diagnosis of ovarian cancer and cytological examination proved negative result of the peritoneal irrigation liquid.	Operative time, intraopetrative blood loss, number of pelvic lymph node resected, surgical complications	C

Study	Method	Participants	Interventions	Outcomes	Quality
Tozzi *et al.*, 2004	Cohort prospective. They did not write about sample size calculation and powe and precision. The study was carried out between May 1996 until June 2003 in Departament of Gynecology, Friedrich Shiller, Jena, Germany	Forty two patients were elegible to enter the study, but 18 patients were excluded because of tumor rupture (n = 5), presense of peritoneal tumor implants (n = 3), ovarian surface invasion either macroscopic (n = 4), or microscoic at frozen (n = 2) or because of tumor size (n = 4), exceeding 11 x 8 cm, which is the diameter of biggest endobag. All these condicions mantaded conversion to lapatotomy. Thus, 24 patients with FIGO stage IA-B underwent either primary treatment or completation of staging by laparoscopy. All women were submitted a preliminary workup, including ultrasonography, CA 125, as well as colour Doppler ultrasonography. Age (years) in the laparoscopy group: 36.8 (SD 13.5; range 19-76) and body mass index 27.3 (SD 4.37; 20.2-38.6).	24 patients with ovarian FIGO stage IA-B were managed by laparoscopy. All patients underwent bowel preparation. The procedure was started by laparoscopy with peritoneal washing and careful inspection of the entire adbomen including diaphragm, liver, gallbaldder, small bowel, reto-sigmoide colon, paracolic gutter and abdominal wall. Any suspicious lesion was biopsied abd sent for frozen section. All specimens were retrieved via endobag to avoid contant with port site.. Once the diagnosis of cancer was confirmed, intraperitoneal spread was excluded by laparoscopy and tumor rupture avoided and laparoscopic assisted vaginal hysterectomy (LAVH) with contralateral salpingo-ophorectomy, appendectomy, particial resection of the omentum, pelvic lymphadenectomy and infrarenal para aortic bilateral lymphadenectomy. Also appendectomy was performed.	Operative complications, surgery time, mean number of pelvic lymph nodes, hospital stay, survival and progression-free survival at five years.	C

Table 2. Newcastle-Ottawa quality of cases-control studies

Numbered item	Hua *et al.*, 2005	Guezzi *et al.*, 2007
SELECTION		
1) Is the case definition adequate?		
(a) Yes, with independent validation?*	Yes*	Yes*
(b) Yes, e.g., record linkage or based on self reports		
(c) No description		
2) Representativeness of the cases		
(a) Consecutive or obviously representative series of cases*	Yes*	Yes*
(b) Potential for selection biases or not stated		
3) Selection of Controls		
(a) Community controls *		
(b) Hospital controls	Yes	Yes
(c) No description		
4) Definition of controls		
(a) No history of disease (endpoint)*	Yes*	Yes*
(b) No description of source		
COMPARABILITY		
1) Comparability of cases-control on the basis of design or analysis		
(a) Study controls for _____selected the most important factor.*	Yes*	Yes*
(b) Study controls for any additional factor * (This criteria could be modified to indicate specific control for a second important factor)		
EXPOSURE	Hua *et al.*, 2005	Guezzi *et al.*,2007
1) Ascertaiment of exposure		
(a) Secure record (e.g., surgical records) *	Yes*	Yes*
(b) Strutured interview where blind to case/control status*		
(c) Interview not blinded to case/controls status		
(d) Written self report or medical record only		
(e) No description		
2) Same method os ascertaiment for cases and controls		
(a) Yes *		
(b) No	No	No
3) Non-response rate		
(a) Same rate for both groups*	No	No
(b) No respondents describe		
(c) Rate different and no designation	Yes	Yes

Table 3. Newcastle-Ottawa quality of cohort studies

Numbered item	Ghezzi *et al.*, 2004
SELECTION	
1) Representativeness of the exposed cohort	
(a) Truly representative of the average ____(describe) in the commnity.*	Yes*
(b) Somewhat representative of average ___ in the community*	
(c) Select group of users e.g., nurses, volunteers	
(d) No description of the derivation of the cohort	
2) Selection of the non exposed cohort	
(a)Drawn from the same community as the exposed cohort *	Yes*
(b) Drawn from a different source	
(c) No description of the derivation of the non exposed cohort	
3) Ascertainment of exposure	
(a) Secure record (e.g., surgical records) *	Yes*
(b) Strutured interview*	
(c) Writen self report	
(d) No description	
4) Demonstration that outcome of interest was not present as start of study	
(a) Yes *	Yes*
(b) No	
COMPARABILITY	
1) Comparability of cohorts on the basis of design or analysis	
(a) Study controls for _____(selected the most important factor)*	Yes*
(b) Study controls for any additional factor * (This criteria could be modified to indicate specific control for a second important factor)*	Yes*
OUTCOMES	
1) Assessment of outcome	
(a) Independently blind assessment *	Yes*
(b) Record lincage *	
(c) self report	
(d) no description	
2) Was follow up enough for outcomes to occur	
(a) Yes (selected adequate follow up period for outcome of interest)*	Yes*
(b) No	
3) Adequacy of follow up of cohorts	
(a) Complete follow up - all subjects accounted for *	Yes*
(b) Subjects lost to follow up unlikely to introduce bias - small number lost >_____% (select an adequate%) follow up, or decription provide of those lost *	
(c) Follow up rate < % (selected an adequante %) and no description of those lost	
(d) No statement	

Table 4. Strengthening the Reporting of Observational studies in Epidemiology (STROBE)

Item	Item number	Recomendations	Hua *et al.*, 2005	Guezzi *et al.*, 2007	Tozzi *et al.*, 2004
Title and Abstract	1	(a) indicate the study's design with a commonly used term in the title or the abstract.	No	No	No
		(b) provide in the abstract and informative and balaced summary of what was done and what was found.	Yes	Yes	Yes
Introduction/ Background	2	Explain the scientific background and rationale for the investigations being reported.	Yes	Yes	Yes
Objetives	3	State specific objetives, including any prespecified hypotheses.	Yes	Yes	Yes
Methods					
Study design	4	Present key elements of study design early in paper.	Yes	Yes	Yes
Settings	5	Describe the settings, locations and relevant dates, including periods of recruitment, exposure collection.	Yes	Yes	Yes
Participants	6	(a)Cohort study: give eligibility criteria, and the sources and methods of selection or participants methods of follow up	–	–	Yes
		(b) Case control study: give the eligibility criteria, and the sources and methods of case ascertaiment and control selection.	Yes	Yes	–
Variables	7	Clearly define all outcomes, exposures, predictors, potencial confunders, and effect modifiers.	No	No	No
Data sources /measuments	8	For each variables of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group. Give such information separately for cases and controls studies, and, if applicable, for exposed and unexposed groups in cohort and cross sectional studies.	Yes	Yes	Yes
Bias	9	Describe any efforts to address potential sources bias.	No	No	No
Study Size	10	Explain how the study size was arrived at.	No	No	No
Quantitaive variables	11	Explain how quantitative variables were handled in the analysis. If applicable, describe which groupings were chosen and why.	Yes	Yes	Yes
Statistical Methods	12	(a) Describe all statistical methods, including those used to control for confunding.	No	No	No
		(b) Describe any methods used to examine subgroups and interations.	No	No	No
		(c) Explain how missing data were adderessed.	No	No	No
		(d) Cohort study: if applicable, explain how loss to follow up was addressed	–	–	No
		Case-control study, if applicable, explain how matching of cases and controls was adderessed	No	No	–
		e) Describe any sensibility analysis	No	No	No

Item	Item number	Recomendations	Hua *et al.*, 2005	Guezzi *et al.*, 2007	Tozzi *et al.*, 2004
Results /Participants	13*	a)report the number of individuals at each stage of the study- e.g., numbers potencially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow up, and analysis.	Yes	Yes	Yes
		b) give reasons for nonparticipation at each stage	Yes	Yes	Yes
		c) consider use of a flow diagram	No	No	No
		*give such information separately for cases and controls in case control studies, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.	Yes	Yes	–
Results/ Descripitive data	14*	a)give characteristics of study participants (e.g., demographic, clinical, social) and information on exposures and potential confounders/	Yes	Yes	Yes
		(b) indicate the number of participants with missing data for each variable of interest.	No	No	No
		c) Cohort study: summarize follow up time - e.g., average and total amount	–	–	Yes
		*give such information separately for cases and controls in case control studies, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.	No	No	No
Results/Outcome data	15*	(a) Cohort studies: reported numbers of outcome events or summary measures of exposure.	–	–	Yes
		(b) Case- control study: reported numbers in each exposure category or summary measures of exposure	Yes	Yes	–
		*give such information separadely for cases and controls in case control studies, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.			
Results/ Main Results	16	(a) give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g., 95% confidence intervals). Make clear which confounders were adjusted for and why they were included	No	No	No
		(b) report categories boundaries when continuous were categorized	–	–	–
		(c) If relevant, consider translating estimates or relative risk into absolute risk for a meaningful time period.	No	No	No
Results/ Other Analysis	17	Report other analysis done, e.g., analysis of subgroups and interactions and sensibility analysis	No	No	No
Discussion					
Key results	18	Summarize key results with reference to study objectives	Yes	Yes	Yes
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or impression. Discuss both directions and magnitude of any potential bias.	Yes	Yes	Yes
Interpretation	20	Give a cautions overall interpretation or results considering objectives, limitations, multiplicity of analysis, results from similar studies, and other relevant evidence.	Yes	Yes	Yes
Generalization	21	Discuss the generalizability (external validity) of the study results	Yes	Yes	Yes
Other infomation					
Funding	22	Give the source of funding and the role of funders for the present study and, if applicable, for the original study on which	Yes	Yes	Yes

Table 5. Comparison between studies of cases-control

Outcomes	n	Laparoscopy Hua *et al.*, 2005	n	Laparotomy Hua *et al.*, 2005	p	n	Laparoscopy Ghezzi *et al.*, 2007	n	Laparotomy Ghezzi *et al.*, 2007	p
Operative time	10	298 min (SD 60 min)	11	182 min (SD 43 min)	< 0.05	15	377 min (SD 47 min)	19	272 min (SD 81 min)	0.002
Blood loss (ml)	10	280 ml (SD 280 ml)	11	346 ml (SD 170 ml)	< 0.05	15	250 ml (SD 225; range 50-1000 ml)	19	400 ml (SD 201; range150-1000 ml)	0.28
Number pelvic lymph nodes	10	25 (SD 5)	11	27 (SD 7)	> 0,05	15	25.2 (SD 9.3)	19	25.1 (SD 5.8)	> 0.05
Pos-operative complications	10	2 (20%)	11	7 (72.7%)	0,05	15	2 (13.3%)	19	8 (42.1%)	0.13

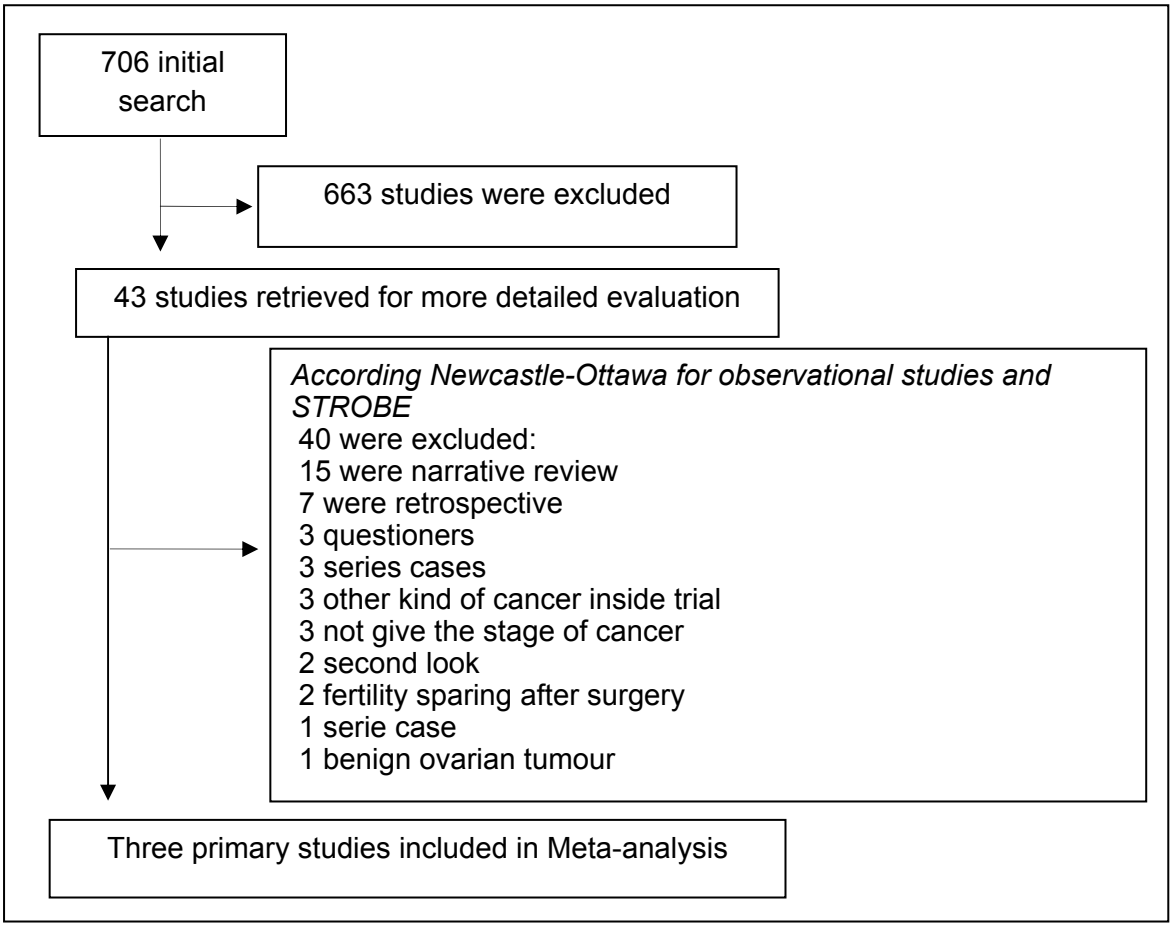


Figure 1. Study selection process.

5.2 Artigo 2

ACCURACY OF FROZEN SECTION ANALYSIS IN THE DIAGNOSIS OF OVARIAN TUMORS: A SYSTEMATIC QUANTITATIVE REVIEW

Publicado na revista Int J Gynecol Cancer 2005,15, 192-202

Accuracy of Frozen Section Analysis in the Diagnosis of Ovarian Tumors: A Systematic Quantitative Review

Author: Lidia Rosi Medeiros, MD, MSc[(1)]

Address: José de Alencar 1244 apt 1009, Porto Alegre, RS, Brasil, CEP 90880-480
lidia.rosi@terra.com.br

Co-authors:

Daniela Dornelles Rosa, MD[(2)]
Maria Isabel Edelweiss, MD, PhD[(2, 5)]
Airton Tetelbon Stein, MD, PhD[(2, 3, 6)]
Mary Clarisse Bozzetti, MD, PhD[(1, 2, 4)]
Alice Zelmanowicz, MD, PhD[(7)]
Paula R.Pohlmann, MD, PhD[(1)]
Luise Meurer, MD, PhD[(5)]
Mariana Teixeira Carballo[(1)]

Affiliations of all authors:

[1] Postgraduate Program in Epidemiology at Federal University of Rio Grande do Sul, Porto Alegre Brazil
[2] Postgraduate Program in Medicine: Medical Sciences at Federal University of Rio Grande do Sul, Porto Alegre, Brazil
[3] Public Health Postgraduate Course at Universidade Luterana do Brasil
[4] Department of Social Medicine, Faculty of Medicine at Federal University of Rio Grande do Sul, Porto Alegre Brazil
[5] Pathology Unit of Hospital de Clinicas de Porto Alegre, Porto Alegre, Brazil
[6] Postgraduate Program in Medical Science at Porto Alegre Federal Faculty of Medical Science, Porto Alegre, Brazil
[7] Cancer Prevention Center, Complexo Hospitalar Santa Casa, Porto Alegre, Brazil

RESUMO

Foi realizada uma revisão sistemática quantitativa para estimar a acurácia diagnóstica do exame anatomopatológico de congelação. Foram incluídos estudos que comparam o resultado desse exame com o resultado final da histologia no exame de parafina nesta revisão sistemática. Os 14 estudos analisados compreenderam 3.659 mulheres. Lesões ovarianas benignas *vs* lesões limítrofes ou malignas de ovário possuem, no somatório final de todos os estudos inclusos, uma razão de verossimilhança positiva de 8,7 (IC 95% de 7,3-10,4) e probabilidade pós-teste para lesões benignas de 95% (IC 95% de 94%-96%). Exame anatomopatológico de congelação comparando resultados de malignidade *vs* lesões benignas possui uma razão de verossimilhança positiva, no somatório geral de todos estudos inclusos, de 303 (IC 95% de 101-605) com aumento da probabilidade pós-teste para malignidade de 98% (IC 95% de 97%-99%). Já na comparação de tumores com malignidade limítrofe (*borderline*) *vs.* lesões benignas ovarianas foi encontrado no somatório geral de todos os estudos, uma razão de verossimilhança de 69 (IC 95% de 45-106) com aumento de probabilidade pós-teste para tumores com malignidade limítrofe de 79% (IC 95% de 71%-85%). Ao se comparar o resultado do somatório de todos os estudos com tumores com malignidade limítrofe *vs* tumores malignos, tem-se um valor de verossimilhança bem menor em relação aos anteriores, sendo de 18 (IC95% de 13-26) tendo uma probabilidade pós-teste somente de 51% (IC 95% de 42%-60%). Conclui-se, assim, que a acurácia do exame de anatomopatológico de congelação é elevada no diagnóstico de tumores malignos e benignos, sendo baixa, porém, para as lesões com malignidade limítrofe.

Palavras-chave: diagnóstico por congelação, revisão sistemática, metanálise, acurácia diagnóstica.

ABSTRACT

A quantitative systematic review was performed to estimate the diagnostic accuracy of frozen section in ovarian tumors. Studies that compared frozen section and paraffin sections within subjects for diagnosis of ovarian tumors were included. Fourteen primary studies were analyzed, which included 3,659 women. For benign ovarian *vs.* borderline/or malignant tumor cases, occurrence of a positive frozen section result for benignity (pooled LR, 8.7; 95% CI, 7.3-10.4) and post-test probability for benignity diagnosis was 95% (95% CI, 94%-96%). A positive frozen section result for malignity *vs.* benign (pooled LR, 303; 95% CI, 101-605) increased the probability of ovarian cancer to 98% (95% CI, 97%-99%). In borderline *vs.* benign ovarian tumor cases, a positive frozen section result (pooled LR, 69; 95% CI, 45-106) increased the probability of borderline tumors to 79% (95% CI, 71%-85%). In borderline *vs.* malignant ovarian tumor cases, a positive frozen section result (pooled LR, 18; 95% CI, 13-26) increased the probability of borderline tumors to 51% (95% CI, 42%-60%). We conclude that diagnostic accuracy rates for frozen section analysis is high for malignant and benign ovarian tumors, but the accuracy rates in borderline tumors remain relatively low.

Keywords: frozen section diagnosis, systematic review, meta-analysis, accuracy, diagnosis

INTRODUCTION

Malignant ovarian neoplasms are responsible for 4% of all cancers affecting women and are the second most common cause of death from gynecological cancer and the fourth most common cause of death from all types of cancer affecting women[1]. Benign, borderline and malignant lesions have been identified within the same surgical specimen[2]. However, the frequency and speed of the evolution from dysplasia into cancer remain unknown[3].

The use of frozen sections has had a great impact on the care of gynecologic patients and has become indispensable in diagnosing malignancy. It also helps to determine the staging, as well as the appropriate surgical technique for each case. In ovarian cancer, surgery usually involves total abdominal hysterectomy, bilateral salpingo-oophorectomy, omentectomy, pelvic and para-aortic lymphadenectomy[4]. In borderline tumors, a fertility sparing approach can be used particularly in patients who desire fertility preservation[5]. The accuracy of frozen section diagnosis of benign and malignant ovarian tumors is generally quite good. However, the diagnosis of borderline ovarian tumors has often been more difficult, particularly in frozen section examination[6]. Therefore, we undertook a quantitative systematic review to ascertain the accuracy of frozen section diagnoses of benign, borderline and malignant ovarian tumors and to explore the reasons for the ongoing debate about this issue.

METHODS

Identification of studies

A comprehensive search of the MEDLINE, CANCERLIT, LILACS and EMBASE databases was made from January 1984 to December 2003. The medical subjects heading (MeSH) and text words for the terms "*ovarian neoplasm"* and "*frozen section"* were combined with the MeSH term *diagnosis* ("sensitivity and specificity"). The search was limited to human studies but had no language restrictions. In addition, the Cochrane Library was searched. Reference lists of all available primary studies were reviewed to identify additional relevant citations. The authors who published the studies were not contacted.

Selection criteria

This review focused on observational studies in which the results of the diagnostic test of interest were compared with the results of a reference standard. The case studies were women treated surgically for ovarian tumors. The diagnostic test was analysis of frozen tissue sections and the diagnostic reference was the result of later histological analysis of standard paraffin sections. A frozen section diagnosis was considered correct if it did not differ from that of the paraffin section. For inclusion criteria, in each trial it was necessary that the final paraffin section histological assessment designated each case as benign, borderline or malignant. We excluded the studies that compared accuracy of frozen with paraffin sections that described only borderline ovarian tumors or malignant ovarian tumor, and other kind of tumor (not ovarian) and, also the studies that lacked data to construct 2 x 2 contingency tables.

The final diagnoses of ovarian lesions were compared in four ways: (1) agreement as to benign *versus* borderline and/or malignant, (2) agreement as to malignant *versus* benign, (3) agreement as to borderline *versus* benign and (4) agreement as to borderline *versus* malignant. To calculate the diagnostic accuracy were excluded cases that were deferred because of uncertain frozen diagnoses in each study. Thus, the primary outcome measure was the accuracy of ovarian tumor diagnoses of benign, borderline and malignant from frozen section analysis. A secondary outcome was the distribution of histological types of ovarian tumors according to paraffin diagnosis. The reviewed studies were identified independently by four investigators (L.R.M, D.D.R, M.I.E and M.C.B). Final inclusion or exclusion was made with reference to a selection criteria checklist. Disagreements about study inclusion or exclusion were initially solved by consensus, and when this was not possible, they were resolved by arbitration with a fifth reviewer (A.T.S). The agreement statistics among reviewers were computed.

Quality assessment

All articles meeting the eligibility criteria were assessed for their methodological quality. This assessment involved scrutinizing the study designs and the relevant features of patient population, the diagnostic test and the reference standard[7-10]. These features included the methods of data collection and patient selection, description of frozen sections and the histological reference standard, and

presence of verification bias[8, 9].The quality assessment results for the included studies are summarized in the "*Scoring of Study Quality*" column of table 1[7]. Studies were further assessed for methodological quality with reference to the *Oxford Centre for Evidence Based Medicine Level of Evidences Classification* rubric*.* Only studies with Oxford Evidence Levels 1 to 3 were considered to be of high quality while those with levels 4 and 5 were excluded[11].

Data abstraction

Four investigators (L.R.M, D.D.R, M.I.E and M.C.B) independently abstracted data regarding the prevalence of benign, borderline and malignant ovarian lesions and the sensitivities, specificities, likelihood ratios and post-test probabilities from the primary studies of frozen section diagnosis. The assessment of English-language articles was performed by 2 reviewers (L.R.M, D.D.R) while assessment of those articles published in languages other than English were performed independently by 2 other reviewers (M.I.E, M.C.B) following translation (when necessary). Any disagreement was resolved by consensus for English and non-English studies. Three ovarian tumor diagnosis outcomes were considered: benign, borderline and malignant. Benign lesion diagnosis was the primary outcome followed by lower incidences of malignant and borderline lesions. Data for benign tumor cases were abstracted as 2 x 2 tables of frozen section diagnosis (positive *versus* negative for benign lesions) and paraffin section diagnosis (benign *versus* malignant or borderline). Similarly, contingency tables were produced for borderline tumor cases comparing frozen section diagnosis (positive for borderline *versus* benign or malignant lesions) and paraffin section diagnosis (borderline *versus* benign or malignant) as well as for malignant ovarian tumor cases comparing frozen section diagnosis (malignant and benign) and paraffin section diagnosis (malignant *versus* benign) (Table 2).

Data synthesis and statistical analysis

To evaluate agreement between study eligibility and methodological quality assessments as well as agreement between frozen and paraffin section analyses the observed percentage agreement and κ coefficient for inter-rater reliability were calculated[12]. For each study, we constructed 2 x 2 contingency tables in which all biopsies were classified as benign, borderline or malignant lesions. We calculated

the true-positive rate (TPR; sensitivity), false-positive rate (FPR; 1 – specificity), likelihood ratios (LRs) and post-test probability for each study along with 95% confidence intervals (CIs). When 2 x 2 tables contained 0 cells, 0.5 was added to each cell to enable our calculations to be made.

Meta-analysis to produce summary pooled estimates of sensitivity and specificity were performed if these measures were found to be independent. The association between sensitivity and specificity was calculated with *Spearman's correlation coefficient* for benign, borderline and malignant ovarian tumors[13]. In this case, estimates of sensitivity and specificity can be calculated because there are two categories of results (negative or positive test outcome) and there is not variability of the diagnostic threshold[14]. LRs can be estimated either from the summary estimates of sensitivity and specificity by using LR positive = sensitivity /1 – specificity and LR negative= 1 – sensitivity /specificity[15]. LRs indicate how much a given frozen section finding increases or decreases the probability of a final diagnosis of a lesion as benign, borderline or malignant[15]. In addition, the pooled post-test probability estimates were calculated by multiplying the pre-test probability (prevalence) by the pooled LR[14]. The heterogeneity of the sensitivities and specificities from the studies was assessed using the Q_T (Cochran) test for χ^2 distributions with N-1 degrees of freedom. Because sensitivity and specificity were homogeneous, a fixed effect model was used and these terms were pooled with a 95% CI[14, 16].

Sensitivity analysis

To assess whether study quality affected diagnostic accuracy of frozen section analysis we excluded those studies that met less than 50% of the criteria quality and that were sub-level 3 by *Oxford Centre for Evidence Based Medicine Level of Evidences Classification*[11, 16] .

To analyze publication bias, inverted funnel plots of individual study log *Odds ratio* (OR) were plotted against sample size[16; 17]. The robustness of the results was tested by repeating the analysis with a different statistical model (random effects model)[16, 17].

RESULTS

Study identification and eligibility

The process of study selection is summarized in figure 1. Our initial search identified 582 potentially relevant articles. We excluded 557 published studies after reviewing their titles and abstracts, because the four independent reviewers considered that they did not relate to the question under review. Twenty-five full-text articles were retrieved including six non-English-language studies; 11 were excluded after further scrutiny. A complete list of excluded studies is available from the authors. Fourteen primary studies (three non-English language studies), including 3,659 women, met the criteria for inclusion and were analyzed (Table 1)[18-31]. Interrater agreement for study eligibility and methodological quality was 79% (κ = 0.64), indicating good agreement[12]. Disagreement between reviewers occurred during analysis of the 25 studies and they were related to inclusion or exclusion criteria (characteristics and scoring of study quality), but it was solved by consensus.

Study description

Details of the participants, interventions, and quality assessments of the studies selected for meta-analysis are summarized in table 1. The mean age of participants across studies was 49 (range 1-95). All studies were non-blind and retrospective from a narrow population, but included sufficient experimental details and the proper diagnostic tests and diagnostic reference standards. Of the 14 included studies, there were five studies[18, 19, 22, 24, 31] with high methodological quality, satisfying ≥ 55% of the criteria for study quality and with an Oxford Evidence Level of 2B[11]. However, nine trials[20, 21, 23, 25-30] were classified as level 3B, because details of the patient populations including age were not reported and because they met < 50% of the criteria.

Benign ovarian tumors were found in 2,593 patients (71%), borderline tumors in 201 patients (5.5%) and malignant tumors in 832 patients (22.9%) (Table 2). Interrater agreement between frozen and paraffin sections was 94% (κ=0.86), indicating very good agreement (Table 2)[12]. Table 2 showed the results of contingency tables (FN, FP, TP and TN) from each study considered. The difference on total numbers on tables 1 and 2 are due to exclusion of 48 misdiagnosed cases by frozen section and of 33 deferred cases, in which frozen section diagnosis was

uncertain. From the 48 misdiagnosed cases, 9 malignant cases by frozen section examination had a final diagnosis of borderline tumors. The other 39 cases had previous result of borderline tumors that was changed to malignant by paraffin examination.

There were no significant correlations between sensitivity and specificity for benign *vs.* malignant or borderline (r_s = 0.42; P = 0.14), borderline *vs.* benign (r_s = 0.1; P = 0.76), malignant *vs.* benign (r_s = 0.36; P = 0.21) and borderline *vs.* malignant (r_s = 0.1; P = 0.7) ovarian tumors. Because sensitivity and specificity did not correlate, a summary receiver operating characteristic curve was not generated[14, 15].

The comparison between the final diagnosis and frozen section was set in four ways (benign *vs.* borderline/malignant, malignant *vs.* benign, borderline *vs.* benign and borderline *vs.* malignant). Sensitivity, specificity, LR and post-test probability data for benign, borderline and malignant lesions are summarized in Tables 3, 4, 5, 6. For malignant *vs.* benign ovarian tumors, the pre-test probability of cancer increased from 17.5% overall to 98% (95% CI, 97%-99%) with a positive result for malignancy in the frozen section diagnosis and decreased to 1.6% (95% CI, 1.1%-1.9%) with a negative result (Table 4). For borderline *vs.* benign ovarian tumors, when frozen section diagnosis was positive for borderline ovarian tumors the pre-test probability increased from 5.5% overall to 79% (95% CI, 78%-82%) and if the result was negative, it decreased to 1.9% (95% CI, 1.6%-2.3%) (Table 5). For borderline *vs.* malignant ovarian tumors, if the frozen section diagnosis was positive for borderline ovarian tumors, the pre-test probability increased from 5.5% overall to 51% (95% CI, 42%-60%) and if the result was negative, it decreased to 0.5% (95% CI, 0.2%-0.9%) (Table 6).

The distribution of tumor histological types according to paraffin diagnosis was calculated from seven of the trials that described 2,281 ovarian tumors[19, 20-22, 26, 28, 31] (Table 7). In the final diagnosis epithelial tumors constituted 1,326 (58%) of these cases, germ cell tumors 41 (17%), tumor-like lesions 318 (14%), sex cord-stromal tumors 141 (6%), secondary metastatic lesions 39 (1.7%) and tumors with uncertain origin five (0.2%) of these cases. Of the malignant ovarian tumors, 75.7% of the cases were epithelial tumors, 8.9% were secondary metastatic tumors and 8.4% were germ cell tumors.

Sensitivity analysis

The robustness of the results was tested by repeating the analysis using a different statistical model (random effects model)[16, 17]. There was homogeneity among the fourteen studies for benign, borderline and malignant ovarian tumors. Heterogeneity testing showed that application of the fixed or the random model did not change the results: benign *vs.* malignant or borderline (Q_T = 12.10, p = 0.52), borderline *vs.* benign (Q_T = 7.14, p = 0.79), malignant *vs.* benign (Q_T = 3.5, p = 0.52) and borderline *vs.* malignant (Q_T = 8.8, p = 0.64) ovarian tumor. The pooling of sensitivity, specificity and LR data from the five studies with high methodological quality did not alter the accuracy rate for diagnosis by frozen section analysis in benign and malignant ovarian tumors[18, 19, 22, 24, 31]. For borderline tumors *vs.* benign, however, the pooled sensitivity changed to 57% (95% CI, 47-67)[18, 19, 22, 24, 31], with homogeneity between studies (Q_T = 0.96 , p = 0.92) by fixed and random effects models. For borderline tumors *vs.* malignant, the pooled sensitivity changed to 89% (95% CI, 79-95)[18, 19, 22, 24, 31], with homogeneity between studies (Q_T = 3.9 , p = 0.42) by fixed and random effects models. The post-test probability with frozen section diagnosis of a borderline tumor though was only slightly changed by selective inclusion of the five higher evidence level studies to 77% (95% CI 75-79) in borderline *vs.* benign and to 57% (95% CI 44-69) for borderline *vs.* malignant ovarian tumor. Therefore, all 14 selected[18-31] studies were included in the sensitivity analysis. Inverted funnel plots showed asymmetry.

COMMENT

The reliability of frozen section diagnosis of ovarian tumors is a critical determinant in selection of the appropriate surgical procedure and provides important information about patient prognosis[32]. Correct intra-operative histopathologic diagnosis is critical to assessing an appropriate surgical plan and for prevention of under – and overtreatment of patients[20].

If ovarian tumors can be accurately diagnosed in the operating room surgeons can give patients the option of fertility preserving surgery when possible[20]. This meta-analysis determined that frozen section diagnosis is safe for benign and malignant ovarian tumors, but has low sensitivity and post-test probability agreement for borderline tumors. The incidence of borderline ovarian tumors however is relatively low when compared to that of benign and malignant ovarian tumors[33].

Among the broad sample of ovarian tumors in this systematic review, 71% were benign, 5.9% were borderline and 22.7% were malignant. These data are in accord with the literature[33].

The pooled sensitivity rates for benign and malignant ovarian tumors was 99% and 94%, which represents the probability that frozen section diagnosis results will be positive for benignity and malignancy in patients with benign or malignant diseases, respectively. There is no definitive variation in the sensitivity of frozen section diagnoses of malignant or benign ovarian tumors. The group of studies reviewed here included 8 cases (0.23%) falsely diagnosed as positive for cancer and 45 cases (1.3%) falsely diagnosed as negative for cancer. Thus in the cases of a false positive result for malignancy the surgeons were misled into doing an extensive unwarranted surgical procedure. Generally though, agreement between frozen section and subsequent paraffin section diagnoses has been good as positive frozen section results for malignancy increased the probability of ovarian cancer to 98% (95% CI, 97%-99%), whereas a negative result reduced the probability of cancer to 1.6% (95% CI, 1%-2%).

The pooled sensitivity for borderline and benign ovarian tumors however was low (66%; 95% CI, 59%-72%). This relatively low sensitivity was due to the greater incidence of false negative results (2.5%) relative to false negative rates for benign (0.8%) and malignant (1.3%) cases. Insufficient tumor removal, technical problems in sampling and the range of experience of the pathologists may also have contributed to the reduced sensitivity among borderline cases[19]. The differentiation between borderline and benign ovarian tumor is based on at least two of the following microscopic features: (1) detachment of cells clusters in cystic spalls of the primary tumor; (2) stratification or multilayering of the epithelium; (3) increased mitotic activity; and (4) cytonuclear atypia[34]. We have found mucinous tumors in approximately 40% of cases and sometimes inadequate sampling of frozen sections may occur due to the larger mucinous tumor dimensions, resulting in possible overlooked malignant areas[19] . Moreover, as a further complication a single mucinous ovarian tumor may sometimes contain benign, borderline, and malignant components, in contrast with serous tumors[24, 26].

The histopathologic criteria used for differential diagnosis between borderline and malignant lesion vary in different countries; the absence or presence of microinvasion growth is regarded as the sole criterion, which is another problem in

frozen section concerned with borderline *vs.* malignant ovarian tumor[34]. Moreover, mucinous tumors invasive growth could be difficult to distinguish from non invasive proliferation, severe nuclear atypia and multilayering of nuclei of more than three layers is used to classify a case as invasive carcinoma[34]. We found that the pooled likelihood for a positive result of frozen section diagnosis for borderline *vs.* malignant was 18 (95% CI, 13-26) and post-test probability of borderline tumors was only 51% (95% CI, 42%-60%). For borderline *vs.* benign tumors the pooled likelihood for a positive result of frozen section diagnosis was 68 (95% CI, 44-105), increasing the post-test probability of borderline tumors to 79% (95% CI, 71%-85%). We could conclude that frozen diagnosis is more accurate for discriminating between borderline and benign lesions then between borderline and malignant tumor. This result represents a lesser agreement than diagnoses with the post-test probabilities that were found for benign and malignant cases. This finding indicates that accuracy of frozen section diagnosis in borderline tumors is modest and therefore diagnoses cannot be made with a high level of certainty and further testing is indicated.

Thus, in cases were the result is borderline in the frozen section the surgeon should be alerted to the necessity of a standard approach for staging, because determination of the extent of disease in patients with malignant ovarian tumor is important to plan the appropriate surgical approach and to predict the prognosis. The surgical procedure comprehends: (1) visual and tactile exploration of the peritoneal cavity, and of the retroperitoneal structures; (2) infracolic omentectomy; (3) peritoneal washings; and (4) biopsy of suspicious areas[5].

A possible limitation of this systematic review is a potential bias in that all trials were retrospective and that there was a lack of blinding in their assessment. In the sensitivity analysis, it was found inverted funnel plots showing asymmetry consistent with a publication bias. The probable reasons for this bias are the small number of trials included and the lower quality of some of the trials (nonconsecutive, insufficient population details)[20, 21, 23, 25-30]. On the other hand, this meta-analysis complied with the criteria for performing a rigorous systematic review planned a priori[14, 15, 17]. These included the use of study quality assessment[7] and investigation of homogeneity with fixed and random models to test the robustness of the results[14, 15, 17].

Intraoperative analysis of frozen sections for diagnosis of ovarian tumors is of great value in preventing under – and overtreatment of patients, but surgeons and pathologists need to be aware of its limitations[21, 31]. The specificity of frozen section

diagnosis is, in experienced hands, high[24, 32]. Thus, we suggest that blinded prospective studies should be performed to investigate the diagnostic accuracy of frozen section analysis in ovarian tumors with trained gynecologic pathologists.

This quantitative review provides precise evaluation of the accuracy of using frozen sections in the diagnosis of benign, borderline and malignant ovarian tumors. The present results indicate that this method is highly accurate in the diagnosis of benign and malignant ovarian tumors, and is only moderately useful in the diagnosis of borderline ovarian tumors. Therefore, better methods to improve the accuracy rates of borderline ovarian tumor diagnosis by gynecological pathologists are needed.

Acknowledgments

The authors thank Ms. Ruth Buist who performed a comprehensive search in EMBASE.

Reference List

(1) Yancik R. Ovarian cancer. Age contrasts in incidence, histology, disease stage at diagnosis, and mortality. Cancer 1993 Jan 15;71(2 Suppl):517-23.

(2) Scully RE. Early de novo ovarian cancer and cancer developing in benign ovarial lesions. Int J Gynecol Cancer 1995;49(suppl):S9-S15.

(3) Scully RE. Influence of origin of ovarian cancer on efficacy of screening. Lancet 2000;355:1028-9.

(4) National Institutes of Health Consensus Development Conference Statement. Ovarian cancer: screening, treatment, and follow-up. Gynecol Oncol 1994 Dec;55(3 Pt 2):S4-14.

(5) Robinson WR, Curtin JP, Morrow CP. Operative staging and conservative surgery in the management of low malignant potential ovarian tumors. Int J Gynecol Cancer 1992 May;2(3):113-8.

(6) Spann CO, Kennedy JE, Musoke E. Intraoperative consultation of ovarian neoplasms. J Natl Med Assoc 1994 Feb;86(2):141-4.

(7) Lijmer JG, Mol BW, Heisterkamp S, Bonsel GJ, Prins MH, van der Meulen JH, et al. Empirical evidence of design-related bias in studies of diagnostic tests. JAMA 1999 Sep 15;282(11):1061-6.

(8) Jaeschke R, Guyatt G, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. A. Are the results of the study valid? Evidence-Based Medicine Working Group. JAMA 1994 Feb 2;271(5):389-91.

(9) Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD Initiative. Ann Intern Med 2003 Jan 7;138(1):40-4.

(10) Cochrane Methods Working Group on Systematic Reviews of Screening and Diagnostic Tests. Recommended methods. Cochrane Library 2004 January 25 [cited 2004 Jan 25];Available from: URL: http://www.cochrane.org/cochrane/sadtdoc1.htm

(11) Phillips B. Oxford Centre for evidence-based Medicine Level of evidence Grades of recommendations (may 2001). Oxford Center 2007 November 23 [cited 2007 Nov 23];Available from: URL: http://www.cebm.net/index.aspx?o=1025

(12) Altman DG. Some common problems in medical research. In: Altman DG, editor. Practical statistics for medical research.London: Chapman & Hall; 1999. p. 403-9.

(13) Altman DG. Relation between two continuous variables. In: Altman DG, editor. Practical statistics for medical research.London: Chapman & Hall; 1999. p. 277-99.

(14) Deeks JJ. Systematic reviews of evaluation of diagnostic and screening tests. In: Egger M SGADG, editor. Systematic Reviews in Health care: Meta-analysis in context. 2 ed. London: BMJ Publishing; 2001. p. 248-82.

(15) Irwig L, Tosteson AN, Gatsonis C, Lau J, Colditz G, Chalmers TC, et al. Guidelines for meta-analyses evaluating diagnostic tests. Ann Intern Med 1994 Apr 15;120(8):667-76.

(16) Sutton AJ. Randon effects methods for combining study estimates. In: Sutton AJ AKJDSTSF, editor. Methods for Meta-Analysis in Medical Research. 1 ed. Chichester: John Wiley; 2000. p. 73-86.

(17) Deekes JJ ADBM. Statistical methods for examining heterogeneity and combining results from several studies in meta-analysis. In: Egger M SGAD, editor. Systematic Reviews in Health care: Meta-analysis in context.London: BMJ Publishing; 2001. p. 285-312.

(18) Gol M, Baloglu A, Yigit S, Dogan M, Aydin C, Yensel U. Accuracy of frozen section diagnosis in ovarian tumors: Is there a change in the course of time? Int J Gynecol Cancer 2003 Sep;13(5):593-7.

(19) Yeo EL, Yu KM, Poddar NC, Hui PK, Tang LC. The accuracy of intraoperative frozen section in the diagnosis of ovarian tumors. J Obstet Gynaecol Res 1998 Jun;24(3):189-95.

(20) Hamed F, Badia J, Chuaqui R, Wild R, Barrena N, Oyarzun E, et al. [Role of frozen section biopsy in the diagnosis of adnexal neoplasms]. Rev Chil Obstet Ginecol 1993;58(5):361-4.

(21) Torres JP, Suso JP, Perea E, Tafur L, Agudelo M. Tumores ováricos: correlación entre los informes de estudios solicitados por congelacion y la histopatologia definitiva. Hospital Universitário del Valle 1994-1997. Rev Colomb Obstet Gynecol 1998;49:149-51.

(22) Wakahara F, Kikkawa F, Nawa A, Tamakoshi K, Ino K, Maeda O, et al. Diagnostic efficacy of tumor markers, sonography, and intraoperative frozen section for ovarian tumors. Gynecol Obstet Invest 2001;52(3):147-52.

(23) Slavutin L, Rotterdam HZ. Frozen section diagnosis of serous epithelial tumors of the ovary. Am J Diagnostic Gynecol Obstet 1979;1(1):89-94.

(24) Rose PG, Rubin RB, Nelson BE, Hunter RE, Reale FR. Accuracy of frozen-section (intraoperative consultation) diagnosis of ovarian tumors. Am J Obstet Gynecol 1994 Sep;171(3):823-6.

(25) Usubutun A, Altinok G, Kucukali T. The value of intraoperative consultation (frozen section) in the diagnosis of ovarian neoplasms. Acta Obstet Gynecol Scand 1998 Nov;77(10):1013-6.

(26) Pinto PB, Andrade LA, Derchain SF. Accuracy of intraoperative frozen section diagnosis of ovarian tumors. Gynecol Oncol 2001 May;81(2):230-2.

(27) Puls L, Heidtman E, Hunter JE, Crane M, Stafford J. The accuracy of frozen section by tumor weight for ovarian epithelial neoplasms. Gynecol Oncol 1997 Oct;67(1):16-9.

(28) Cuello MF, Galleguillos GL, Zárate CR. Biopsia rápida por congelación en el diagnóstico de tumores de ovario: correlación diagnóstica según diámetro y peso en tumores de origem epitelial. Rev Méd Chile 1999;127:1199-205.

(29) Obiakor I, Maiman M, Mittal K, Awobuluyi M, DiMaio T, Demopoulos R. The accuracy of frozen section in the diagnosis of ovarian neoplasms. Gynecol Oncol 1991 Oct;43(1):61-3.

(30) Lim FK, Yeoh CL, Chong SM, Arulkumaran S. Pre and intraoperative diagnosis of ovarian tumours: how accurate are we? Aust N Z J Obstet Gynaecol 1997 May;37(2):223-7.

(31) Twaalfhoven FC, Peters AA, Trimbos JB, Hermans J, Fleuren GJ. The accuracy of frozen section diagnosis of ovarian tumors. Gynecol Oncol 1991 Jun;41(3):189-92.

(32) da Cunha BA, Salvatore CA, Faria RM. Frozen section biopsy of ovarian neoplasms. Int J Gynaecol Obstet 1983 Apr;21(2):103-10.

(33) DiSaia PJ, Creasman WT. The adnexal mass and early ovarian cancer. In: DiSaia PJ, Creasman WT, editors. Clinical gynecologic oncology. 6th edn ed. St Louis: Mosby; 2002. p. 259-88.

(34) Burger CW, Prinssen HM, Baak JP, Wagenaar N, Kenemans P. The management of borderline epithelial tumors of the ovary. Int J Gynecol Cancer 2000 May;10(3):181-97.

Table 1. Participant characteristics and scoring criteria in studies included

Study, Year	Mean age (range)	Period of study	*n**	Scoring of study quality[7]	Oxford evidence level[13]
Cuello *et al.*,[28] 1999	Not reported	1988-1998	489	Narrow population, verification complete, nonblinded, consecutive, retrospective, details test sufficient, details reference test sufficient, details population insufficient	Cohort 3B
Gol et al,[18] 2003	44.8 (14-78)	2000-2002	222	Narrow population, verification complete, nonblinded, consecutive, retrospective, details test sufficient, details reference test sufficient, details population sufficient	Cohort 2B
Hamed et al,[20] 1993	46.7 (9-81)	1987-1992	324	Narrow population, verification complete, nonblinded, nonconsecutive, retrospective, details test sufficient, details reference test sufficient, details population sufficient	Cohort 3B
Lim et al,[30]1997	Not reported	1988-1994	173	Narrow population, verification complete, nonblinded, consecutive, retrospective, details test sufficient, details reference test sufficient, details population insufficient	Cohort 3B
Obiakor et al,[29] 1991	Not reported	1980-1989	311	Narrow population, verification complete, nonblinded, nonconsecutive, retrospective, details test sufficient, details reference test sufficient, details population insufficient	Cohort 3B
Pinto et al,[26] 2001	Not reported	1994-1999	243	Narrow population, verification complete, nonblinded, nonconsecutive, retrospective, details test sufficient, details reference test sufficient, details population insufficient	Cohort 3B
Puls et al,[27] 1997	Not reported	12 years (years?)	294	Narrow population, verification complete, nonblinded, consecutive, retrospective, details test sufficient, details reference test sufficient, details population insufficient	Cohort 3B
Rose et al,[24] 1994	47 (1-95)† 61 (23-78)‡ 58 (16-87)§	1983-1993	383	Narrow population, verification complete, nonblinded, consecutive, retrospective, details test sufficient, details reference test sufficient, details population sufficient	Cohort 2B
Slavutin et al,[23] 1979	Not reported	1975-1977	55	Narrow population, verification complete, nonblinded, consecutive, retrospective, details test sufficient, details reference test sufficient, details population insufficient	Cohort 3B
Torres et al,[21] 1998	Not reported	1994-1997	126	Narrow population, verification complete, nonblinded, consecutive, retrospective, details test sufficient, details reference test sufficient, details population insufficient	Cohort 3B
Twaalfhoven et al,[31] 1991	51.4 (18-86)	1984-1990	176	Narrow population, verification complete, nonblinded, consecutive, retrospective, details test sufficient, details reference test sufficient, details population sufficient	Cohort 2B
Usubütün et al,[25]1998	Not reported	1991-1996	360	Narrow population, verification complete, nonblinded, nonconsecutive, retrospective, details test sufficient, details reference test sufficient, details population insufficient	Cohort 3B
Wakahara et al,[22] 2001	40.3 (11-79)	1994-1999	187	Narrow population, verification complete, nonblinded, consecutive, retrospective, details test sufficient, details reference test sufficient, details population sufficient	Cohort 2B
Yeo et al,[19] 1998	43.9 (3-89)	1990-1995	316	Narrow population, verification complete, nonblinded, consecutive, retrospective, details test sufficient, details reference test sufficient, details population sufficient	Cohort 2B

* Total number of ovarian tumor submitted to frozen section

† Benign ovarian tumor

‡ Borderline ovarian tumor

§ Malignant ovarian tumor

Table 2. Contingency tables for benign borderline and malignant ovarian tumors

Study	False positive	False negative	True negative	True positive
Benign versus borderline or malignant ovarian tumors*				
Cuello *et al.*,[28] 1999	7	3	81	392
Gol *et al.*,[18] 2003	7	5	66	139
Hamed *et al.*,[20] 1993	2	1	62	258
Lim *et al.*,[30]1997	2	1	41	127
Obiakor *et al.*,[29] 1991	11	0	68	224
Pinto *et al.*,[26] 2001	8	0	75	156
Puls *et al.*,[27] 1997	9	3	58	213
Rose *et al.*,[24] 1994	20	3	124	231
Slavutin *et al.*,[23] 1979	1	1	15	33
Torres *et al.*, [21] 1998	14	0	28	78
Twaalfhoven *et al.*,[31] 1991	8	1	62	90
Usubütün *et al.*,[25]1998	7	3	82	256
Wakahara *et al.*,[22] 2001	5	1	64	117
Yeo *et al.*,[19] 1998	11	1	52	249
Total	**112**	**23**	**878**	**2563**
Borderline versus benign ovarian tumors				
Cuello *et al.*,[28] 1999	6	6	392	14
Gol *et al.*,[18] 2003	3	5	139	10
Hamed *et al.*,[20] 1993	1	2	258	7
Lim *et al.*,[30]1997	1	1	127	7
Pinto *et al.*,[26] 2001	0	6	156	11
Puls *et al.*,[27] 1997	3	8	213	31
Rose *et al.*,[24] 1994	3	15	231	13
Slavutin *et al.*,[23] 1979	1	1	33	1
Twaalfhoven *et al.*,[31] 1991	1	6	90	8
Usubütün *et al.*,[25]1998	1	1	256	1
Wakahara *et al.*,[22] 2001	1	5	117	10
Yeo *et al.*,[19] 1998	1	8	249	12
Total	**22**	**64**	**2261**	**125**
Malignant versus benign ovarian tumors				
Cuello *et al.*,[28] 1999	3	4	392	67
Gol *et al.*,[18] 2003	2	2	139	56
Hamed *et al.*,[20] 1993	0	0	258	55
Lim *et al.*,[30]1997	1	1	127	34
Obiakor *et al.*,[29] 1991	0	11	224	68
Pinto *et al.*,[26] 2001	0	2	156	64
Puls *et al.*,[27] 1997	0	1	213	27
Rose *et al.*,[24] 1994	0	5	231	111
Slavutin *et al.*,[23] 1979	0	0	33	14
Torres *et al.*, [21] 1998	0	6	78	28
Twaalfhoven *et al.*,[31] 1991	0	2	90	54
Usubütün *et al.*,[25]1998	2	6	256	81
Wakahara *et al.*,[22] 2001	0	2	117	57
Yeo *et al.*,[19] 1998	0	3	249	40
Total	**8**	**45**	**2563**	**753**
Borderline versus malignant ovarian tumors				
Cuello *et al.*,[28] 1999	3	3	67	14
Gol *et al.*,[18] 2003	5	2	56	10
Hamed *et al.*,[20] 1993	0	1	55	7
Lim *et al.*,[30]1997	0	0	34	7
Pinto *et al.*,[26] 2001	3	1	64	11
Puls *et al.*,[27] 1997	10	1	27	31
Rose *et al.*,[24] 1994	4	1	111	13
Slavutin *et al.*,[23] 1979	5	0	14	1
Twaalfhoven *et al.*,[31] 1991	4	0	54	8
Usubütün *et al.*,[25]1998	0	0	81	1
Wakahara *et al.*,[22] 2001	0	0	54	10
Yeo *et al.*,[19] 1998	0	3	40	12
Total	**34**	**12**	**657**	**125**

* In benign *versus* borderline or malignant ovarian tumors, were not included the results that were diagnosed differently in frozen and paraffin section on between borderline and malignant ovarian tumor and, also were excluded cases that were deferred.

Table 3. Diagnostic accuracy of frozen section for benign *vs.* borderline or malignant ovarian tumors

Study	*n*	Sensitivity, % (95% CI)	Specificity, % (95% CI)	Likelihood (95% CI) Positive	Likelihood (95% CI) Negative	Posttest Probability (95% CI), %* Positive	Posttest Probability (95% CI), %* Negative
Cuello *et al.*,[28] 1999	483	99 (97-99)	92 (84-96)	12 (6-25)	0.008 (0.003-0.02)	96 (93-98)	1.8 (0.7-4.5)
Gol *et al.*,[18] 2003	217	96 (92-98)	90 (81-95)	10 (5-20)	0.039 (0.016-0.09)	96 (93-99)	8.5 (3-20)
Hamed *et al.*,[20] 1993	323	99 (97-100)	96 (89-124)	31.8 (8 -124)	0.004 (0.001-0.02)	98 (96-99)	0.9 (0.2-4)
Lim *et al.*,[30] 1997	171	99 (95-100)	95 (84-98)	21 (5.5-99)	0.008 (0.001-0.05)	98 (92-100)	1.8 (0.2-0.1)
Obiakor *et al.*,[29] 1991	303	98 (97-100)	85 (76-91)	6.9 (4-11.8)	0.03 (0 -0.04)	94 (91-97)	6.7 (0-8.7)
Pinto *et al.*,[26] 2001	239	99 (97-100)	89 (81-94)	9.8 (5.2-18.6)	0.004 (0-0.05)	95 (93-97)	0.9 (0-10.7)
Puls *et al.*,[27] 1997	283	98 (96-99)	86 (76-92)	7.3 (3.9-13.4)	0.016 (0.005-0.05)	94 (91-96)	3.6 (1-10.7)
Rose *et al.*,[24] 1994	378	98 (96-99)	86 (79-90)	7.1 (4.7-10.6)	0.015 (0.005-0.04)	94 (91-96)	3.4 (1-8.7)
Slavutin *et al.*,[23] 1979	50	97 (85-99)	93 (72-99)	15.5 (2-103)	0.032 (0.005-0.2)	97 (82-99)	7 (1.1-32)
Torres *et al.*,[21] 1998	120	99 (94-99.8)	66 (51-78)	2.9 (2-4.4)	0.01(0.001-1.15)	87 (82-91)	0.02 (0-0.7)
Twaalfhoven *et al.*,[31] 1991	161	98 (94-99)	88 (79-94)	8.6 (4.5-16.6)	0.012 (0.002-0.08)	95 (92-98)	2.7 (0.4-16)
Usubütün et al.,[25] 1998	348	98 (96-99)	92 (84-90)	12 (6-25)	0.013 (0.004-0.03)	96 (94-98)	3 (0.9-6.7)
Wakahara *et al.*,[22] 2001	187	99 (95-100)	92 (84-96)	13.6 (5.8-32)	0.009 (0.001-0.06)	97 (93-98)	2 (0.2-12.5)
Yeo *et al.*,[19] 1998	313	99 (97-100)	82 (71-90)	5.7 (3.3-9.7)	0.005 (0.001-0.04)	93 (90-96)	1 (0.2-8.7)
Total	**3576**	**99 (98-99)**	**88 (86-90)**	**8.7 (7.3-10.4)**	**0.01 (0.007-0.015)**	**95 (94-96)**	**2.3 (1.6-3.4)**

* The 95% CIs for posttest probability were calculated by using pretest odds and limits of 95% CIs of likelihood ratios.
Pretest odds=prevalence/1-prevalence
Posttest odds= pretest odds x likelihood ratio
Posttest probability= posttest odds/1+posttest odds

Table 4. Diagnostic accuracy of frozen section for malignant *vs.* benign ovarian tumors

Study	*n*	Sensitivity,% (95% CI)	Specificity,% (95% CI)	Likelihood (95% CI)		Posttest Probability (95% CI), %*	
				Positive	Negative	Positive	Negative
Cuello *et al.*,[28] 1999	466	94 (86-97)	99 (97-99)	124 (40-384)	0.057 (0.02-0.14)	97 (95-98)	1.6 (0.5-3.9)
Gol *et al.*,[18] 2003	199	96 (88-99)	98 (95-99)	68 (17-269)	0.035 (0.009-0.13)	95 (92-98)	1 (0.2-3.6)
Hamed *et al.*,[20] 1993	313	99 (92-99)	99 (98-100)	513 (32-8186)	0.009 (0.001-0.14)	99 (98-100)	0.2 (0.02-3.9)
Lim *et al.*,[30] 1997	163	97 (85-99)	99 (95-100)	124 (17-876)	0.029 (0.004-0.19)	97 (95-99)	0.8(0.11-5.2)
Obiakor *et al.*,[29] 1991	303	85 (76-91)	99 (97-100)	385 (24-6150)	0.144 (0.08-0.24)	99 (98-100)	4 (2.2-6.5)
Pinto *et al.*,[26] 2001	222	96 (88-98)	99 (97-100)	302 (19-4813)	0.052 (0.01-0.24)	98 (97-99)	1.4 (0.2-6.5)
Puls *et al.*,[27] 1997	241	94 (80-98)	99 (97-100)	405 (25-6476)	0.016 (0.005-0.05)	99 (98-100)	0.4 (0.1-1.4)
Rose *et al.*, [24] 1994	347	95 (89-97)	99 (98-100)	442 (27-7050)	0.047 (0.02-0.1)	99 (98-100)	1.3 (0.5-2.8)
Slavutin *et al.*,[23] 1979	47	96 (74-99)	98 (87-99)	65 (4-1031)	0.187 (0.09-0.37)	95 (89-100)	5 (2.5-9.6)
Torres *et al.*,[21] 1998	112	81 (65-90)	99 (94-99)	128 (10-2762)	0.04 (0.013-0.14)	97 (94-100)	1 (0.3-3.9)
Twaalfhoven *et al.*,[31] 1991	146	95 (86-98)	99 (95-99)	174 (4.5-16.6)	0.012 (0.002-0.08)	98 (96-100)	0.3 (0.5-2.2)
Usubütün *et al.*,[25] 1998	345	93 (85-97)	99 (97-99.8)	120 (30-478)	0.07 (0.03-0.15)	97 (96-98)	1.9 (0.8-4.1)
Wakahara *et al.*,[22] 2001	173	95 (86-98)	99 (96-100)	332 (20-5300)	0.04 (0.013-0.14)	99 (97-100)	1.1 (0.3-3.9)
Yeo *et al.*,[19] 1998	292	92 (80-97)	99 (98-100)	470 (28-7347)	0.058 (0.017-0.19)	99 (98-100)	1.6 (0.4-5.2)
Total	**3369**	**94 (92-95)**	**99 (98-100)**	**303 (151-605)**	**0.057 (0.04-0.07)**	**98 (97-99)**	**1.6 (1.1-1.9)**

* The 95% CIs for posttest probability were calculated by using pretest odds and limits of 95% CIs of likelihood ratios.

Table 5. Diagnostic accuracy of frozen section for borderline *vs.* benign ovarian tumors

Study	*n*	Sensitivity,% (95% CI)	Specificity,% (95% CI)	Likelihood (95% CI)		Posttest Probability (95% CI), %*	
				Positive	Negative	Positive	Negative
Cuello *et al.*,[28] 1999	418	70 (48-85)	98 (96-99)	46 (19-108)	0.3 (0.15-0.59)	72 (52-86)	1.7 (0.8-3.3)
Gol *et al.*,[18] 2003	157	66 (41-84)	98 (94-99)	33 (10-106)	0.34 (0.14-0.62)	65 (36-86)	1.9 (0.8-3.5)
Hamed *et al.*,[20] 1993	268	77 (45-93)	99 (97-100)	201 (27-1469)	0.22 (0.06-0.75)	92 (61-98)	1.2 (0.3-4.1)
Lim *et al.*,[30] 1997	136	87 (52-99)	99 (95-100)	112 (15-802)	0.12 (0.02-0.78)	86 (46-97)	0.7 (0.1-4.3)
Pinto *et al.*,[26] 2001	173	63 (41-81)	99 (97-100)	200 (12-3263)	0.36 (0.19-0.67)	92 (65-99)	2.1 (1.1-3.7)
Puls *et al.*,[27] 1997	255	79 (64-89)	98 (96-99)	57 (18-178)	0.2 (0.11-0.38)	77 (51-91)	1.1 (0.6-2.1)
Rose *et al.*,[24] 1994	262	46 (29-64)	98 (96-99)	36 (10-119)	0.5 (0.38-0.76)	67 (36-87)	2.8 (2.1-4.2)
Slavutin *et al.*,[23] 1979	36	50 (48-99)	97 (85-99)	17(1.5-183)	0.5 (0.12-2)	50 (8-91)	2.8 (0.7-10.3)
Twaalfhoven *et al.*,[31] 1991	105	57 (32-78)	98 (94-99)	52 (7-384)	0.43 (0.23-0.79)	75 (39-96)	2.4 (1.3-4.3)
Usubütün *et al.*,[25] 1998	259	50 (25-97)	99 (97-100)	128 (11-1412)	0.5 (0.12-2)	88 (38-98)	2.8 (0.7-10.3)
Wakahara *et al.*,[22] 2001	133	66 (41-84)	99 (95-100)	78 (10-572)	0.3 (0.16-0.68)	82 (36-97)	1.7 (0.9-3.8)
Yeo *et al.*,[19] 1998	270	60 (38-78)	99 (97-100)	150 (20-1095)	0.4 (0.2-0.68)	89 (53-98)	2.2 (1.1-3.7)
Total	**2472**	**66 (59-72)**	**99 (98-100)**	**68 (44-105)**	**0.34 (0.28-0.41)**	**79 (71-85)**	**1.9 (1.6-2.3)**

* The 95% CIs for posttest probability were calculated by using pretest odds and limits of 95% CIs of likelihood ratios.

Table 6. Diagnostic accuracy of frozen section for borderline *vs.* malignant ovarian tumors

Study	*n*	Sensitivity,% (95% CI)	Specificity,% (95% CI)	Likelihood (95% CI)		Posttest Probability (95% CI), %*	
				Positive	Negative	Positive	Negative
Cuello *et al.*,[28] 1999	87	82 (59-93)	95 (88-98)	19 (6-59)	0.18 (0.06-0.5)	52 (26-77)	1 (0.3-2.8)
Gol *et al.*,[18] 2003	73	83 (55-95)	91 (82-96)	10 (4-24)	0.18 (0.05-0.6)	36 (18-58)	1 (0.2-3.3)
Hamed *et al.*,[20] 1993	63	83 (50-96)	99 (92-99)	93 (5.8-1496)	0.16 (0.04-0.7)	84 (25-98)	0.9 (0.2-4)
Lim *et al.*,[30] 1997	41	93 (59-99)	98 (87-99)	65 (4-1034)	0.06 (0.004-0.9)	79 (18-98)	0.3 (0.02-5)
Pinto *et al.*,[26] 2001	79	91 (64-99)	95 (87-98)	20 (6-62)	0.08 (0.01-0.6)	53 (25-78)	0.4 (0.05-3)
Puls *et al.*,[27] 1997	69	96 (84-99)	73 (57-84)	3.5 (2-6)	0.04 (0.006-0.3)	17 (10-26)	0.2 (0.03-1.7)
Rose *et al.*,[24] 1994	129	92 (68-99)	96 (91-98)	26 (10-70)	0.07 (0.01-0.49)	60 (36-80)	0.4 (0.05-2.7)
Slavutin *et al.*,[23] 1979	20	75 (19-97)	72 (50-87)	2.7 (0.9-8)	0.3 (0.03-3.8)	7.2 (4.9-31)	1.7 (0.17-18)
Twaalfhoven *et al.*,[31] 1991	66	94 (62-99)	92 (82-96)	12 (5-30)	0.06 (0.004-0.9)	41 (22-64)	0.3 (0.02-5)
Usubütün *et al.*,[26] 1998	82	75 (20-97)	99 (94-99)	123 (7-2184)	0.2 (0.02-2.7)	87 (28-99)	1 (0.1-13)
Wakahara *et al.*,[22] 2001	64	95 (67-99)	99 (91-99)	105 (6.6-1662)	0.04 (0.003-0.7)	86 (27-99)	0.2 (0.01-4)
Yeo *et al.*,[19] 1998	55	96 (71-99)	92 (80-97)	12 (4.4-33)	0.04 (0.003-0.6)	41 (20-66)	0.2 (0.01-3)
Total	**828**	**91 (85-99)**	**95 (93-96)**	**18 (13-26)**	**0.09 (0.05-0.15)**	**51 (42-60)**	**0.5 (0.2-0.9)**

* The 95% CIs for posttest probability were calculated by using pretest odds and limits of 95% CIs of likelihood ratios.

Table 7. Distribution of histological types in ovarian tumors according to paraffin diagnosis

Classification	Benign (%)	Borderline (%)	Malignant (%)	Total (%)
Epithelial tumour				
Serous cystadenoma[18,19,20,22,23,28,31]	544 (58.8)	51(42.5)	128 (38.6)	
Mucinous cystadenoma[18,19,20,22,28,31]	202 (21.8)	49 (40.8)	27 (8.1)	
Serous cystadenofibroma[20,23,28,31]	28 (3)	1 (0.8)		
Mucinous cystadenofibroma[20,28,31]	8 (0.8)			
Endometriod cystadenomafibroma[20,28]	1 (0.1)		1 (0.3)	
Carcinoma endometriod[18,19]			19 (5.7)	
Mixed seromucinous cystadenoma[31]	4 (0.43)			
Brenner tumor[19,28,31]	7 (0.75)			
Clear cell [19,20,22,23,28,31]			42 (12.6)	
Transitional cell tumour[28]			1 (0.3)	
Endometrioid[18,22,23,28,31]	65 (7)		34 (10.2)	
Mixed epithelial carcinoma[20,28,31]			7 (2.1)	
Undifferentiated carcinoma[19,20,22,23,28,31]			33 (9.9)	
Unclassified epithelial carcinoma[22,31]			4 (1.2)	
Summary of one study[26]	66 (7)	19 (15.8)	35 (10.5)	
Total of epithelial tumour*	**925 (40)**	**120 (5.2)**	**331 (14.5)**	**1326 (58)**
Sex cord-stromal tumours				
Fibrothecoma[18,19,20,22,23,26,28,31]	37 (32.4)		1 (4)	
Fibroma[18,19,20,22,23,26,28,31]	75 (65.7)		1 (4)	
Granulosa cell[18,19,20,22,23,26,28,]		2 (100)	20 (80)	
Sex cord stromal unclassified[18,26]	2 (1.7)			
Androblastoma[19,26]			3 (12)	
Total of sex cord-stromal tumour*	**114 (4.9)**	**2 (0.08)**	**25 (1.1)**	**141 (6)**
Germ cells				
Dermoid cyst [18,20,26,28,31]	103 (28.2)			
Teratoma[20,22,28,31]	192 (52.6)		3 (8.1)	
Struma ovarii[19,22,]	8 (2.1)			
Immature teratoma [18,19,26]	62 (16.9)		6 (16.2)	
Dysgerminoma[18,19,20,22,26,28]			16 (43.2)	
Mix[18]			5 (13.5)	
Yolk sac tumour (endodermal)[20,22,26,28]			7 (18.9)	

Classification	Benign (%)	Borderline (%)	Malignant (%)	Total (%)
Total of Germ cells*	**365 (16)**		**37 (1.7)**	**402 (17)**
Secondary Metastatic[18,20,22,26,28]			39 (1.7)	**39 (1.7)**
Tumours of uncertain origin[18]			5 (0.2)	**5 (0.2)**
Tumor-like conditions				
Para-ovarian[20,23]	21 (6.6)			
Folicular cysts[19,22,28]	65 (20.4)			
Endometriosis[18,19,20,28,31]	181 (56.9)			
Tubo-ovarian abscess[18,22]	6 (1.8)			
Torsion of adnexa[18]	5 (1.5)			
Simple cysts[18]	40 (12.5)			
Corpus luteum cyst[19,20,22,28,31]	82 (25.7)			
Inflammatory lesions[20,31]	6 (1.8)			
Total of tumor-like conditions*	**318 (14)**			**318 (14)**
Total of permanent section[18,19,20,22,23,26,28,31]				**2281**

* The total rate of histological type was calculated according of total number in permanent section

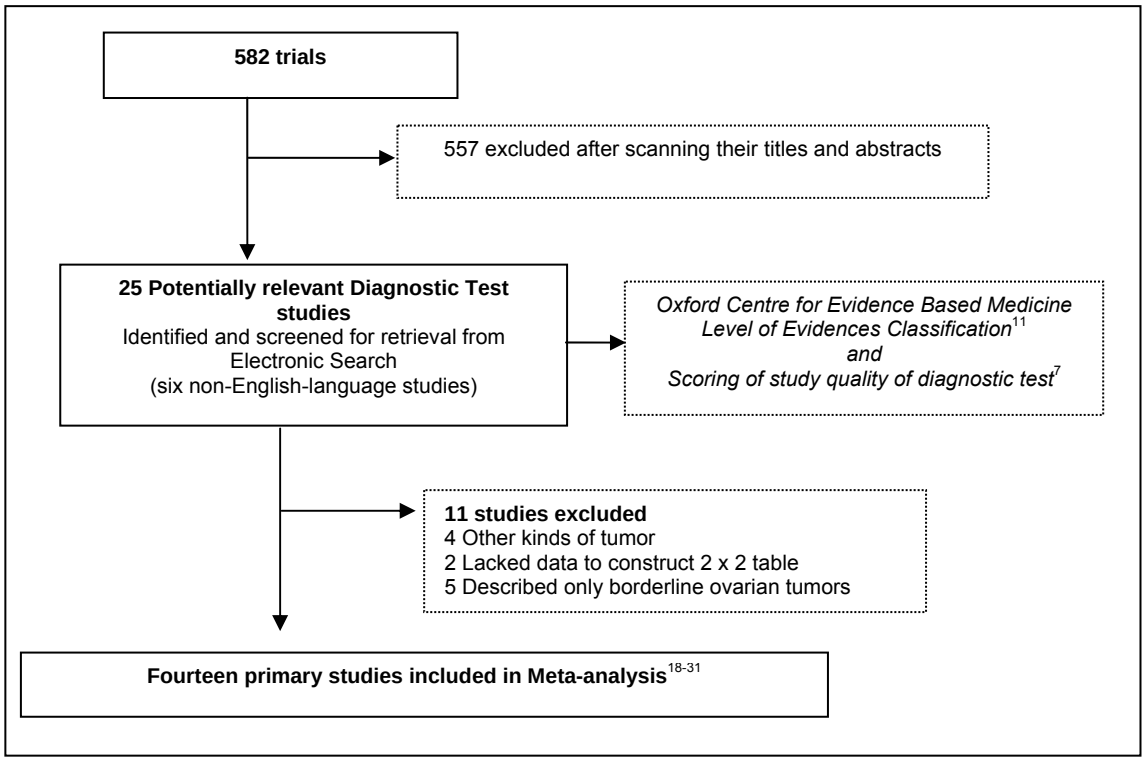


Figure 1. Study selection process.

5.3 Artigo 3

ACCURACY OF ULTRASONOGRAPHY WITH COLOR DOPPLER IN OVARIAN TUMOR: A SYSTEMATIC QUANTITATIVE REVIEW

International Journal of Gynecological Cancer
Manuscript ID: IJGC-2008-00021
Title: Accuracy of ultrasonography with color Doppler in ovarian tumor: a systematic quantitative review
Authors: Medeiros, Lídia
Rosa, Daniela D.
Rosa, Maria
Bozzetti, Mary
Date Submitted: 09-Jan-2008

Accuracy of Ultrasonography with Color Doppler in Ovarian Tumor: A Systematic Quantitative Review

Author: Lidia Rosi Medeiros, MD, MSc(1)
Address: José de Alencar 1244 apt 1009, Porto Alegre, RS, Brasil, CEP 90880-480
lidia.rosi@terra.com.br

Co-authors
Daniela Dornelles Rosa, MD, PhD(2,3)
Maria Inês da Rosa, MD, PhD(1, 4)
Mary Clarisse Bozzetti, MD, PhD(1, 2, 5)

Affiliations of all authors:
[1] Postgraduate Program in Epidemiology at Federal University of Rio Grande do Sul, Porto Alegre Brazil
[2] Hospital Fêmina and Hospital Moinhos de Vento, Porto Alegre, Brazil
[3] Postgraduate Program in Medicine: Medical Sciences at Federal University of Rio Grande do Sul, Porto Alegre, Brazil
[4] Medical School at University of Extremo Sul Catarinense, Criciúma, Brazil
[5] Department of Social Medicine, Faculty of Medicine at Federal University of Rio Grande do Sul, Porto Alegre Brazil

RESUMO

Foi realizada uma revisão sistemática quantitativa para estimar a acurácia da ultra-sonografia com Doppler colorido nas tumorações ovarianas, tendo sido incluídos todos estudos que avaliam esse exame e comparam o resultado final com o diagnóstico anatomopatológico, considerado exame de referência. Foram analisados 12 estudos que incluíram 2.398 mulheres. O somatório de estudos para o cálculo de sensibilidade encontrou um valor de 0,87 (IC95% de 0,84-0,90); para especificidade esse valor foi de 0,92 (IC 95% de 0.87-0.90). A DOR para tumores ovarianos ou com malignidade limítrofe *vs* lesões benignas foi de 125 (95% CI de 55-283). Foi realizado um sumário da curva ROC (SROC) devido à heterogeniedade encontrada na DOR. Entretanto, embora haja heterogeneidade na comparação dos tumores malignos ou malignidade límitrofe *vs.* lesões benignas, a área sob a curva foi de 0,9573. Em conclusão, ultra-sonografia com Dopller colorido é um exame pré-teste importante na predição da natureza da tumoração ovariana – se maligna ou benigna.

Palavras-chave: ultra-sonografia, color Doppler, tumors de ovário**,** revisão sistemática, metanálise, acurácia diagnóstica.

ABSTRACT

A quantitative systematic review was performed to estimate the accuracy ultrasonography with color Doppler in the diagnosis of ovarian tumors. Studies that compared color Doppler ultrasonography with paraffin-embedded sections parameters for the diagnosis of ovarian tumors were included. Twelve studies were analyzed, which included 2,398 women. The pooled sensitivity was 0.87 (IC95% 0.84-0.90); and the specificity was 0.92 (IC 95% 0.87-0.90). The DOR for ovarian cancer and borderline lesions *vs* benign lesions was 125 (95% CI, 55-283). SROC curves were constructed due to heterogeneity in the DOR. For malignant ovarian cancer and borderline *vs*. benign lesions the AUC was 0.9573. In conclusion, ultrasonography with color Doppler is a useful pre-operative test for predicting the diagnosis of pelvic masses.

Keywords: ultrasonography, color Doppler, ovarian tumor**,** systematic review, meta-analysis, accuracy, diagnosis.

INTRODUCTION

Functional and benign ovarian cysts are the most frequent abnormal structural gynaecological findings in women in reproductive age, and are among the five main causes of hospitalisation for gynaecological disease in the United States and in England[1]. Data from the National Hospital Discharge Survey from 1988 to 1990 found that the average annual rates of hospitalization for benign ovarian cysts was 32.7 (95% CI 28.8-36.6) per 10,000 women in reproductive age, and 68% of these women were submitted to surgical procedures[2].

There is still a need for a non-invasive procedures that effectively discriminate between malignant and benign adnexal masses[3]. The use of transvaginal color flow imaging has facilitated the study of vascular change within the pelvis through the identification of areas of angiogenesis. Malignant tumors exhibit the reliance on new vessel formations. Low impedance of the intramutumoral flow is usually observed in cases of primary ovarian cancer and the resistance index (RI) is commonly was below 0.4 in these cases[4]. In addition, the pulsatility index (PI) is a useful way of expressing blood flow impedance distal to the point of sampling. PI values ≤ 1 suggested the presence of borderline or malignant ovarian tumors[5].

To avoid false-positive findings like angiogenesis in the corpus luteum, premenopausal women should be examined between 3th and the 8th day of their menstrual cycle[5]. Kurjak *et al*. showed the ability of transvaginal color Doppler sonography in the detection of ovarian cancer as early as stage I, both in asymptomatic women and in the morphologically normal ovaries[6].

We performed a systematic quantitative review of the literature to ascertain of the accuracy of transvaginal ultrasound with color Doppler in the diagnosis of ovarian cancer and to explore the reasons for the ongoing controversies about this issue.

METHODS

Identification of studies

A comprehensive search of the MEDLINE, CANCERLIT, LILACS and EMBASE databases was made from January 1990 to December 2007. The medical subjects heading (MeSH) and text words for the terms "*ovarian neoplasm"* and

"*transvaginal ultrasound with color Doppler*" were combined with the MeSH term *diagnosis* ("sensitivity and specificity"). The search was limited to human studies but had no language restrictions. In addition, the Cochrane Library was searched. Reference lists of all available primary studies were reviewed to identify additional relevant citations. The authors who published the studies were not contacted.

Selection criteria

This review focused on observational studies evaluating clinically suspected adnexal masses with 5-MHz, transvaginal probe ultrasonography with color Doppler in which the results of the diagnostic test of interest were compared with the results of a reference standard. Malignancy was suspected when the resistance index was ≤ 0.5. The cases studies were women treated surgically due to ovarian tumors. The diagnostic test was the transvaginal ultrasound using color Doppler and the diagnostic reference was the result of the histological analysis of standard paraffin-embedded sections after surgery. A transvaginal ultrasound using color Doppler diagnosis was considered correct if it did not differ from that of the paraffin section. For inclusion in the systematic review, the final histological diagnosis should have been designated as benign, borderline or malignant in the selected studies. We excluded the studies that analyzed only borderline ovarian tumors and malignant ovarian tumor as well as the studies that lacked data to construct 2 x 2 contingency tables.

The final diagnoses of ovarian lesions were grouped and compared as malignant or borderline *versus* benign. To calculate the diagnostic accuracy were excluded the cases that were deferred due to uncertain transvaginal ultrasound using Doppler. Thus, the primary outcome analyzed was the accuracy of ovarian tumor diagnosis (benign, borderline and malignant) by transvaginal ultrasound using Doppler analysis. A secondary outcome was the distribution of histological types of ovarian tumors according to paraffin-embedded tissue diagnosis. The reviewed studies were identified independently by three investigators (L.R.M, D.D.R, M.I.R). Final inclusion or exclusion was made with reference to a selection criteria checklist. Disagreements about study inclusion or exclusion were initially solved by consensus, and when this was not possible, they were resolved by arbitrarily by a fourth reviewer (M.C.B). Statiscal agreement among reviewers was computed.

Quality assessment

All articles meeting the eligibility criteria were assessed for their methodological quality. This assessment involved scrutinization the study designs and of the relevant featuress of the patient population, the diagnostic test and the reference standard[7-10]. These features included the methods of data collection and patient selection, description of transvaginal ultrasound using color Doppler and the histological reference standard, and presence of verification of biases[8, 9]. The results of the quality assessment for the included studies are summarized in the "*Scoring of Study Quality*" column of table 1[10]. Studies were further assessed for methodological quality with reference to the *Oxford Centre for Evidence Based Medicine Level of Evidences Classification* rubric*.* Only studies with Oxford Evidence Levels 1 to 3 were considered to be of high quality while those with levels 4 and 5 were excluded[11].

Data abstraction

Three investigators (L.R.M, D.D.R, M.I.R) independently abstracted data regarding the prevalence of benign, borderline and malignant ovarian lesions. They also calculated the sensitivities, specificities, true-positive rates (TPR; sensitivity) and the false-positive rates (FPR; 1 – specificity) from the primary studies of transvaginal ultrasound with Doppler diagnosis. The assessment of English-language articles was performed by all reviewers (L.R.M, D.D.R, M.I.R) while assessment of those articles published in languages other than English was performed independently by a fourth reviewer (M.C.B) following translation (when necessary). Any disagreement was resolved by consensus for both English and non-English studies. Three ovarian tumors diagnosis outcomes were considered: benign, borderline and malignant (Table 2). Data for the accuracy of transvaginal ultrasound with color Doppler in the diagnosis of malignant tumor cases were abstracted as 2 x 2 tables of diagnosis (malignant or borderline ovarian tumors *versus* benign lesions) and paraffin-embedded section diagnosis.

Data synthesis and statistical analysis

To evaluate the agreement between study eligibility and assessment of methodological quality, as well as between the results of transvaginal ultrasound with Doppler and paraffin-embedded section analyses, the observed percentage of agreement and the κ coefficient for inter-rater reliability were calculated[12, 13]. For each study, 2 x 2 contingency tables were constructed, in which all biopsies were classified as normal or benign lesion, borderline and malignant ovarian tumors . The true-positive (TPR; sensitivity) and the false-positive rates (FPR; 1 – specificity) were calculated. When 2 x 2 tables had 0 cells, the value of 0.5 was added to each cell to enable calculations to be made; when a study contained two cells with the 0 value, it was excluded from the analysis.

The summary weighted sensitivity and specificity were calculated as the sum of sensitivities and specificities reported for each study, multiplied by the number of subjects in the study, divided by the total number of subjects in all studies. The 95% CIs for mean weighted results were calculated using the exact method. The association between sensitivity and specificity for normal or benign ovarian lesions, borderline, and malignant ovarian cancer was calculated with the *Spearman's correlation coefficient test*[13-15]. When there was no correlation, pooling sensitivities and specificities were calculated, since there were two categories of results (negative or positive test)[14]. In the case of correlation, a summary receiver operating characteristic curve (SROC) was generated using data from all thresholds, using the Littenberg and Moses method[14-16].

The diagnostic odds ratio (DOR) can relate to different combinations of sensitivities and specificities. The DOR describes the odds of the positive test results in participants with disease compared with the odds of positive test results in those without disease. A single diagnostic odds ratio corresponds to a set of sensitivities and specificities depicted by the SROC. It can change according to the threshold and to the ROC curve used to define an abnormal examination resulted in the expected trade-off between sensitivity and specificity. Also, the area under the curve (AUC) can summarize the inherent capacity of a test for discriminating a diseased from a non-diseased subject. Perfect tests usually have AUCs close to 1 and poor tests usually have AUCs close to 0.5[14-16]. The heterogeneity of the sensitivities and specificities of the studies were assessed using the Q_T (Cochran) test for χ^2

distributions with N -1 degrees of freedom. Since sensitivity and specificity were heterogeneous, a random effect model was used and the terms were pooled with a 95% CI[14-17]. The statistical analysis was performed with the software Meta-DiSc® (version Beta 1.1.1)[18].

RESULTS

The process of study selection is summarized in figure 1. Our initial search identified 312 potentially relevant articles. We excluded 254 published studies after reviewing their titles and abstracts, because the three independent reviewers considered that they did not relate to the question under review. Fifty-eight full-text articles were retrieved studies; 46 were excluded after further scrutiny. A complete list of excluded studies is available from the authors. Twelve primary studies, including 2,398 women, met the criteria for inclusion and were analyzed (Table 1)[19-30]. Interrater overall agreement for study eligibility and methodological quality was 88% (κ = 0.67), indicating good agreement[13]. Disagreement between reviewers occurred during analysis of the twelve studies and they were related to inclusion or exclusion criteria (characteristics and scoring of study quality). However, it was solved by consensus.

Studies description

Details of the participants, interventions, and quality assessments of the studies selected for the meta-analysis are summarized in Table 1. The mean age of participants was reported in six studies[21, 23, 26, 28-30]. Six studies were non-blinded and prospective with a small population, but included sufficient experimental details, proper diagnostic tests and diagnostic reference standards[20, 21, 23, 26, 28, 30]. Of the twelve included studies, six had high[20, 21, 23, 26, 28, 30] methodological quality, satisfying ≥ 55% of the criteria for study quality and with an Oxford Evidence Level of 2B[11]. However, six studies[19, 22, 24, 25, 27, 29] were classified as level 3B, because details of the patients, including age, were not reported and because they met < 50% of the quality criteria.

Benign ovarian tumors were found in 1,836 patients (77.1%), borderline tumors in 74 patients (3.1%) and malignant tumors in 488 patients (20.4%) (Table 2).

Table 3 showed the results of contingency tables (FN, FP, TP and TN) from each study considered in the systematic review. In Guerriero *et al*.[21] the difference in the total numbers presented in tables 1 and 3 are due to exclusion of 19 borderline cases when analyzing the IR in transvaginal probe sonography with color Doppler.

The association between sensitivity and specificity for benign and borderline or malignant ovarian cancer was calculated with the *Spearman's correlation coefficient test*[13]. The correlations coefficient was 0.042, p = 0.897. When there was no correlation, pooling sensitivities and specificities were calculated, since there were two categories of results (negative or positive test)[14, 16]. Since sensitivity and specificity were heterogeneous, a random effect model was used and the terms were pooled with a 95% CI[14-17]. Interrater overall agreement between ultrasonography with Doppler and paraffin sections was 88% (k = 0.87), indicating very good agreement (Table 3)[12].

Sensitivity

Ultrasonography with color Doppler was better at detecting ovarian cancer or borderline lesions when compared with benign lesions in ovarian with a pooled sensitivity was 0.87 (IC95% 0.84-0.90). The estimates for heterogeneity were highly consistent across studies: for normal or benign lesion *vs*. borderline or malignant lesions (Q_T = 74.4 *p* = 0.001, inconsistency I^2 = 85%) (table 4).

Specificity

Specificity for ultrasonography with color Doppler was higher for normal or benign lesion *vs*. borderline or malignant ovarian cancer with a pooled specificity was 0.90 (IC 95%, 0.87-0.90). The estimates for heterogeneity were highly consistent across studies: for normal or benign lesion *vs*. borderline or malignant lesions (Q_T = 167, *p* = 0.001, inconsistency I^2 = 93%) (table 4).

DOR

The DOR in ultrasonography with color Doppler among borderline and malignant ovarian cancer *vs*. benign lesions was 125 (IC95% 55-283; Q_T = 33.7, *p* = 0.001; inconsistency I^2 = 67%). SROC curves were constructed due to heterogeneity in the DOR[14-17]. For malignant ovarian cancer or borderline tumors *vs*. benign lesions the AUC was 0.9573 for ultrasonography with Doppler. The AUC for the ROC curve was estimated by trapezoidal rule (Meta-DiSc®; version Beta 1.1.1)[18]

(Figure 2). An alternative summary ROC curve allowing for variation in the diagnostic odds ratio with diagnostic threshold can be estimate by the method of Moses and Littenberg[14-17]. First, the log of the diagnostic odds ratio (D) on the measure of the diagnostic threshold (*S*), weighting by study size, produces an estimate of the parameters *a* and *b* from the regression equation, D = *a*+*bS,* the parameter estimates are found in table 5. The results suggest that diagnostic odds ratio do not change with the threshold (p = 0.13). ROC curves, including meta-analysis SROC curves, do not specify the exact operating point (the threshold and its associated TRP and FRP) that is the best.

Sensitivity analysis

The robustness of the results was tested by repeating the analysis using a different statistical model (random effects model)[14-17]. The pooling of sensitivity, specificity and DOR from the six studies with high methodological quality did not alter the accuracy rate for diagnosis by ultrasonography with color Doppler analysis for benign, borderline or malignant ovarian tumors[20, 21, 23, 26, 28, 30]. Therefore, the twelve studies were included in the sensitivity analysis[19-30]. Inverted funnel plots showed asymmetry.

DISCUSSION

In summary, this systematic review showed that ultrasonography with Doppler can detect malignant or borderline lesion when the resistance index (RI) was below 0.5. The sensitivity for detecting malignant ovarian cancer with ultrasonography with Doppler was 87% and the specificity was 90%. The DOR was 125 and the AUC was 0.9661 for detection of malignant or borderline *vs* benign ovarian tumor. Therefore, ultrasonography with Doppler is a useful pre-operative test for the prediction of benign or malignant nature of pelvic mass.

The differentiation of benign from malignant adnexal masses is of great value, since the therapeutic approach is different for these two entities. Benign ovarian masses can be managed with a more conservative approach, either closed observation or laparoscopic surgery. On the contrary, when the tumor is the malignant there is a need for urgent laparotomy. Bourne *et al*.,[4] showed that all

cases of invasive ovarian tumors had neovascularization with low resistance index below 0.5. Kurtz *et al.*, found the same results in discrimination between cases of benign and case malignant disease with that ultrasonography with Doppler, computed tomography and magnetic resonance imaging had overall pooled areas under the ROC curve was 0.91 (95% CI 0.87-0.95) for the three different methods[31]. Ultrasonography with Doppler is often the initial imaging study in the evaluation of a suspected ovarian abnormality.

Valentin *et al*. showed that subjective evaluation of the color content of the tumor scan is the best Doppler method for distinguishing between benign and malignant tumors and that low color content suggests benignity. They concluded that blood flow velocity was the better discriminator between benign and malignant pelvic tumors[31, 32]. For Kurjak *et al*. reported that when the vessel were located in the central, septal, or papillary projections, together with a diffuse vascular arrangement, and the RI was less than 0.4, the tumors was likely to be malignant[33].

Brown *et al*. showed the four sonographic features that are statistically significant contributors to the distinction between benign an malignant adnexal tumors: presence and nature of any solid component; presence and location of flow at conventional color Doppler; presence and amount of free intraperitoneal fluid; presence and thickness of septations[33, 34].

It is important to note that a possible limitation of this systematic review is a potential bias in that all trials were retrospective and that there was a lack of blinding in their assessment. The probable reasons for this bias are the small number of trials included and the lower quality of some of the trials (nonconsecutive, insufficient population details)[9, 10]. On the other hand, this meta-analysis complied with the criteria for performing a rigorous systematic review planned a priori[14-16]. These included the use of study quality assessment[10] and investigation of homogeneity with fixed and random models to test the robustness of the results[14-16]. SROC curves were constructed in cases of heterogeneity in the DOR with calculated the AUC[14-18].

We suggest that blinded prospective studies should be performed to investigate the diagnostic accuracy of ultrasonography with color Doppler analysis in the diagnosis of ovarian tumors associated with another exam as such as CA 125 correlated to age (premenopausal or posmenopausal patients).

This quantitative review provides precise evaluation of the accuracy of ultrasonography with color Doppler in the diagnosis of benign, borderline and malignant ovarian tumors. The area under the ROC curve (AUC = 0.9577) reflect the tendencies of readers to overstage an understage the disease. However, the data should be analyzed according the clinical process. Better methods to improve the accuracy rates of the diagnosis of borderline ovarian tumors are needed.

Contributors

LRM had full access to all data and took responsibility for the integrity and the accuracy of the analysis. DDR, MIR and MCB organized the study concept and the design. LRM acquired, analyzed and interpreted the data. LRM, DDR, MIR and MCB drafted the manuscript. LRM and DDR performed the statistical analysis.

Conflict of interest statement

We declare no conflict of interest.

Reference List

(1) Westhoff CL, Beral V. Patterns of ovarian cyst hospital discharge rates in England and Wales, 1962-79. Br Med J (Clin Res Ed) 1984 Nov 17;289(6455):1348-9.

(2) Velebil P, Wingo PA, Xia Z, Wilcox LS, Peterson HB. Rate of hospitalization for gynecologic disorders among reproductive-age women in the United States. Obstet Gynecol 1995 Nov;86(5):764-9.

(3) Tailor A, Jurkovic D, Bourne TH, Collins WP, Campbell S. Sonographic prediction of malignancy in adnexal masses using multivariate logistic regression analysis. Ultrasound Obstet Gynecol 1997 Jul;10(1):41-7.

(4) Bourne TH. Transvaginal color Doppler in gynecology. Ultrasound Obstet Gynecol 1991 Sep 1;1(5):359-73.

(5) Bourne T, Campbell S, Steer C, Whitehead MI, Collins WP. Transvaginal colour flow imaging: a possible new screening technique for ovarian cancer. BMJ 1989 Dec 2;299(6712):1367-70.

(6) Kurjak A, Shalan H, Matijevic R, Predanic M, Kupesic-Urek S. Stage I ovarian cancer by transvaginal color Doppler sonography: a report of 18 cases. Ultrasound Obstet Gynecol 1993 May 1;3(3):195-8.

(7) Cochrane Methods Working Group on Systematic Reviews of Screening and Diagnostic Tests. Recommended methods. Cochrane Library 2004 January 25 [cited 2004 Jan 25];Available from: URL: http://www.cochrane.org/cochrane/sadtdoc1.htm

(8) Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD Initiative. Ann Intern Med 2003 Jan 7;138(1):40-4.

(9) Jaeschke R, Guyatt G, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. A. Are the results of the study valid? Evidence-Based Medicine Working Group. JAMA 1994 Feb 2;271(5):389-91.

(10) Lijmer JG, Mol BW, Heisterkamp S, Bonsel GJ, Prins MH, van der Meulen JH, et al. Empirical evidence of design-related bias in studies of diagnostic tests. JAMA 1999 Sep 15;282(11):1061-6.

(11) Phillips B. Oxford Centre for evidence-based Medicine Level of evidence Grades of recommendations (may 2001). Oxford Center 2007 November 23 [cited 2007 Nov 23];Available from: URL: http://www.cebm.net/index.aspx?o=1025

(12) Altman DG. Some common problems in medical research. In: Altman DG, editor. Practical statistics for medical research.London: Chapman & Hall; 1999. p. 403-9.

(13) Altman DG. Relation between two continuous variables. In: Altman DG, editor. Practical statistics for medical research.London: Chapman & Hall; 1999. p. 277-99.

(14) Irwig L, Tosteson AN, Gatsonis C, Lau J, Colditz G, Chalmers TC, et al. Guidelines for meta-analyses evaluating diagnostic tests. Ann Intern Med 1994 Apr 15;120(8):667-76.

(15) Sutton AJ. Randon effects methods for combining study estimates. In: Sutton AJ AKJDSTSF, editor. Methods for Meta-Analysis in Medical Research. 1 ed. Chichester: John Wiley; 2000. p. 73-86.

(16) Deeks JJ. Systematic reviews of evaluation of diagnostic and screening tests. In: Egger M SGADG, editor. Systematic Reviews in Health care: Meta-analysis in context. 2 ed. London: BMJ Publishing; 2001. p. 248-82.

(17) Deekes JJ ADBM. Statistical methods for examining heterogeneity and combining results from several studies in meta-analysis. In: Egger M SGAD, editor. Systematic Reviews in Health care: Meta-analysis in context.London: BMJ Publishing; 2001. p. 285-312.

(18) Zamora J AV, Muriel A, Khan KS, Coomarasamy A. Meta-DiSc: a software for meta-analysis of test accuracy data. BMC Medical Research Methodology 20066:31.

(19) Chou CY, Chang CH, Yao BL, Kuo HC. Color Doppler ultrasonography and serum CA 125 in the differentiation of benign and malignant ovarian tumors. J Clin Ultrasound 1994 Oct;22(8):491-6.

(20) Emoto M, Udo T, Obama H, Eguchi F, Hachisuga T, Kawarabayashi T. The blood flow characteristics in borderline ovarian tumors based on both color Doppler ultrasound and histopathological analyses. Gynecol Oncol 1998 Sep;70(3):351-7.

(21) Guerriero S, Alcazar JL, Ajossa S, Lai MP, Errasti T, Mallarini G, et al. Comparison of conventional color Doppler imaging and power doppler imaging for the diagnosis of ovarian cancer: results of a European study. Gynecol Oncol 2001 Nov;83(2):299-304.

(22) Guerriero S, Alcazar JL, Coccia ME, Ajossa S, Scarselli G, Boi M, et al. Complex pelvic mass as a target of evaluation of vessel distribution by color Doppler sonography for the diagnosis of adnexal malignancies: results of a multicenter European study. J Ultrasound Med 2002 Oct;21(10):1105-11.

(23) Itakura T, Kikkawa F, Kajiyama H, Mitsui T, Kawai M, Mizutani S. Doppler flow and arterial location in ovarian tumors. Int J Gynaecol Obstet 2003 Dec;83(3):277-83.

(24) Kurjak A, Zalud I, Jurkovic D, Alfirevic Z, Miljan M. Transvaginal color Doppler for the assessment of pelvic circulation. Acta Obstet Gynecol Scand 1989;68(2):131-5.

(25) Kurjak A, Schulman H, Sosic A, Zalud I, Shalan H. Transvaginal ultrasound, color flow, and Doppler waveform of the postmenopausal adnexal mass. Obstet Gynecol 1992 Dec;80(6):917-21.

(26) Kurjak A, Predanic M. New scoring system for prediction of ovarian malignancy based on transvaginal color Doppler sonography. J Ultrasound Med 1992 Dec;11(12):631-8.

(27) Predanic M, Vlahos N, Pennisi JA, Moukhtar M, Aleem FA. Color and pulsed Doppler sonography, gray-scale imaging, and serum CA 125 in the assessment of adnexal disease. Obstet Gynecol 1996 Aug;88(2):283-8.

(28) Tepper R, Lerner-Geva L, Altaras MM, Goldberger S, Ben-Baruch G, Markov S, et al. Transvaginal color flow imaging in the diagnosis of ovarian tumors. J Ultrasound Med 1995 Oct;14(10):731-4.

(29) Weiner Z, Thaler I, Beck D, Rottem S, Deutsch M, Brandes JM. Differentiating malignant from benign ovarian tumors with transvaginal color flow imaging. Obstet Gynecol 1992 Feb;79(2):159-62.

(30) Wu CC, Lee CN, Chen TM, Shyu MK, Hsieh CY, Chen HY, et al. Incremental angiogenesis assessed by color Doppler ultrasound in the tumorigenesis of ovarian neoplasms. Cancer 1994 Feb 15;73(4):1251-6.

(31) Kurtz AB, Tsimikas JV, Tempany CM, Hamper UM, Arger PH, Bree RL, et al. Diagnosis and staging of ovarian cancer: comparative values of Doppler and conventional US, CT, and MR imaging correlated with surgery and histopathologic analysis-report of the Radiology Diagnostic Oncology Group. Radiology 1999 Jul;212(1):19-27.

(32) Valentin L. Gray scale sonography, subjective evaluation of the color Doppler image and measurement of blood flow velocity for distinguishing benign and malignant tumors of suspected adnexal origin. Eur J Obstet Gynecol Reprod Biol 1997 Mar;72(1):63-72.

(33) Kurjak A, Shalan H, Kupesic S, Predanic M, Zalud I, Breyer B, et al. Transvaginal color Doppler sonography in the assessment of pelvic tumor vascularity. Ultrasound Obstet Gynecol 1993 Mar 1;3(2):137-54.

(34) Brown DL, Doubilet PM, Miller FH, Frates MC, Laing FC, DiSalvo DN, et al. Benign and malignant ovarian masses: selection of the most discriminating gray-scale and Doppler sonographic features. Radiology 1998 Jul;208(1):103-10.

Table 1. Characteristics of included studies of trasvaginal with Doppler in diagnosis of ovarian tumor in systematic review.

Study, Year	Mean age (SD)	Period of study	n	Scoring of study quality[8]	Oxford evidence level[14]
Chou *et al.*, 1994[19]	Not reported	January 1991 to February 1993	108	Small population, verification complete, nonblinded, consecutive, prospective, test details sufficient, reference test details sufficient, population details insufficient	3B
Emoto *et al.*, 1997[20]	70.6% premenopausal and 29,4% posmenopausal	April 1993 to August 1996	143	Small population, verification complete, nonblinded, consecutive, prospective, test details sufficient, reference test details sufficient, population details insufficient	2B
Guerriero *et al.*, 2001[21]	43 (SD 15)	December 1997 to December 1999	347	Small population, verification complete, nonblinded, consecutive, prospective, test details sufficient, reference test details sufficient, population details insufficient	2B
Guerriero *et al.*, 2002[22]	Not reported	April 1997 to July 2000	826	Small population, verification complete, nonblinded, consecutive, prospective, test details sufficient, reference test details sufficient, population details insufficient	3B
Itakura *et al.*, 2003[23]	51.6 years	June 1998 to July 2000	95	Small population, verification complete, nonblinded, consecutive, prospective, test details sufficient, reference test details sufficient, population details insufficient	2B
Kurjak *et al.*, 1989[24]	Not reported	Not reported	56	Small population, verification complete, nonblinded, not reported if consecutive and prospective, test details sufficient, reference test details sufficient, population details insufficient	3B
Kurjak *et al.*, 1992[25]	48 years (mean)	September 1990 to September 1991	174	Small population, verification complete, nonblinded, consecutive, prospective, test details sufficient, reference test details sufficient, population details insufficient	2B
Kurjak *et al.*, 1992[26]	Not reported	Not reported	83	Small population, verification complete, nonblinded, not reported if consecutive and prospective, test details sufficient, reference test details sufficient, population details insufficient	3B
Predanic *et al.*,1996[27]	Not reported	Not reported	83	Small population, verification complete, nonblinded, not reported if consecutive and prospective, test details sufficient, reference test details sufficient, population details insufficient	3B
Tepper *et al.*,1995[28]	43 years (SD 16)	1990-1993	217	Small population, verification complete, nonblinded, consecutive, prospective, test details sufficient, reference test details sufficient, population details insufficient	2B
Weiner *et al.*, 1992[29]	Range 20-69 years	Not reported	53	Small population, verification complete, nonblinded, not reported if consecutive and prospective, test details sufficient, reference test details sufficient, population details insufficient	3B
Wu *et al.*, 1994[30]	Range 11-72 years	July 1990 to February 1993	228	Small population, verification complete, nonblinded, consecutive, prospective, test details sufficient, reference test details sufficient, population details insufficient	2B

Table 2. Distribution of histological types in ovarian tissue according to diagnosis in paraffin-embedded block

Study, year	Normal or benign	Borderline	Ovarian cancer	Total N
Chou *et al.*, 1994[19]	83	–	25	108
Emoto *et al.*, 1998[20]	100	12	31	143
Guerriero *et al.*, 2001[21]	258	19	70	347
Guerriero *et al.*, 2002[22]	679	19	128	826
Itakura *et al.*, 2003[23]	64	3	28	95
Kurjak *et al.*, 1989[24]	33	–	8	41
Kurjak *et al.*, 1992[25]	136	–	38	174
Kurjak *et al.*, 1992[26]	54	–	29	83
Predanic *et al.* 1996[27]	76	0	7	83
Teper *et al.*, 1995[28]	165	14	38	217
Weiner *et al.*, 1992[29]	36	3	14	53
Wu *et al.*, 1994[30]	152	4	72	228
Total	**1,836**	**74**	**488**	**2,398**

Table 3. Contingency tables malignant or borderline ovarian cancer *vs* benign ovarian tumor

Malignant or borderline versus benign ovarian tumors				
Study	**False positive**	**False negative**	**True negative**	**True positive**
Chou *et al.*, 1994[19]	0	1	83	24
Emoto *et al.*, 1998[20]	4	4	96	39
Guerriero *et al.*, 2001[21a]	12	29	246	41
Guerriero *et al.*, 2002[22]	108	6	571	141
Itakura *et al.*, 2003[23]	16	4	48	27
Kurjak *et al.*, 1989[24]	0	1	33	7
Kurjak *et al.*, 1992[25]	0	1	136	37
Kurjak *et al.*, 1992[26]	19	2	35	27
Predanic *et al.*, 1996[27]	11	0	65	7
Teper *et al.*, 1995[28]	0	15	165	37
Weiner *et al.*, 1992[29]	1	1	35	16
Wu *et al.*, 1994[30]	29	4	123	72

[a] The difference on total numbers on tables 1 and 3 are due to exclusion of 19 borderline cases in the analysis of IR in transvaginal probe sonography with Doppler.

Table 4. Accuracy of transvaginal ultrasound with Doppler for benign *vs.* borderline or malignant ovarian tumor

Malignant or borderline vs benign ovarian tumors				
Study	**N**	**Sensitivity (95% CI)**	**Specificity (95% CI)**	**DOR (95%CI)**
Chou *et al.*, 1994[19]	108	0.96 (0.8-1)	1 (0.95-1)	2727 (107-69,102)
Emoto et al, 1997[20]	143	0.90 (0.7-0.9)	0.96 (0.9-0.98)	234 (55.7-982)
Guerreiro *et al.*2001,[21]	347	0.58 (0.4-0.7)	0.95 (0.9-0.97)	28.9 (13.6-61)
Guerreiro *et al.*,2002 [22]	826	0.95 (0.9-0.98)	0.84 (0.8-0.86)	124.2 (53.5-288)
Itakura *et al.*2003, [23]	95	0.87 (0.7-0.96)	0.75 (0.62-0.85)	20.2 (6.1-66.7)
Kurjak *et al.*,1989[24]	56	0.87 (0.4-1)	1 (0.89-1)	335 (12.3-9,056.4)
Kurkak et al, 1992[25]	174	0.93 (0.7-1)	0.74 (0.62-0.83)	38.3 (8.3-176)
Kurjak *et al.*, 1992[26]	83	0.97 (0.8-1)	1 (0.97-1)	6825 (272.4-170,979)
Predanic *et al.*, 1996[27]	83	1 (0.59-1)	0.85 (0.76-0.92)	85.4 (4.5-1,600)
Tepper *et al.*, 1995[28]	217	0.7 (0.56-0.82)	1 (0.9-1)	800.8 (46.8-13,684)
Weiner et al, 1992[29]	53	0.94 (0.7-1)	0.97 (0.85-1)	560 (32.9-9,529.6)
Wu *et al.*, 1994[30]	222	0.94 (0.87-0.98)	0.8 (0.73-0.86)	76.3 (25.7-225.9)
Total	**2,400**	**0.87 (0.84-0.9)**	**0.9 (0.87-0.9)**	**125.43 (55.4-283.9)**

Table 5. Calculating a summary ROC curve by method of Moses and Littenberger[14]

Parameter	Estimate	SE	T	P value
a	4.717	0.466	10.127	0.00
b	-0.3284	0.228	1.685	0.13

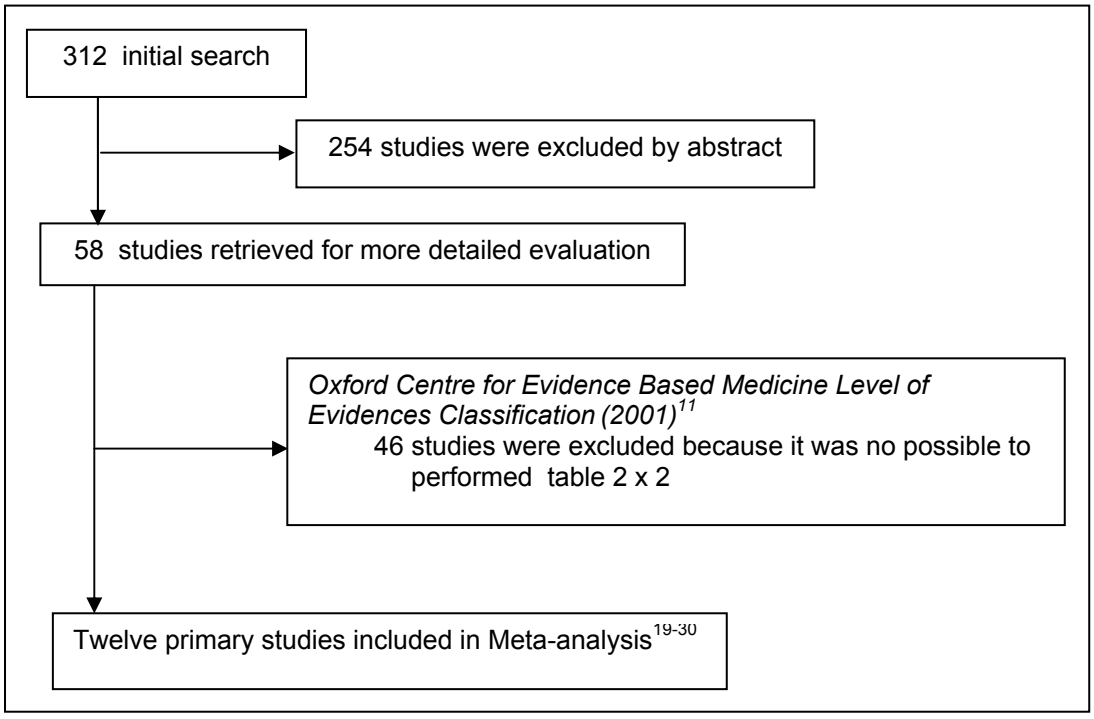


Figure 1. Study selection process.

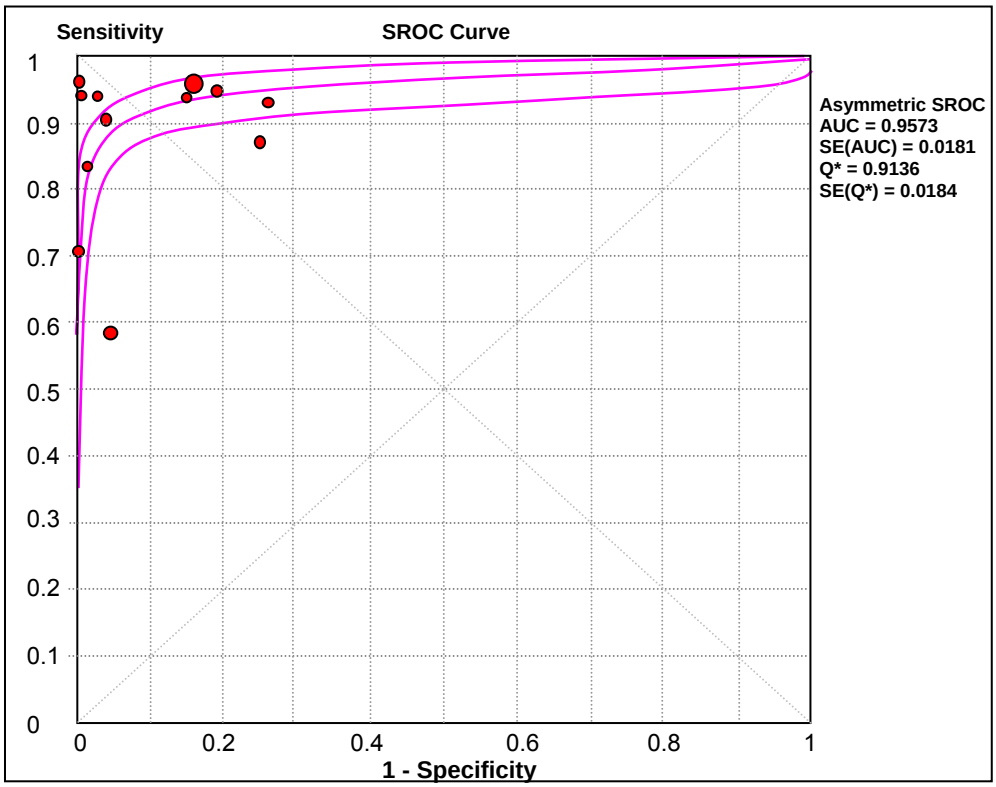


Figure 2. SROC plot of ultrasonography with Doppler for detecting ovarian cancer, with 95%CI.

5.4 Artigo 4

ACCURACY OF CA 125 IN OVARIAN TUMOR: A SYSTEMATIC QUANTITATIVE REVIEW

International Journal of Gynecological Cancer
Manuscript ID: IJGC-2008-00048
Title: Accuracy of CA 125 in the diagnosis of ovarian tumors: a systematic quantitative review
Authors: Medeiros, Lídia
Rosa, Daniela D.
Rosa, Maria
Bozzetti, Mary
Date Submitted: 19-Jan-2008

Accuracy of CA 125 in Ovarian Tumor: A Systematic Quantitative Review

Author: Lidia Rosi Medeiros, MD, MSc(1)
Address: José de Alencar 1244 apt 1009, Porto Alegre, RS, Brasil, CEP 90880-480
lidia.rosi@terra.com.br

Co-authors
Daniela Dornelles Rosa, MD, PhD(2,3)
Maria Inês da Rosa, MD, PhD(1, 4)
Mary Clarisse Bozzetti, MD, PhD(1, 3, 5)

Affiliations of all authors:
[1] Postgraduate Program in Epidemiology at Federal University of Rio Grande do Sul, Porto Alegre Brazil
[2] Hospital Fêmina and Hospital Moinhos deVento, Porto Alegre, Brazil
[3] Postgraduate Program in Medicine: Medical Sciences at Federal University of Rio Grande do Sul, Porto Alegre, Brazil
[4] Medical School at University of Extremo Sul Catarinense, Criciúma, Brazil
[5] Department of Social Medicine, Faculty of Medicine at Federal University of Rio Grande do Sul, Porto Alegre Brazil

RESUMO

Foi realizada uma revisão sistemática quantitativa para estimar a acurácia do CA 125 nas tumorações ovarianas, tendo sido incluídos estudos que avaliaram os níveis séricos de CA 125 em pacientes com tumoração ovariana e foram comparados com os resultados finais do exames anatomopatógicos incluídos nesta revisão. Os 17 estudos analisados compreenderam 2.374 mulheres. No cálculo que sumariza os resultados da sensibilidade de todos os estudos, encontrou-se um valor de 0,80 (IC95% de 0,76-0,82); na especificidade esse valor foi de 0,75 (IC 95% de 0,73-0,77). O valor da DOR para diagnóstico de câncer ovariano ou de lesões limítrofes *vs* tumorações benignas foi de 21,2 (IC 95% de 12-37). O sumário da curva ROC foi contruído devido à heterogeneidade da DOR. Para diagnóstico de câncer ovariano ou de lesões limítrofes *vs* tumorações benignas, a área sob a curva (AUC) foi de 0,8877. Em conclusão, a avaliação dos níveis séricos do CA 125 ≥ 35 U/ml é importante, nas tumorações ovarianas, para predizer se a lesão ovariana é de natureza benigna ou maligna.

Palavras-chave: CA 125, tumor de ovário**,** revisão sistemática, metanálise, acurácia diagnóstica.

ABSTRACT

A quantitative systematic review was performed to estimate the accuracy of CA 125 assay in the diagnosis of ovarian tumors. Studies that evaluated CA 125 levels for the diagnosis of ovarian tumors and compared it to paraffin-embedded sections as the diagnostic standard were included. Seventeen studies were analyzed, which included 2,374 women. The pooled sensitivity was 0.80 (IC95% 0.76-0.82) and the specificity was 0.75 (IC 95% 0.73-0.77). The DOR for ovarian cancer and borderline lesions *vs* benign lesions was 21.2 (95% CI, 12-37). SROC curves were constructed due to heterogeneity in the DOR. For malignant ovarian tumors and borderline *vs*. benign lesions the AUC was 0.8877. In conclusion, a CA 125 level ≥ 35 U/ml is a useful pre-operative test for predicting the benign or malignant nature of pelvic mass.

Keywords: CA 125, ovarian tumor**,** systematic review, meta-analysis, accuracy, diagnosis.

INTRODUCTION

Malignant ovarian neoplasms are responsible for 4% of all cancers affecting women. They are the second most common cause of death from gynecological cancer and the fourth most common cause of death from all types of cancer affecting women[1]. Benign, borderline and malignant lesions have been identified within the same surgical specimen[2]. However, the frequency and velocity of progression from dysplasia into cancer remain unknown[3].

The diagnosis of ovarian neoplasms is a common problem in the clinical practice. Although the majority of adnexal masses are benign, the main aim of the diagnostic evaluation is to exclude or to confirm the diagnosis of malignancy. The methods used in the evaluation of woman with suspected adnexal mass are physical examination, ultrasound and serum CA 125 levels determination[4]. CA 125 is an antigenic determinant on a high molecular-weight glycoprotein recognized by a monoclonal antibody (OC 125). This is expressed by epithelial ovarian tumors as well as by other tissues of Müllerian origin (peritoneum, pleura and pericardium)[5]. A CA 125 level ≥ 35 U/ml is considered suspicious for malignancy[4]. In general, CA 125 is elevated in 80-85% of women with epithelial ovarian cancer and in 65% of patients with mucinous carcinoma of the ovary[5]. The diagnostic accuracy of ovarian tumors is significantly improved by combining transvaginal ultrasound with Doppler and CA 125 findings. Using these methods together an erroneous diagnosis may occur in only 6% of cases[6].

We undertook a systematic quantitative review of the literature to ascertain the accuracy of CA 125 levels in the diagnosis of ovarian cancer and to explore the reasons for the ongoing controversies about this issue.

METHODS

Identification of studies

A comprehensive search of the MEDLINE, CANCERLIT, LILACS and EMBASE databases was perfomed from January 1990 to December 2007. The medical subjects heading (MeSH) and text words for the terms "*ovarian neoplasm"* and "*CA 125*" were combined with the MeSH term *diagnosis* ("sensitivity and specificity"). The search was limited to human studies but had no language

restrictions. In addition, the Cochrane Library was searched. Reference lists of all available primary studies were reviewed to identify additional relevant citations. The authors who published the studies were not contacted.

Selection criteria

This review focused on observational studies evaluating clinically suspected adnexal masses through the evaluation of CA 125 levels. Results of the diagnostic test of interest were compared with the results of a reference standard in all studies. The case were defined as women treated surgically for ovarian tumors. The diagnostic test was CA 125 levels with a cutoff of 35 U/ml and the diagnostic reference was the result of histological analysis of standard paraffin-embedded sections. A CA 125 levels diagnosis was considered correct if it did not differ from that of the paraffin section. For inclusion criteria in this systematic review, the final histological assessment of paraffin-embedded sections should have been designated each case as benign, borderline or malignant. We excluded studies that compared the accuracy of CA 125 levels only for borderline or malignant ovarian tumors, of for non ovarian tumors, as well as studies that lacked data to construct 2 x 2 contingency tables.

Primary outcome measure was the accuracy of diagnosis of benign, borderline and malignant ovarian tumors according to CA 125 levels. The secondary outcome was the distribution of histological types of ovarian tumors according to paraffin-embedded tissue diagnosis. The studies were reviewed independently by three investigators (L.R.M, D.D.R, M.I.R). Final inclusion and exclusion criteria were defined with reference to a selection criteria checklist. Disagreements on study inclusion or exclusion were initially solved by consensus, and when this was not possible, they were solved by arbitration by a fourth reviewer (M.C.B). The agreement statistics among reviewers were computed.

Quality assessment

All articles meeting the eligibility criteria were assessed for their methodological quality. The assessment involved scrutinizing the study designs and the relevant features of the patient population, the diagnostic test and the reference standard[7-10]. Relevant features included the methods of data collection and patient selection, description of the CA 125 levels, description of the histological reference

standard, and presence of verification bias[8, 9]. The quality assessment results for the included studies are summarized in the "*Scoring of Study Quality*" column of table 1[10]. Studies were further assessed for methodological quality with reference to the *Oxford Centre for Evidence Based Medicine Level of Evidences Classification* rubric. Only studies with Oxford Evidence Levels 1 to 3 were considered to be of high quality while those with levels 4 and 5 were excluded[11].

Data abstraction

Three investigators (L.R.M, D.D.R, M.I.R) independently abstracted data from the primary studies regarding the prevalence of benign, borderline and malignant ovarian lesions as well as the sensitivities, specificities, true-positive (TPR; sensitivity) and the false-positive rates (FPR; 1 – specificity) of CA 125 levels. The assessment of English-language articles was performed by three reviewers (L.R.M, D.D.R, M.I.R) while assessment of those articles published in languages other than English were performed independently by another reviewers (M.C.B) following translation (when necessary). Any disagreement was resolved by consensus for both English and non-English studies. Three possibilities of ovarian tumor diagnosis were considered: benign, borderline and malignant. Data were abstracted as 2 x 2 tables relating CA 125 levels with paraffin-embedded tissue diagnosis regarding malignant or borderline disease *versus* benign lesions section (Table 2).

Data synthesis and statistical analysis

To evaluate the agreement between study eligibility and assessment of methodological quality, as well as the agreement between the results of CA 125 and those from paraffin-embedded section analysis, the observed percentage of agreement and the κ coefficient for inter-rater reliability were calculated[12, 13]. For each study, 2 x 2 contingency tables were constructed, in which all biopsies were classified as normal or benign, borderline and malignant. The true-positive (TPR; sensitivity) and the false-positive rates (FPR; 1 – specificity) were calculated. When 2 x 2 tables had 0 cells, the value of 0.5 was added to each cell to enable calculations; when a study contained two cells with the 0 value, it was excluded from the analysis.

The summary weighted sensitivity and specificity were calculated as the sum of sensitivities and specificities reported for each study, multiplied by the number of subjects in the study and divided by the total number of subjects in all studies. The 95% CIs for mean weighted results were calculated using the exact method. The association between sensitivity and specificity for benign and borderline or malignant ovarian lesions was calculated using the *Spearman's correlation coefficient test*[13]. When there was no correlation, pooling sensitivities and specificities were calculated, since there were two categories of results (negative or positive test)[14]. In the case of correlation or heterogeneity between sensitivity and specificity, a summary receiver operating characteristic curve (SROC) was generated using data from all thresholds, by the Littenberg and Moses method[14, 15] .

The diagnostic odds ratio (DOR) can relate to different combinations of sensitivity and specificity. The DOR describes the odds of the positive test results in participants with disease compared with the odds of positive test results in those without disease. A single diagnostic odds ratio corresponds to a set of sensitivities and specificities depicted by SROC. It can change according to the threshold and ROC curve used to define an abnormal examination resulting in the expected trade-off between sensitivity and specificity. Also, the area under the curve (AUC) can summarize the inherent capacity of a test for discriminating a diseased from a non-diseased subject. Perfect tests usually have AUCs close to 1 and poor tests usually have AUCs close to 0.5[14, 15]. The heterogeneity of the sensitivities and specificities of the studies were assessed using the Q_T (Cochran) test for χ^2 distributions with N -1 degrees of freedom. Since sensitivity and specificity were heterogeneous, a random effect model was used and the terms were pooled with a 95% CI[16, 17]. The statistical analysis was performed with the software Meta-DiSc® (version Beta 1.1.1)[18].

Sensitivity analysis

To assess whether study quality affected the diagnostic accuracy of frozen section analysis we excluded those studies that met less than 50% of the quality criteria and those that were sub-level 3 by *Oxford Centre for Evidence Based Medicine Level of Evidences Classification*[11, 16] .To analyze publication bias, inverted funnel plots of individual study logarithmic *odds ratio* (OR) were plotted against sample size. The robustness of the results was tested by repeating the analysis with a different statistical model (random effects model)[16, 17].

RESULTS

Study identification and eligibility

The process of study selection is summarized in figure 1. Our initial search identified 785 potentially relevant articles. We excluded 757 published studies after reviewing their titles and abstracts, since that they did not relate to the subject under review. Twenty-eight full-text articles were retrieved; 11 were excluded after further scrutiny. A complete list of excluded studies is available from the authors. Seventeen primary studies, including 2,374 women, met the criteria for inclusion and were analyzed (Table 1)[6, 19-34]. Interrater overall agreement for study eligibility and methodological quality was 94% (κ = 0.82), indicating excellent agreement[13]. Disagreements among reviewers occurred during analysis of the 28 retrieved studies and were related to inclusion or exclusion criteria (characteristics and scoring of study). The disagreement, was solved by consensus. The difference on total numbers on tables 1 and 3 are due to exclusion of cases in which CA 125 levels were uncertain.

Studies description

Details of the participants, interventions, and quality assessments of the studies selected for the meta-analysis are summarized in table 1. The mean age of participants was reported in seven studies[6, 19, 22-24, 27, 33]. Three studies were non-blinded and prospective with a small population, but included sufficient experimental details, proper diagnostic tests and diagnostic reference standards[22, 23, 25]. Of the 17 included studies, three[22, 23, 25] had high methodological quality, satisfying ≥ 55% of the criteria for study quality and with an Oxford Evidence Level of 2B. However, 14 studies[6, 19-21, 24, 26-34] were classified as level 3B, because details of the patients, including age, were not reported and also because they met < 50% of the quality criteria.

The distribution of histological diagnosis according paraffin-embedded blocks analysis was retrieved from seventeen trials[6, 19-34] (table 2). Normal or benign lesions were found in 1,695 patients (71.3%), borderline in 73 patients (3.07%), and ovarian cancer in 606 patients (25.5%). Table 3 shows the contingency tables (FN, FP, TP and TN) for each study.

The association between sensitivity and specificity for benign and borderline or malignant ovarian tumors was calculated with the *Spearman's correlation coefficient test*[13]. The correlations coefficient was 0.343, p = 0.178. When there was no correlation, pooling sensitivities and specificities were calculated, since there were two categories of results (negative or positive test)[14]. Since sensitivity and specificity were heterogeneous, a random effect model was used and the terms were pooled with a 95% CI[16, 17]. Interrater overall agreement between CA 125 level and paraffin sections was 77% (k = 0.51), indicating moderate agreement (Table 3)[12].

Sensitivity

CA 125 was better at detecting ovarian cancer or borderline lesions when compared with benign lesions with a pooled sensitivity was 0.80 (IC95% 0.76-0.82) (table 4). The estimates for heterogeneity were highly consistent across studies: for normal or benign lesion *vs*. borderline or malignant lesions (Q_T = 121, p = 0.001; inconsistency I^2 = 89%) (table 4).

Specificity

In general, specificity was lower for normal or benign lesion *vs*. borderline or malignant ovarian tumors with a pooled specificity was 0.75 (IC 95% 0.73-0.77). The estimates for heterogeneity were highly consistent across studies: for normal or benign lesion *vs*. borderline or malignant lesions (Q_T = 409, p = 0.001, inconsistency I^2 = 96%) (table 4).

DOR

The DOR between borderline and malignant ovarian cancer *vs*. benign lesions was 21.2 (IC95% 12-37; Q_T = 57, p = 0.001; inconsistency I^2 = 72.4%) (table 4). SROC curves were constructed due to heterogeneity in the DOR[15]. For malignant ovarian tumors or borderline *vs*. benign lesions the AUC was 0.8877. The AUC for the ROC curve was estimated by a trapezoidal rule (Meta-DiSc®; version Beta 1.1.1)[18] (Figure 2). An alternative summarized ROC curve allowing for variation in the DOR with a diagnostic threshold may be estimated by the method of Moses and Littenberg. First, the logarithm of the diagnostic odds ratio (D) in the measure of the diagnostic threshold (*S*), weighting by study size, produce an estimate of the parameters *a* and *b* from the regression equation, D = *a*+*bS* (table 5). The results

suggested that the DOR not change with the threshold (p = 0.23). ROC curves, including meta-analysis SROC curves, do not specific the best operating point (the threshold and its associated TRP and FRP)[14-17].

Sensitivity Analysis

The robustness of the results was tested by repeating the analysis using a different statistical model (random effects model)[16, 17]. Pooling sensitivity, specificity and the DOR from the three studies with high methodological quality did not alter the accuracy rate of CA 125 levels analysis in benign, borderline or malignant ovarian tumors[27, 30, 31]. Therefore, all 17 selected[6, 19-34] studies were included in the sensitivity analysis. Inverted funnel plots showed asymmetry.

DISCUSSION

In summary, this systematic review showed that CA 125 ≥ 35 U/ml may detect malignant or borderline ovarian tumors. However, CA 125 levels may be also elevated in benign lesions: endometriosis, uterine miomas, acute and chronic salpingites, pelvic inflammatory disease. Therefore, the sensitivity for detecting malignant ovarian tumors using CA 125 levels was 80%, but the specificity was 75%. Due to the fact that the DOR was 21.2, that CA 125 is a useful pre-operative test for predicting a benign or a malignant nature of a pelvic mass.

Combined data from 15 different studies and showed that CA 125 levels were increased in 50% of cases of patients with ovarian cancer FIGO stage I, in 90% of stage II, in 92% of stage III and in 94% of stage IV disease[35]. Over the past two decades, the value of CA 125 has been tested in monitoring treatment, detecting recurrent disease, and screening for early stage ovarian cancer. The CA 125 has the ability to distinguish malignant from benign pelvic masses, mainly in the preoperative assessment of postmenopausal women[35]. The ability to predict whether a tumor is malignant or benign before surgery is important to decide between laparoscopy and laparotomy. Potential screening procedures for the diagnosis of ovarian cancer include transvaginal Doppler ultrasound, and serum CA 125 levels. The use of CA 125 levels alone, has a number several serious limitations. The sensitivity of serum CA 125 measurement for the diagnosis of ovarian cancer is limited by several factors including tumour expression CA 125, tumor size and stage of disease.

A possible limitation of this systematic review is a potential bias regarding the retrospective nature of all analyzed trials and the lack of blinding in their assessment. Other sources for biases are the small number of trials included and the lower quality of some of the trials (nonconsecutive, insufficient population details)[9, 10]. On the other hand, this meta-analysis complied with the criteria for performing a rigorous systematic review planned a priori[14-17]. These included the use of study quality assessment tools[10] and investigating homogeneity with fixed and random models to test the robustness of the results[14-16]. Moreover, SROC curves were constructed in cases of heterogeneity in the DOR[15] with calculus of the AUC[14-18]. Thus, we suggest that blinded prospective studies should be performed to investigate the diagnostic accuracy of CA 125 analysis in the diagnosis of ovarian tumors associating another exam transvaginal Doppler ultrasound.

This quantitative review provides precise evaluation of the accuracy of CA 125 in the diagnosis of benign, borderline and malignant ovarian tumors. At present, CA 125 is the best available marker for epithelial cancer, although it lacks sensitivity for stage I disease and lacks specificity for the early diagnosis of ovarian cancer. CA 125 levels alone do not provide good sensitivity or specificity for the distinction of malignant from benign pelvic masses. An accurate preoperative prediction of the origin and character of pelvic tumors is essential for optimal pre and intraoperative surgical management, which may provide, in the case of malignancy, the best chance for a long disease-free interval or for cure. Better methods to improve the accuracy rates for the diagnosis of borderline ovarian tumor are needed.

Contributors

LRM had full access to all data and take responsibility for the integrity and the accuracy of the analysis. DDR, MIR and MCB organized the study concept and the design. LRM acquired, analyzed and interpreted the data. LRM, DDR, MIR and MCB drafted the manuscript. LRM and DDR performed the statistical analysis.

Conflict of interest statement

We declare no conflict of interest

Reference List

(1) Yancik R. Ovarian cancer. Age contrasts in incidence, histology, disease stage at diagnosis, and mortality. Cancer 1993 Jan 15;71(2 Suppl):517-23.

(2) Scully RE. Early de novo ovarian cancer and cancer developing in benign ovarian lesions. Int J Gynaecol Obstet 1995 Jul;49 Suppl:S9-15.

(3) Scully RE. Influence of origin of ovarian cancer on efficacy of screening. Lancet 2000 Mar 25;355(9209):1028-9.

(4) Alcázar JL, Errasti T, Zornosa A, Mínguez JA, Galán MJ. Transvaginal color Doppler ultrasonogrphy and CA 125 in suspicious adnexal masses. Int J Gynecol Obstet 1999;66:255-61.

(5) Jacobs I, Bast RC, Jr. The CA 125 tumour-associated antigen: a review of the literature. Hum Reprod 1989 Jan;4(1):1-12.

(6) Maggino T, Sopracordevole F, Matarese M, Di PC, Tambuscio G. CA-125 serum level in the diagnosis of pelvic masses: comparison with other methods. Eur J Gynaecol Oncol 1987;8(6):590-5.

(7) Cochrane Methods Working Group on Systematic Reviews of Screening and Diagnostic Tests. Recommended methods. Cochrane Library 2004 January 25 [cited 2004 Jan 25]; Available from: URL: http://www.cochrane.org/cochrane/sadtdoc1.htm

(8) Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD Initiative. Ann Intern Med 2003 Jan 7;138(1):40-4.

(9) Jaeschke R, Guyatt G, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. A. Are the results of the study valid? Evidence-Based Medicine Working Group. JAMA 1994 Feb 2;271(5):389-91.

(10) Lijmer JG, Mol BW, Heisterkamp S, Bonsel GJ, Prins MH, van der Meulen JH, et al. Empirical evidence of design-related bias in studies of diagnostic tests. JAMA 1999 Sep 15;282(11):1061-6.

(11) Phillips B. Oxford Centre for evidence-based Medicine Level of evidence Grades of recommendations (may 2001). Oxford Center 2007 November 23 [cited 2007 Nov 23]; Available from: URL: http://www.cebm.net/index.aspx?o=1025

(12) Altman DG. Some common problems in medical research. In: Altman DG, editor. Practical statistics for medical research.London: Chapman & Hall; 1999. p. 403-9.

(13) Altman DG. Relation between two continuous variables. In: Altman DG, editor. Practical statistics for medical research.London: Chapman & Hall; 1999. p. 277-99.

(14) Deeks JJ. Systematic reviews of evaluation of diagnostic and screening tests. In: Egger M SGADG, editor. Systematic Reviews in Health care: Meta-analysis in context. 2 ed. London: BMJ Publishing; 2001. p. 248-82.

(15) Irwig L, Tosteson AN, Gatsonis C, Lau J, Colditz G, Chalmers TC, et al. Guidelines for meta-analyses evaluating diagnostic tests. Ann Intern Med 1994 Apr 15;120(8):667-76.

(16) Deekes JJ ADBM. Statistical methods for examining heterogeneity and combining results from several studies in meta-analysis. In: Egger M SGAD, editor. Systematic Reviews in Health care: Meta-analysis in context.London: BMJ Publishing; 2001. p. 285-312.

(17) Sutton AJ. Randon effects methods for combining study estimates. In: Sutton AJ AKJDSTSF, editor. Methods for Meta-Analysis in Medical Research. 1 ed. Chichester: John Wiley; 2000. p. 73-86.

(18) Zamora J AV, Muriel A, Khan KS, Coomarasamy A. Meta-DiSc: a software for meta-analysis of test accuracy data. BMC Medical Research Methodology 2006:31.

(19) Chen DX, Schwartz PE, Li XG, Yang Z. Evaluation of CA 125 levels in differentiating malignant from benign tumors in patients with pelvic masses. Obstet Gynecol 1988 Jul;72(1):23-7.

(20) Einhorn N, Bast RC, Jr., Knapp RC, Tjernberg B, Zurawski VR, Jr. Preoperative evaluation of serum CA 125 levels in patients with primary epithelial ovarian cancer. Obstet Gynecol 1986 Mar;67(3):414-6.

(21) Einhorn N, Knapp RC, Bast RC, Zurawski VR, Jr. CA 125 assay used in conjunction with CA 15-3 and TAG-72 assays for discrimination between malignant and non-malignant diseases of the ovary. Acta Oncol 1989;28(5):655-7.

(22) Erdogan N, Ozcelik B, Serin IS, Akgun M, Ozturk F. Doppler ultrasound assessment and serum cancer antigen 125 in the diagnosis of ovarian tumors. Int J Gynaecol Obstet 2005 Nov;91(2):146-50.

(23) Guerriero S, Ajossa S, Lai MP, Alcazar JL, Paoletti AM, Marisa O, et al. The diagnosis of functional ovarian cysts using transvaginal ultrasound combined with clinical parameters, CA125 determinations, and color Doppler. Eur J Obstet Gynecol Reprod Biol 2003 Sep 10;110(1):83-8.

(24) Hata K, Hata T, Manabe A, Sugimura K, Kitao M. A critical evaluation of transvaginal Doppler studies, transvaginal sonography, magnetic resonance imaging, and CA 125 in detecting ovarian cancer. Obstet Gynecol 1992 Dec;80(6):922-6.

(25) Maggino T, Gadducci A, D'Addario V, Pecorelli S, Lissoni A, Stella M, et al. Prospective multicenter study on CA 125 in postmenopausal pelvic masses. Gynecol Oncol 1994 Aug;54(2):117-23.

(26) Malkasian GD, Jr., Knapp RC, Lavin PT, Zurawski VR, Jr., Podratz KC, Stanhope CR, et al. Preoperative evaluation of serum CA 125 levels in premenopausal and postmenopausal patients with pelvic masses: discrimination of benign from malignant disease. Am J Obstet Gynecol 1988 Aug;159(2):341-6.

(27) Mancuso A, De VA, Triolo O, Irato S. The role of transvaginal ultrasonography and serum CA 125 assay combined with age and hormonal state in the differential diagnosis of pelvic masses. Eur J Gynaecol Oncol 2004;25(2):207-10.

(28) Mogensen O, Mogensen B, Jakobsen A. CA 125 in the diagnosis of pelvic masses. Eur J Cancer Clin Oncol 1989 Aug;25(8):1187-90.

(29) Predanic M, Vlahos N, Pennisi JA, Moukhtar M, Aleem FA. Color and pulsed Doppler sonography, gray-scale imaging, and serum CA 125 in the assessment of adnexal disease. Obstet Gynecol 1996 Aug;88(2):283-8.

(30) Roman LD, Muderspach LI, Stein SM, Laifer-Narin S, Groshen S, Morrow CP. Pelvic examination, tumor marker level, and gray-scale and Doppler sonography in the prediction of pelvic cancer. Obstet Gynecol 1997 Apr;89(4):493-500.

(31) Soper JT, Hunter VJ, Daly L, Tanner M, Creasman WT, Bast RC, Jr. Preoperative serum tumor-associated antigen levels in women with pelvic masses. Obstet Gynecol 1990 Feb;75(2):249-54.

(32) Vasilev SA, Schlaerth JB, Campeau J, Morrow CP. Serum CA 125 levels in preoperative evaluation of pelvic masses. Obstet Gynecol 1988 May;71(5):751-6.

(33) Weiner Z, Thaler I, Beck D, Rottem S, Deutsch M, Brandes JM. Differentiating malignant from benign ovarian tumors with transvaginal color flow imaging. Obstet Gynecol 1992 Feb;79(2):159-62.

(34) Yedema C, Massuger L, Hilgers J, Servaas J, Poels L, Thomas C, et al. Pre-operative discrimination between benign and malignant ovarian tumors using a combination of CA125 and CA15.3 serum assays. Int J Cancer Suppl 1988;3:61-7.

(35) Duffy MJ, Bonfrer JM, Kulpa J, Rustin GJ, Soletormos G, Torre GC, et al. CA125 in ovarian cancer: European Group on Tumor Markers guidelines for clinical use. Int J Gynecol Cancer 2005 Sep;15(5):679-91.

Table 1. Characteristics of included studies of CA 125 levels in ovarian tumor in systematic review

Study, Year	Mean age (SD)	Period of study	n	Scoring of study quality[8]	Oxford evidence level[14]
Chen *et al.*, 1988[19]	37.4 years (SD 6.5)	Not reported	211	Small population, verification complete, nonblinded, not reported if consecutive and prospective, test details sufficient, reference test details sufficient, population details insufficient	3B
Einhorn *et al.*, 1986[20]	Not reported	1983	100	Small population, verification complete, nonblinded, not reported if consecutive and prospective, test details sufficient, reference test details sufficient, population details insufficient	3B
Einhorn *et al.*, 1989[21]	Not reported	Not reported	219	Small population, verification complete, nonblinded, not reported if consecutive and prospective, test details sufficient, reference test details sufficient, population details insufficient	3B
Erdogan *et al.*, 2005[22]	48,2 years (SD15)	No reported	63	Small population, verification complete, nonblinded, was prospective and consecutive, test details sufficient, reference test details sufficient, population details insufficient	2B
Guerriero *et al.*, 2003[23]	33.4 years (SD 9.1)	From Abril 1997 until December 2002	229	Small population, verification complete, nonblinded, was consecutive and prospective, test details sufficient, reference test details sufficient, population details insufficient	2B
Hata *et al.*, 1992[24]	47.4 years (SD 13)	Not reported	63	Small population, verification complete, nonblinded, not reported if consecutive and prospective, test details sufficient, reference test details sufficient, population details insufficient	3B
Maggino *et al.*, 1987[6]	Premenopausal (SD 12.8 years; range 20-74 years)	Not reported	290	Small population, verification complete, nonblinded, not reported if consecutive and prospective, test details sufficient, reference test details sufficient, population details insufficient	3B
Maggino *et al.*, 1994[25]	Postmenopausal	From March 1991 to March 1992	290	Small population, verification complete, nonblinded, consecutive, prospective, test details sufficient, reference test details sufficient, population details insufficient	2B
Malkasian *et al.*, 1988[26]	Not reported	Not reported	158	Small population, verification complete, nonblinded, not reported if consecutive and prospective, test details sufficient, reference test details sufficient, population details insufficient.	3B

Study, Year	Mean age (SD)	Period of study	n	Scoring of study quality[8]	Oxford evidence level[14]
Mancuso *et al.* 2004[27]	42.2 years (SD 15.2; range 18-82)	Not reported	125	Small population, verification complete, nonblinded, not reported if consecutive and prospective, test details sufficient, reference test details sufficient, population details insufficient	3B
Mogensen *et al.*, 1989[28]	Not reported	Not reported	151	Small population, verification complete, nonblinded, not reported if consecutive and prospective, test details sufficient, reference test details sufficient, population details insufficient	3B
Predanic *et al.*, 1996[29]	Not reported	Not reported	83	Small population, verification complete, nonblinded, not reported if consecutive and prospective, test details sufficient, reference test details sufficient, population details insufficient	3B
Roman *et al.*, 1997[30]	Not reported (181 premenopausal and 45 postmenopausal)	July 1992 until March 1994	223	Small population, verification complete, nonblinded, not consecutive and prospective, test details sufficient, reference test details sufficient, population details insufficient	3B
Soper *et al.*, 1990[31]	Not reported	January 1985 until January 1986	88	Small population, verification complete, nonblinded, not reported if consecutive and prospective, test details sufficient, reference test details sufficient, population details insufficient	3B
Vasilev *et al.*, 1988[32]	Not reported	March 1984 until February 1986	182	Small population, verification complete, nonblinded, not reported consecutive and prospective, test details sufficient, reference test details sufficient, population details insufficient	3B
Weiner *et al.*, 1992[33]	Range 20-69 years	Not reported	53	Small population, verification complete, nonblinded, not reported consecutive and prospective, test details sufficient, reference test details sufficient, population details insufficient	3B
Yedema *et al.*, 1988[34]	Not reported	Not reported	70	Small population, verification complete, nonblinded, not reported if consecutive and prospective, test details sufficient, reference test details sufficient, population details insufficient	3B

Table 2. Distribution of histological types in ovarian tissue according to diagnosis in paraffin-block

Study, year	Benign	Borderline	Ovarian cancer	Total N
Chen *et al.*, 1988[19]	153	-	58	211
Einhorn *et al.*, 1986[20]	77	4	19	100
Einhorn *et al.*, 1989[21]	165	27	27	219
Erdogan *et al.*, 2005[22]	42	-	21	63
Guerriero *et al.*, 2003[23]	206	-	23	229
Hata *et al.*, 1992[24]	36	-	27	63
Maggino *et al.*, 1987[6]	45	2	19	66
Maggino *et al.*, 1994[25]	184	3	103	290
Malkasian *et al.*, 1988[26]	90	3	65	158
Mancuso *et al.*, 2004[27]	111	-	14	125
Mogensen *et al.*, 1989[28]	52	8	91	151
Predanic *et al.*, 1996[29]	76	-	7	83
Roman *et al.*, 1997[30]	180	17	26	223
Soper *et al.*, 1990[31]	46	3	39	88
Vasilev *et al.*, 1988[32]	164	3	15	182
Weiner *et al.*, 1992[33]	36	3	14	53
Yedema *et al.*, 1988[34]	32	-	38	70
Total	**1,695**	**73**	**606**	**2,374**

Table 3. Contingency tables benign *vs* borderline or malignant ovarian tumors

Malignant or borderline versus benign ovarian tumors				
Study	**False positive**	**False negative**	**True negative**	**True positive**
Chen *et al.*, 1988[19]	61	10	92	48
Einhorn *et al.*, 1986[20]	6	4	72	18
Einhorn *et al.*, 1987[21]	26	10	139	44
Erdogan *et al.* 2005[22]	1	11	31	20
Guerriero *et al.*, 2003[23]	111	0	95	23
Hata *et al.*,[24]	11	3	33	16
Maggino *et al.* 1987[6]	4	4	38	20
Maggino *et al.*, 1994[25]	39	23	145	83
Malkasian *et al.*, 1988[26]	1	4	89	64
Mancuso *et al.*, 2004[27]	23	0	88	14
Mogensen *et al.*, 1989[28]	2	6	46	97
Predanic *et al.*, 1996[29]	3	3	73	4
Roman *et al.*,[30]	51	14	129	29
Soper *et al.*, 1990[31]	22	4	24	38
Vasilev *et al.*, 1988[32]	4	36	128	14
Weiner *et al.*, 1992[33]	14	3	22	14
Yedema *et al.*, 1988[34]	11	8	24	27

Table 4. Accuracy of CA 125 for borderline or malignant ovarian tumor *vs* benign ovarian tumor

Malignant or borderline vs benign ovarian tumors				
Study	**N**	**Sensitivity (95% CI)**	**Specificity (95% CI)**	**DOR (95%CI)**
Chen *et al.*, 1988[19]	211	0.82 (0.7-0.9)	0.6 (0.5-0.6)	7.2 (3-15.)
Einhorn *et al.*, 1986[20]	100	0.81 (0.6-0.9)	0.84 (0.7-0.9)	23.5 (10-52)
Einhorn *et al.*, 1987[21]	219	0.78 (0.5-0.9)	0.92 (0.8-0.9)	43.2 (11-157)
Erdogan *et al.*, 2005[22]	63	0.64 (0.4-0.8)	0.96 (0.8-0.9)	56.3 (6-470)
Guerriero *et al.*, 2003[23]	229	1 (0.8-1)	0.46 (0.3-0.5)	9.9 (0.5-168)
Hata *et al.*,[24]	63	0.84 (0.6-0.9)	0.75 (0.6-0.8)	16 (10-210)
Maggino *et al.*, 1987[6]	66	0.83 (0.6-0.9)	0.9 (0.7-0.9)	47.5 (10-210)
Maggino *et al.*, 1994[25]	290	0.78 (0.6-0.8)	0.78 (0.7-0.8)	13.4 (7-24)
Malkasian *et al.*, 1988[26]	158	0.94 (0.8-0.9)	0.98 (0.9-1)	1424 (155-1,304)
Mancuso *et al.*, 2004[27]	125	1 (0.7-1)	0.79 (0.7-0.8)	109 (6-1,898)
Mogensen *et al.*, 1989[28]	151	0.94 (0.8-0.9)	0.95 (0.8-1)	371 (72-1,913)
Predanic *et al.*, 1996[29]	83	0.57 (0.1-0.9)	0.96 (0.8-1)	32.4 (4.8-214)
Roman *et al.*,[30]	223	0.67 (0.5-0.8)	0.71 (0.6-0.7)	5.2 (2.5-10)
Soper *et al.*, 1990[31]	88	0.9 (0.7-0.9)	0.5 (0.3-0.6)	9.5 (2.8-31.1)
Vasilev et al, 1988[32]	182	0.28 (0.1-0.4)	0.97 (0.9-1)	12.4 (3.8-40)
Weiner *et al.*,[33]	53	0.82 (0.5-0.8)	0.61 (0.4-0.7)	7.3 (1.7-30.2)
Yedema *et al.*, 1988[34]	70	0.75 (0.5-0.8)	0.68 (0.5-0.8)	6.5 (2.2-19.1)
Pooled	2,373	0.8 (0.7-0.8)	0.75 (0.73-0.77)	21.2 (2.2-19)

Table 5. Calculating a summary ROC curve by method of Moses and Littenberger [14]

Parameter	Estimate	SE	T	P value
a	3.059	0.324	9.4	0.00
b	-0.203	0.165	1.233	0.23

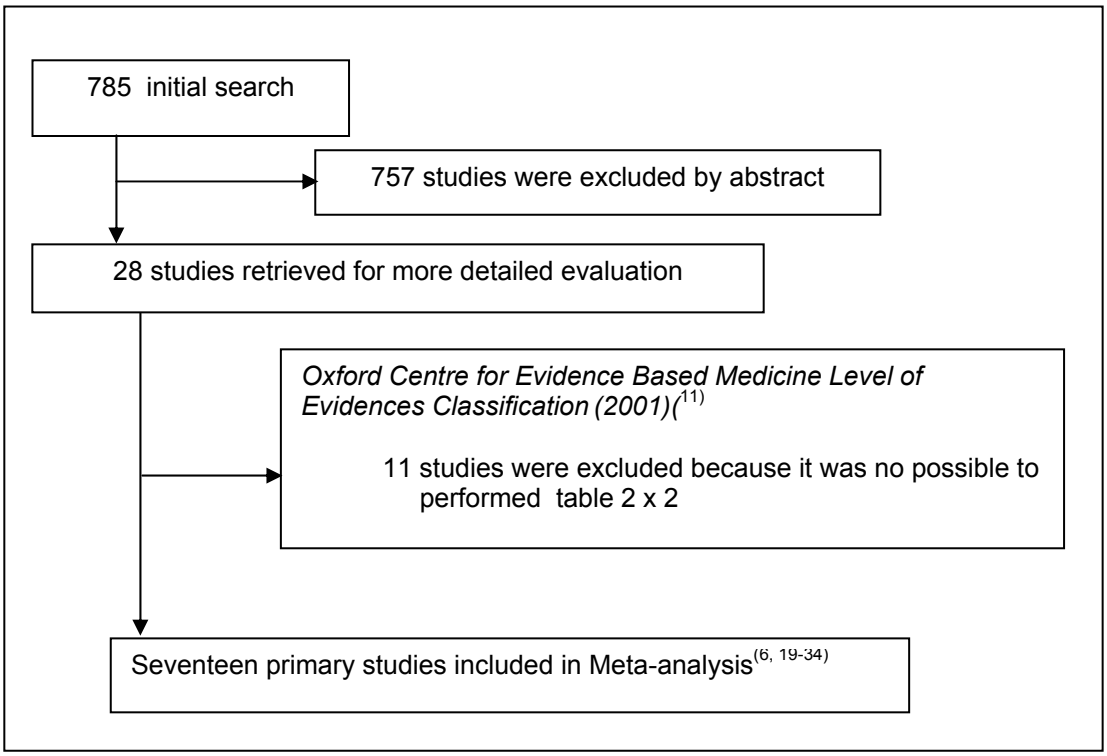


Figure 1. Study selection process.

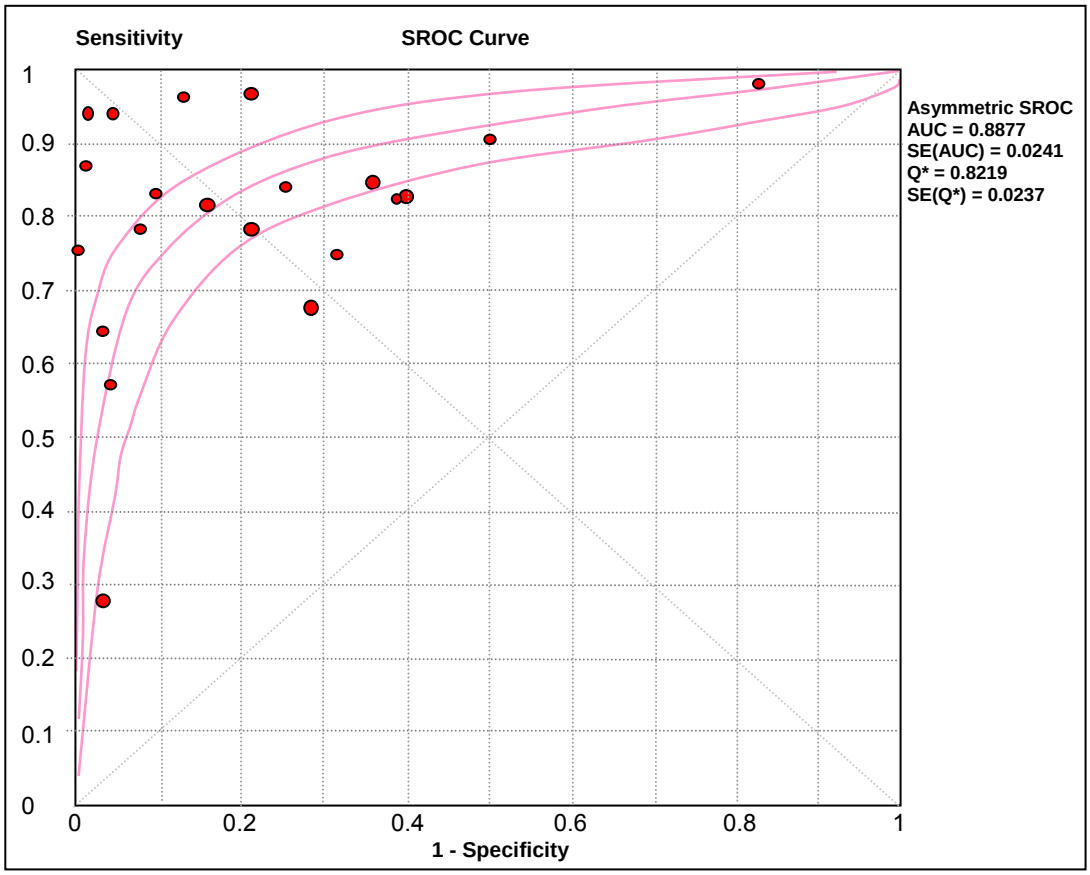


Figure 2. SROC plot of CA 125 level for detecting ovarian cancer, with 95% CI.

6 CONCLUSÕES E CONSIDERAÇÕES FINAIS

6 CONCLUSÕES E CONSIDERAÇÕES FINAIS

As conclusões advindas dos quatro artigos que abordam de maneira sistemática a pesquisa sobre as tumorações ovarianas deixa bem claro que:

- As melhores evidências científicas, mais fortes sobre o melhor tipo de abordagem – se por laparoscopia ou por laparotomia no câncer ovariano, estágio inicial (Ia, Ib e Ic) pela FIGO, indicam que a laparotomia é a melhor opção, pelo *guidelines* da FIGO (BENEDET, 2000). A laparoscopia pesquisada de forma sistemática para esse tipo de patologia não mostrou evidência que aprove essa conduta na prática clínica.
- O exame anatomopatológico de congelação é extremamente importante para definir qual a melhor conduta cirúrgica nas tumorações ovarianas. Mostrou-se extremamente confiável para os diagnósticos de benignidade e malignidade, com probabilidade pós-teste de 95% (IC 95% de 94-96%) e de 98% (IC 95% de 97-99%), respectivamente. Entretanto, quando os tumores possuem malignidade limítrofe, a probabilidade pós-teste é baixa sendo de 51% (IC 95% de CI 42-60%) para definir se a lesão é maligna ou de malignidade limítrofe.
- Os estudos de ultra-sonografia com Doppler colorido que sumarizam os resultados de sensibilidade e especificidade deste exame no câncer ovariano, mostraram os resultados de 0,87 (IC95% de 0,84-0,9) e 0,92 (IC 95% de 0,87-0,90) respectivamente. A DOR para tumores ovarianos malignos ou com malignidade limítrofe *versus* lesões benignas foi de 125 (IC 95% de 55-283). Foi calculada área sob a curva ROC, sendo o resultado de 0,9573. Em conclusão, a ultra-sonografia com Doppler colorido é um exame pré-teste importante na predição da natureza da tumoração ovariana, se maligna ou benigna.
- Os estudos de acurácia diagnóstica do CA 125 com nível sérico ≥ 35U/ml indicaram para sensibilidade e especificidade, respectivamente, valores de 0,8 (IC 95% de 0,76-0,82) e 0,75 (IC 95% de 0,73-0,77). A DOR para diagnóstico de câncer ovariano ou tumores com malignidade limítrofe *versus* lesões benignas foi

de 21,2 (IC 95% de 12-37). Foi calculada área sob a curva ROC foi de 0,8877. Em conclusão, o CA 125 com nível sérico ≥ 35 U/ml é um exame pré-teste importante, nas tumorações ovarianas, para predizer se a natureza da lesão é benigna ou maligna.

Em síntese, estes quatros estudos de revisão sistemática que abordam as tumorações ovarianas, mostraram níveis de evidência baixo para abordagem laparoscópica no câncer ovariano estágios inicias (Ia, Ib e Ic) pela FIGO. E que os três exames diagnósticos mais realizados na presença de tumoração ovariana são extremamente importantes para definir o diagnóstico. Definindo-se o diagnóstico, cabe adotar a melhor opção terapêutica.

7 ANEXOS

ANEXO A

Projeto

PROJETO

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE MEDICINA
PROGRAMA DE PÓS-GRADUAÇÃO EM EPIDEMIOLOGIA

Revisão sistemática com enfoque diagnóstico e terapêutico nas tumorações ovarianas

Lídia Rosi de Freitas Medeiros

Projeto de Pesquisa

Orientador:

Orientadora: Profa. Dra Mary Clarisse Bozzetti

Porto Alegre, 2008

1 - INTRODUÇÃO

Revisão sistemática realiza-se quando existe uma questão clínica que suscite dúvida, sendo seu objetivo principal ajudar na implantação de condutas validadas através da análise crítica dos estudos científicos (Cook *et al.*, 1997, Sackett and Rosenberg, 1995). Na área cirúrgica, existem diferentes estratégias para o tratamento de uma mesma doença, devendo-se avaliar os benefícios, os danos, os custos e a eficácia de cada procedimento escolhido (Kreder, 1999). Preferencialmente, a maior parte dessas informações deve proceder de revisões sistemáticas ou metanálises, a fim de validar a escolha da conduta cirúrgica com base nas melhores evidências científicas (Reeves, 1999,Sauerland *et al.*, 1999)

Em cirurgia ginecológica, o tratamento dos tumores ovarianos é tema que suscita inúmeras controvérsias quanto à melhor abordagem se por laparoscopia ou por laparotomia (Canis *et al.*, 1994, Chi and Curtin, 1999, Childers JM *et al.*, 1994), visto que os benefícios da cirurgia endoscópica são incontestáveis, como rápida recuperação, menos tempo de internação hospitalar e menor intensidade da dor no período pós-operatório (Lorenz *et al.*, 1999, Troidl, 1999)

Entretanto, nos casos de neoplasia maligna de ovário, o manejo por essa via torna-se inadequado, em decorrência das pequenas incisões que dificultam os cuidados operatórios preconizados para essas neoplasias (Benedet *et al.*, 2000). Além do que, o uso do gás carbônico, em cirurgia laparoscópica, propicia a disseminação e a implantação de células neoplásicas na cavidade abdominal (Köhler C *et al.*, 2004,Müller JM *et al.*, 1999,Smidt *et al.*, 2001,Volz *et al.*, 1999). Assim, o inadequado diagnóstico e o tratamento de uma falsa tumoração benigna através da endoscopia são fatores de mau prognóstico para o câncer de ovário (Wang *et al.*, 1999).

No entanto, estudos de casos selecionados para as tumorações ovarianas malignas, em seus estágios iniciais (Ia,Ib e Ic) têm sido publicados com maior freqüência nos últimos anos, dando a entender que a laparoscopia para as tumorações malignas iniciais (Ia, Ib e Ic), pela FIGO, seja o tratamento de escolha (Chi and Curtin, 1999, Curtin JP, 1994, Dottino P *et al.*, 1999). Contudo permanece a controvérsia sobre ser a laparoscopia uma boa escolha para as tumorações ovarianas malignas em estágios iniciais (Vergote I and Trimbos JB, 2003, Vergote and Amant, 2004).

Quando uma tumoração ovariana é detectada, faz-se necessário estabelecer suas características de benignidade ou malignidade, sendo que determinados exames, de certa forma, guiam a conduta terapêutica a ser adotada, apesar de não possuírem 100% de acurácia. Esses exames são:

- A ultra-sonografia transvaginal que usa um sistema de escore baseado nos bordos internos da tumoração, na presença de septações, no tipo de ecogenecidade e no volume do ovário (Sassone *et al.*, 1991). Ovários com volume superior a 20 cm^3 na idade reprodutiva e acima de 10 cm^3 em mulheres no período de pós-menopausa indicam necessidade de investigação (van, Jr. *et al.*, 2000).
- A ultra-sonografia com Doppler colorido que permite avaliar o fluxo sangüíneo nas tumorações através do cálculo do índice de resistência (IR) e do índice de pulsatilidade (PI), com pontos de corte para malignidade menor que 0,5 e 1 respectivamente (Brown *et al.*, 1998).
- O marcador tumoral CA 125 com níveis suspeitos para malignidade quando superiores a 35 U/ml (Maggino *et al.*, 1987,Maggino *et al.*, 1994);
- Durante o ato cirúrgico, impõe-se, conforme o tipo de tumoração, o exame anatomopatológico de congelação que orienta o médico na melhor conduta terapêutica a ser seguida.

Devido às incertezas que suscitam controvérsias quanto à melhor abordagem inicial para o manejo operatório das tumorações ovarianas – se por laparoscopia ou por laparotomia–, planejamos realizar estudos de revisão sistemática para avaliar a acurácia diagnóstica: do CA 125, da ecografia transvaginal com Doppler colorido e do exame anatomopatológico de congelação durante o transoperatório, por serem os principais exames para auxiliar na escolha da melhor conduta cirúrgica a ser seguida nas tumorações ovarianas. Concomitantemente procedemos ao estudo de revisão sistemática comparando laparoscopia e laparotomia, na abordagem cirúrgica do câncer ovariano, em seus estágios iniciais (Ia, Ib e Ic), segundo a Federação Internacional de Ginecologia e Obstetrícia (FIGO).

2 - OBJETIVOS

3.1. Objetivo geral

Realizar estudos de revisão sistemática com enfoque diagnóstico e terapêutico das tumorações ovarianas com pressupostos de malignidade em seus estágios iniciais (Ia, Ib e Ic) pela FIGO, tendo como finalidade principal que seus resultados auxiliem na tomada de decisão ou sugeriram a realização de novos estudos embasados em melhores níveis de evidência.

3.2 Objetivos específicos

3.2.1 Realizar revisão sistemática com enfoque de intervenção cirúrgica das tumorações ovarianas quanto à abordagem por laparoscopia ou por laparotomia.

3.2.2 Realizar revisão sistemática com enfoque diagnóstico avaliando a acurácia diagnóstica do exame anatomopatológico de congelação nas tumorações ovarianas.

3.2.3 Realizar revisão sistemática com enfoque diagnóstico avaliando a acurácia diagnóstica da ultra-sonografia com Doppler colorido nas tumorações ovarianas.

3.2.4 Realizar revisão sistemática com enfoque diagnóstico avaliando a acurácia diagnóstica do marcador tumoral CA 125 nas tumorações ovarianas.

ARTIGO 1

Laparoscopia versus laparotomia nos tumores de ovário estágio I conforme a FIGO.

Introdução

Tumores de ovário são responsáveis por 4% de todos os cânceres que afetam as mulheres e são a segunda causa de morte mais comum por motivos ginecológicos e a quarta causa mais comum de morte entre todos outros tipos que atingem a mulher (Yancik, 1993). Diagnóstico de cânceres de ovário nos estágios iniciais (limitado somente ao ovário – estágios Ia – Ib e Ic) é raro e geralmente esses tipos são descobertos de forma acidental por ultra-sonografia ou durante laparoscopia. A incidência de câncer ovariano insuspeito diagnosticado por laparoscopia está em torno de 6,5 /1000 pacientes com massa ovariana ((Wenzl *et al.*, 1996).

O prognóstico de todos os tipos de tumores ovarianos é afetado de forma independente pelos seguintes fatores: estágio do câncer no momento do diagnóstico, tipo, grau histológico e volume residual de doença (Benedet *et al.*, 2000). O tratamento considerado padrão para pacientes com câncer nos estágios iniciais (Ia, Ib e Ic) é a laparotomia com incisão longitudinal mediana, a qual permite amplo espaço para um perfeito estadiamento cirúrgico (Benedet *et al.*, 2000, Hand *et al.*, 1993, Kosary, 1994). Para todas as pacientes, a proposta cirúrgica de tratamento consiste em histerectomia total com anexectomia bilateral, omentectomia, linfadenectomia pélvica e para-aórtica e, em alguns casos, apendicectomia (Benedet *et al.*, 2000, Vinatier *et al.*, 1996).

O tratamento cirúrgico laparoscópico tem sido considerado um procedimento adequado para pacientes com exames sugestivos de benignidade (Vergote, 2004), sendo entretanto, inapropriado para casos com pressupostos de malignidade por estar associado com mau prognóstico (Lehner R *et al.*, 1998). Apesar disso, nos últimos anos, diversas publicações mostram que a técnica laparoscópica vem sendo a primeira escolha no manejo cirúrgico inicial de doenças ginecológicas malignas (Chi and Curtin, 1999, Dottino P *et al.*, 1999, Kadar N, 1997, Vinatier *et al.*, 1996), mas há controvérsia sobre ser a laparoscopia a melhor conduta para tumores malignos de ovário em seus estágios iniciais (Ia, Ib e Ic) (Vergote, 2004, Vergote and Amant, 2004, Vergote and Amant, 2004).

Em razão de não estar bem estabelecido se a laparoscopia é um tratamento melhor que o convencional (laparotomia) para o manejo do câncer de ovário em seus estágios iniciais, planejou-se elaborar uma revisão sistemática comparando laparoscopia e laparotomia para o manejo desses casos.

Objetivos da revisão sistemática

O objetivo desta revisão sistemática será avaliar o impacto da laparoscopia no tratamento cirúrgico dos tumores de ovários em seus estágios iniciais pela FIGO (Ia, Ib e Ic) quando comparado com a laparotomia.

Os seguintes pontos de pesquisa serão o foco desta revisão sistemática:

(1) É a laparoscopia (grupo de intervenção) efetiva no aumento total de sobrevida quando comparada com a laparotomia (grupo controle) em pacientes em estágio I dos tumores de ovário pela FIGO?

(2) É a laparoscopia (grupo de intervenção) efetiva em aumentar o tempo livre de doença quando comparada com a laparotomia (grupo controle) em pacientes em estágio I dos tumores de ovário pela FIGO?

(3) Sendo a laparoscopia (grupo de intervenção) um procedimento menos traumático, resulta em menor número de complicações quando comparada com a laparotomia (grupo controle) em pacientes com estágio I dos tumores de ovário pela FIGO?

(4) A laparoscopia (grupo de intervenção) tem maior número de recorrência local (*port site*) do que a laparotomia (grupo controle) na linha mediana da incisão em pacientes com estágio I dos tumores de ovário pela FIGO?

(5) A laparoscopia (grupo de intervenção) está associada com maior número de recorrência a distância quando comparada com a laparotomia (grupo controle) em pacientes com estágio I dos tumores de ovário pela FIGO?

(6) A laparoscopia (grupo de intervenção) resulta em maior número de casos de extravasamento de material do cisto durante o ato cirúrgico do que o que ocorre durante a laparotomia (grupo controle) em pacientes com estágio I dos tumores de ovário pela FIGO?

Critérios para considerar estudos para essa revisão sistemática

Critérios de inclusão – tipos de estudos

Foram incluídos estudos com pacientes classificadas como tendo estágio I do câncer ovariano de acordo com a FIGO. Estudos que comparam laparoscopia com laparotomia para tumores de ovários em estágios iniciais passaram a ser realizados somente após 1990.

Já se antecipa que deve ser muito pequeno o número de estudos randomizados que tenham sido conduzidos comparando as duas técnicas. Desta forma, estudos não randomizados, como os de coorte e de casos e controles, serão aceitos, mas de séries de casos ou com controle histórico serão excluídos.

Os subgrupos histológicos das tumorações ovarianas serão considerados quando possível conforme a classificação internacional realizada pela FIGO (Scully RE, 1999):

(1) Tumores da superfície epitelial e estromal:
   a) tipo seroso (malignidade limítrofe e malignos);
   b) tipo mucinoso (malignidade limítrofe e malignos);
   c) tipo endometrial.

(2) Tumores da célula germinativa:
   a) teratoma (imaturo ou do tipo monodermal;
   b) Disgerminoma;
   c) Tumor de *Yolk sac*;
   d) Carcinoma embrionário;
   e) Tumor carcinóide.

(3) Tumores do estroma do cordão sexual:
   a) tumores das células da granulosa;
   b) tumores das células de Sertoli (androblastomas);
   c) tumor do cordão sexual com túbulos anulares;
   d) ginandroblastoma;
   e) Tumores não classificáveis do estroma do cordão sexual;
   f) Tumores de células esteróides.

Critérios de exclusão

Foram excluídos todos estudos com pacientes com tumores iniciais de câncer ovariano que desejassem manter a fertilidade (tratadas com cirurgia conservadora – salpingo-ooforectomia unilateral), bem como aqueles em que o câncer ovariano foi inadequadamente estadiado.

Tipos de participantes

As pacientes deveriam apresentar câncer ovariano com estágio inicial pela FIGO, isto é, doença confinada somente aos ovários, sem comprometimento com linfonodos ou metástases a distância.

A Federação Internacional de Ginecologia e Obstetrícia (FIGO) distingue pacientes com estágios inicias da seguinte forma (Scully RE, 1999):

- Estágio Ia – tumor unilateral
- Estágio Ib – tumor bilateral
- Estágio Ic – tumor com extravasamento para fora do ovário, penetração capsular ou citologia peritoneal positiva

Trata-se de tumores sem linfonodos envolvidos ou metástases a distância

Se possível, os resultados serão estratificados conforme o tipo histológico.

Tipos de Intervenção

Nesta revisão sistemática dois tipos de procedimentos serão comparados para o tratamento cirúrgico dos tumores de ovário em seu estágio inicial: laparoscopia (grupo de intervenção) e laparotomia (grupo controle).

Tipos de desfechos mensurados

Desfechos primários

(1) Sobrevida em 5 anos;
(2) Tempo livre de doença em 5 anos.

Desfechos secundários

(1) Tumor com extravasamento no momento da cirurgia;
(2) Recorrência local (entrada do trocarte na laparoscopia e, na laparotomia no local da incisão na linha mediana);
(3) Recorrência a distância;

(4) Complicações cirúrgicas
   (a) Complicações cirúrgicas (imediatas e tardias):
      (i) injúria (bexiga, ureter, vascular, intestino delgado, intestino grosso);
      (ii) presença de aderências que complicam o tratamento cirúrgico;
      (iii) febre;
      (iv) obstrução intestinal;
      (v) hematoma;
      (vi) infecção;
      (vii) índice de conversão para laparotomia;
   (b) Complicações sistêmicas:
      (i) infecção pulmomar;
      (ii) trombose venosa profunda;
      (iii) embolia pulmonar;
      (iv) parada cardíaca;
      (v) isquemia cardíaca;
      (vi) acidente cerebrovascular;
   (c) Tempo operatório
   (d) Recuperação pós cirurgia:
      (i) tempo de hospitalização;
      (ii) índice de re-admissão por complicações.

Estratégias para identificação dos estudos

Serão conduzidos estudos para identificar todos os artigos randomizados publicados e não publicados comparando laparoscopia e laparotomia para os tumores de ovário em seus estágios iniciais. Serão pesquisadas publicações em todos os idiomas. Quando necessário, estudos que não forem escritos em Inglês serão traduzidos e somente após a tradução é que se decidirá sobre sua inclusão ou não na revisão sistemática.

Os estudos serão identificados através da pesquisa no banco de registros do Cochrane Gynaecological Cancer Group, na Central de Registro de estudos randomizados da Cochrane Library (CENTRAL), (The Cochrane Library Issue 2, 2007), no MEDLINE (janeiro de 1990 até novembro de 2007), no EMBASE (janeiro de 1990 até novembro de 2007), no LILACS (janeiro de 1990 até novembro de 2007), no BIOLOGICAL ABSTRACTS (janeiro de 1990 até novembro de 2007) e CANCERLIT (janeiro de 1990 até novembro de 2007).

No **MEDLINE** os estudos serão pesquisados mediante o uso das seguintes palavras-chave:

1. Randomized controlled trial. pt.
2. Controlled clinical trial.pt
3. Randomizes controlled trials/
4. random allocation/
5. double -blind method/
6. single-blind method/
7. or/1-6

8. clinical trial.pt
9. exp clinical trials/
10. (clin$ adj25 trial$).ti,ab,sh.
11. ((singl$ or doubl$ or tripl$ or trebl$) adj25 (blind$ or masks$)).ti,ab,sh.
12. placebos/
13. placebo$.ti,ab,sh
14. random$.ti,ab,sh.
15. Research design/
16. or/8-15
17. (animal not human).sh
18. 16 not 17
19. comparative study.sh
20. exp evaluation studies
21. follow up studies.sh
22. prospective studies
23. (control$ or prospectiv$).mp or volunter$.ti.ab.
24. exp cohort studies/
25. cohort.tw
26. exp longitudinal studies/
27. (cohort adj5 (stud$ or trial$)).tw
28. (prospectiv$ adj5 (stud$ or trial$)).tw
29. (longitudinal adj5 (stud$ or trials)).tw
30. or/18-29
31. exp Ovarian Neoplasms/
32. (ovar$ adj5 tumo?r).tw
33. (ovar$ adj5 neoplas$).tw
34. (ovar$ adj5 cancer$).tw
35. (ovar$ adj5 carcino$).tw
36. exp Adnexal Diseases/
37. exp Ovarian Cancer/
38. or/31-37
39. exp "early ovarian cancer"/
40. exp "early ovarian neoplasm"/
41. "stage I ovarian cancer"/
42. or/ 39-41
43. 38 and 42
44. exp SURGERY/
45. surg$.tw.
46. laparo$.tw.
47. exp Surgical procedures, Operative/
48. or/44-47
49. 43 and 48
50. 30 and 49

No **EMBASE** os estudos serão pesquisados mediante o uso das seguintes palavras-chave:
1. Controlled study/or Randomized Controlled trial/
2. double blind procedure/
3. single blind procedure/
4. crossover procedure/
5. drug comparison/
6. placebo/
7. random$.ti,ab,hw,tn,mf.
8. latin square.ti,ab,hw,tn,mf.
9. crossover.ti,ab,hw,tn,mf.
10. cross-over.ti,ab,hw,tn,mf.
11. placebo$.ti,ab,hw,tn,mf.
12. ((doubl$ or singl$ or tripl$ or trebl$) adj5 (blind$ or mask$)).ti,ab,hw,tn,mf.
13. (comparative adj5 trial$).ti,ab,hw,tn,mf.
14. (clinical adj5 trial$).ti,ab,hw,tn,mf.
15. or/ 1-14

16. nonhuman/
17. (animal not human)/
18. or/16-17
19. 15 not 18
20. comparative study.ti,ab,hw,tn,mf.
21. follow up studies.ti,ab,hw,tn,mf.
22. prospective studies.ti,ab,hw,tn,mf.
23. (control$ or prospectiv$).mp or volunteer$.ti.ab.
24. cohort studies/
25. cohort.ti,ab,hw,tn,mf.
26. longitudinal studies.ti,ab,hw,tn,mf.
27. (cohort adj5 trial$).ti,ab,hw,tn,mf.
28. (prospectiv$ adj5 trial$).ab,hw,tn,mf.
29. (longitudinal adj5 trials).ti,ab,hw,tn,mf.
30. or/19-29
31. exp Ovarian Neoplasms/
32. (ovar$ adj5 tumo?r).tw
33. (ovar$ adj5 neoplas$).tw
34. (ovar$ adj5 cancer$).tw
35. (ovar$ adj5 carcino$).tw
36. exp Adnexal Diseases/
37. exp Ovarian Cancer/
38. or/ 31-37
39. exp "early ovarian cancer"/
40. exp "early ovarian neoplasm"/
41. "stage I ovarian cancer"/
42. or/39-41
43. 38 and 42
44. exp SURGERY/
45. surg$.tw.
46. laparo$.tw.
47. exp Surgical Technique
48. or/43-46
49. 43 and 48
50. 30 and 49

Através da CENTRAL da *Cochrane Library*, estudos clínicos registrados serão pesquisados em todos os campos usando como palavras-chave: *ovarian cancer*, *laparotomy, laparoscopy, ovarian surgery*.

Listas de citações de relevantes publicações e resumos de jornadas e congressos científicos serão pesquisados mediante busca manual. Quando possível, serão mantidos contatos com os autores das publicações para esclarecer dúvidas. A busca manual será efetuada em importantes publicações na área de ginecologia oncológica tais como: *Gynecologic Oncology*; *International Journal of Gynecological Cancer; Bristish Journal of Cancer; Bristish Cancer Research Meeting; Annual Meeting of the International Gynecologic Cancer Society; Annual Meeting of the American Society of Gynecologic Oncologist; Annual Meeting of The European Society of Medical Oncology (ESMO); Annual Meeting of The American Society of Clinical Oncology (ASCO)*

MÉTODOS DA REVISÃO SISTEMÁTICA

Seleção de estudos

Todos os estudos considerados elegíveis para integrarem a revisão sistemática terão sua qualidade metodológica aferidas por diferentes tipos de instrumentos de avaliação de publicações científicas. Não ocorrerá cegamento dos autores nem das instuições onde os estudos foram realizados porque todos os revisores são bem familiarizados com a literatura sobre câncer ovariano em seus estágios iniciais. Como se pressume que provavelmente haverá muitos poucos estudos randomizados, sobre o tema, serão incorporamos nesta revisão sistemática estudos observacionais de coorte e de casos e controles, deixando-se de incluir apenas series de casos e coorte histórica. A qualidade de alocação dos estudos randomizados será classificada com orientação do Cochrane Gynaecological Cancer Group (Jadad *et al.*, 1996) : a)adequada; b)incerta; c) inadequada.

Para estudos observacionais de coorte e de casos e controles será usado o instrumento conhecido como *Newcastle-Ottawa Scale* (NOS), justamente para avaliar a qualidade metodológica desses estudos. Os revisores avaliarão os artigos de forma independente para aferir a qualidade metodológica de cada um deles, bem como para extrair dados como características clínicas dos participantes, intervenção e desfechos. As diferenças deverão ser resolvidas após discussão e consenso; quando os dados foram insuficientes, será feito contato com o autor do estudo.

ANÁLISE ESTATÍSTICA

A análise estatística deverá ser realizada de acordo com o *guideline* desenvolvido pelo Cochrane Gynaecological Cancer Group. Todos os estudos serão inicialmente inclusos na análise comparando os dois tipos de procedimento cirúrgico: laparoscopia e laparotomia para câncer ovariano em seus estágios iniciais (Ia, Ib e Ic) pela FIGO.

A heterogeneidade entre os resultados dos diferentes estudos será examinada usando-se o teste estatístico *Cochran's Q* onde o valor do P será obtido comparando-o com o teste estatístico da distribuição do qui-quadrado. Deve-se ter cuidado na avaliação do teste do qui-quadrado, desde que tenha baixo poder, pois isso constitui uma situação em metanálise em que os estudos possuem pequeno número de amostra. Considera-se para metanálise não existência de heterogeneidade se $p > 0,10$; no caso, de $p > 0,1$, deve-se usar o modelo de efeitos fixos. Porém, se ocorrer heterogeneidade significante ($p < 0,1$), devem-se explorar os motivos metodológicos de cada estudo, bem como as razões clínicas, usando-se neste tipo de caso o modelo de efeitos randômicos, para sumarizar os cálculos estatísticos (Deeks JJ *et al.*, 2003). Entretanto, uma vez que ocorre heterogeneidade clínica e metodológica, conclui-se que diversidades em uma metanálise são eventos comuns, sendo heterogeneidade inevitável. Uma alternativa para ajudar a quantificar essa heterogeneidade é o que se chama de inconsistência (I^2), dando ao estudo graus de inconsistência em seus resultados com intervalo de confiança de 95% (Higgins *et al.*, 2003). A inconsistência descreve a percentagem de variabilidade do efeito estimado devido à heterogeneidade. Inconsistência com valor 0 indica que não há heterogeneidade, valor acima de 50% pode ser considerados como uma substancial heterogeneidade. Quando existe um grau de inconsistência muito alto, torna-se, muitas vezes, inapropriado sumarizar todos os dados em um único valor devido à heterogeneidade estatística encontrada na revisão sistemática. Nestes casos, realiza-se somente uma revisão sistemática qualitativa.

Se possível deve-se realizar a metanálise dos desfechos primários e secundários. Nos casos, em que forem pesquisados, eventos de tempo como sobrevida e tempo livre de doença, será utilizada a *hazard ratio* (HR). Nos casos de impossibilidade, esta poderá ser estimada de forma indireta a partir de outros cálculos que sumarizam os resultados (Parmar *et al.*, 1998). Quando não for possível, a *Odds ratio* será calculada e interpretada com prudência, pois a possibilidade de mortalidade e morbidade em *Odds ratio* influencia o tempo de seguimento.

Para metanálise de desfechos dicotômicos. Será calculado o risco relativo (RR) com intervalo de confiança de 95% sumarizando-se os resultados da metanálise através do *software* RevMan.

Os dados contínuos serão combinados por metanálise. Serão usados média e desvio padrão para equilibrar o peso entre as diferenças das médias (WMD) com intervalo de confiança de 95%, usando-se modelos de efeitos fixos em casos de homogeneidade; em casos de heterogeneidade, os cálculos serão realizados através do modelo de efeitos randômicos.

Quando possível, a análise de subgrupos será planejada para comparar os resultados dos tipos de intervenção, conforme o tipo histológico e o desenho do estudo clínico.

ARTIGO 2

Acurácia do exame de anatomopatológico de congelação no diagnóstico das tumorações ovarianas: uma revisão sistemática

Introdução

Tumores malignos são responsáveis por 4% de todos os tipos de câncer que afetam a mulher, sendo a segunda causa mais comum de morte por câncer ginecológico e a quarta causa mais comum de mortes por todos os tipos de câncer em mulheres (Yancik, 1993). Lesões benignas, com malignidade limítrofe ou malignas podem ser encontradas em uma única lesão (Scully RE, 1995). Entretanto, a freqüência e a velocidade de evolução de uma lesão benigna para displasia ou câncer não são conhecidas (Scully RE, 1999).

O uso de exame anatomopatológico de congelação tem sido de grande impacto no cuidado das pacientes com tumorações ovarianas, pois é indispensável para o diagnóstico transoperatório de malignidade, podendo ajudar a determinar o estadiamento, assim como auxiliar na decisão sobre a melhor cirurgia a ser realizada em determinadas situações. Nos casos de câncer ovariano, a cirurgia geralmente envolve histerectomia total com anexectomia bilateral, omentectomia e linfadenectomia pélvica e paraáortica.

A acurácia do exame é geralmente muito boa. Entretanto, nos diagnósticos de lesões ovarianas com malignidade limítrofe é considerado um exame particularmente difícil (Spann *et al.*, 1994).

Por tais razões, decidiu-se realizar uma revisão sistemática quantitativa para avaliar a acurácia do exame anatomopatológico de congelação no diagnóstico das tumorações ovarianas e explorar os motivos de determinadas controvérsias.

MATERIAL E MÉTODOS

Uma busca em diversos bancos de dados, como o MEDLINE (PUBMED), o CANCERLIT, o LILACS, a COHRANE LIBRARY e o EMBASE, será feita entre janeiro de 1984 e dezembro de 2003. As expressões usadas para a busca serão: "*ovarian neoplasm*", "*frozen section*", combinadas com termos médicos de diagnóstico "*sensitivity" e "specificity*". A pesquisa será limitada a humanos, mas sem restrição de idioma. Serão também utilizadas as listas de referências de todos os estudos primários quando contiverem citações consideradas relevantes.

Estratégias de busca do "*Frozen section"* no PUBMED

1. "sensitivity and specificity" [all fields]
2. "sensitivity and specificity/standards" [all fields]
3. "specificity" [all fields]
4. "screening" [all fields]
5. "false positive" [all fields]
6. "false negative" [all fields]
7. #1 OR #2 OR #3 OR #4 OR #5 OR #6
8. "accuracy" [all fields]
9. "predictive value" [all fields]
10. "predictive value of tests" [all fields]
11. "reference value" [all fields]
12. "reference values" [all fields]
13. "reference values/standards" [all fields]
14. #8 OR #9 OR #10 OR #11 OR #12 OR #13
15. "roc" [all fields]
16. "roc analysis" [all fields]
17. "roc and" [all fields]
18. "roc area" [all fields]
19. "roc auc" [all fields]
20. "roc characteristics" [all fields]
21. "roc curve" [all fields]
22. "roc curve method" [all fields]
23. "roc curves" [all fields]
24. "roc estimated" [all fields]
25. "roc evaluation" [all fields]
26. "likelihood ratio" [all fields]

27. #16 OR #17 OR #18 OR #19 OR #20 OR #21 OR #22 OR #23 OR #24 OR #25 OR #26
28. #14 OR #27
29. Frozen section
30. Ovarian cysts [mh]
31. ovar*[tw] AND tumo*[tw]
32. ovar*[tw] AND cancer [tw]
33. Adnexal diseases [mh]
34. Ovarian neoplasms [mh]
35. Pelvic*[tw] AND tumo* [tw]
36. Pelvic*[tw] AND masses*[tw]
37. #30 OR #31 OR #32 OR #33 OR #34 OR #35 OR #36
38. #29 AND #37
39. #27 AND #38

Estratégias de busca "*Frozen section*" no EMBASE

1. (sensitiv$ adj5 specific$). af.
2. (sensitive$ adj5 specificity/standards). af.
3. "specificity". af.
4. "screening". af.
5. "false positive". af.
6. "false negative". af.
7. or/1-6
8. "accuracy". af.
9. "predictive value". af.
10. "predictive value of tests". af.
11. "reference value". af.
12. "reference values" . af.
13. "reference values/standards". af.
14. "roc" . af.
15. "roc analysis". af.
16. "roc adj". af.
17. "roc area".af.
18. "roc auc". af.
19. "roc characteristics". af.
20. "roc curve". af.
21. "roc curve method". af.
22. "roc curves". af.
23. "roc estimated". af.
24. "roc evaluation". af.
25. "likelihood ratio". af.
26. OR/8-25
27. #7 OR #26
28. "Frozen section". ti,ab,hw,tn,sh,mf.
29. nonhuman/
30. animal/ not (human/and animal/)
31. or/29-30
32. 27 not 31
33. exp Ovary Cyst/
34. exp Ovarian neoplasms/
35. (ovar$ adj5 tumor?r).tw.
36. (ovar$ adj5 tumo$).tw.
37. (ovar$ adj5 cancer).tw.
38. (ovar$ adj5 neoplasm$).tw.
39. exp Adnexa Disease/
40. (Pelvic$ adj5 tumo$). tw.
41. (Pelvic$ adj5 mass$).tw
42. OR/33-41
43. 28 AND 32
44. 42 AND 43

Através da CENTRAL da *Cochrane Library*, estudos do *The National Research Register* (NRR) e pesquisas clínicas registradas serão procuradas em todos os campos usando como palavras-chave: *ovarian cancer*, *frozen section, ovarian tumor e ovarian tumour.*

Listas de citações de relevantes publicações, resumos de jornadas e de congressos científicos serão pesquisados através de busca manual, procurando-se quando possível, manter contato com os autores das publicações para dirimir dúvidas. A busca manual será realizada em fontes da área de ginecologia oncológica como: *Gynecologic Oncology; International Journal of Gynecological Cance;, Bristish Journal of Cancer; Bristish Cancer Research Meeting; Annual Meeting of the International Gynecologic Cancer Society; Annual Meeting of the American Society of Gynecologic Oncologist; Annual Meeting of The European Society of Medical Oncology (ESMO).*

Critérios de seleção dos estudos

Esta revisão sistemática tem como foco estudos observacionais que comparam o exame anatomopatológico de congelação com o teste padrão-ouro, que é exame diagnóstico de patologia em parafina em pacientes submetidas a cirurgia por apresentarem tumoração ovariana. O exame anatomopatológico de congelação será considerado correto se não diferir do exame de parafina. Para serem incluídos os estudos todos devem ter a designação final de lesões benignas, de malignidade limítrofe ou malignas em exame histológico em parafina. Serão excluídos estudos que comparem a acurácia diagnóstica do exame anatomopatológico de congelação que descrevam somente um tipo das três patologias. Também não farão parte da revisão estudos em que seja impossível construir uma tabela 2 x 2.

O diagnóstico final será comparado de quatro maneiras: (1)concordância entre lesões benignas *vs* lesões limítrofes e/ou malignas; (2) concordância entre lesões malignas *vs* benignas; (3) concordância entre tumores de malignidade limítrofe *vs* lesões benignas; (4) concordância entre tumores de malignidade limítrofe *vs l*esões malignas. Casos em que houver dúvida serão excluídos do cálculo de acurácia final. O desfecho primário será mensurar a acurácia diagnóstica do exame anatomopatológico de congelação para as lesões benignas, para as com malignidade limítrofe e para as malignas. O desfecho secundário a ser pesquisado será a distribuição do tipo histológico de acordo com o exame histopatológico em parafina. Os estudos serão identificados de forma independente por quatro pesquisadores (L.R.M, D.D.R, M.I.E e M.C.B). A inclusão ou exclusão final dos artigos será baseada nos critérios de seleção contidos em um *checklist*. Discordâncias sobre inclusão ou exclusão serão inicialmente resolvidas por consenso e, quando não for possível, um quinto revisor será consultado (A.T.S.). As concordâncias entre os revisores serão computadas.

Avaliação da qualidade metodológica dos estudos

Todos os artigos selecionados terão sua qualidade metodológica avaliada. Essa avaliação consiste em analisar o tipo de delineamento do estudo, os critérios relevantes da população, a presença do teste diagnóstico (anatomopatológico de congelação) e o teste considerado padrão-ouro, que deve ser o histopatológico em parafina. A avaliação dos dados levará em consideração a forma como foram coletados, a seleção de pacientes, a descrição do método anatomopatológico de congelação e do exame de parafina, bem como a presença de viés (Bossuyt *et al.*, 2003, Jaeschke *et al.*, 1994). Os estudos incluídos terão seus dados avaliados através de um instrumento de avaliação de qualidade metodológica denominado de *Scoring of study quality* (Lijmer *et al.*, 1999). Os trabalhos também terão sua qualidade julgada através de outro instrumento para avaliação de qualidade metodológica conhecido como *Oxford Centre for Evidence-Based Medicine Level of Evidence Grades of Recomendantions rubric.* Somente estudos com níveis de evidência entre 1 e 3 serão considerados com alta qualidade, enquanto os classificados entre os níveis 4 e 5 serão excluídos (Phillips, 2007).

Extração dos dados

A extração dos dados será realizada por quatro investigadores de forma independente (L.R.M., D.D.R.,M.I.E. e M.C.B), com base nos seguintes dados: prevalência da lesões benignas, malignidade limítrofe e lesões ovarianas malignas, bem como os dados de sensibilidade, especificidade, razão de verossimilhança e probabilidade pós-teste de cada um dos estudos incluídos na metanálise. A avaliação dos artigos em inglês será feitas por dois revisores (L.R.M., D.D.R). Os artigos publicados em outro idioma que não o inglês terão seus dados extraídos também de forma independente por dois outros revisores (M.I.E e M.C.B.) quando se fizer necessária a tradução. Qualquer discordância será resolvida por consenso tanto para os artigos publicados em inglês como para os publicados em outros idiomas. Serão considerados três tipos de desfechos para as diferentes lesões ovarianas (benignas, malignidade e lesões com malignidade). As benignas serão

consideradas como desfecho primário, seguidas pelas de baixa incidência, como as lesões malignas e as de malignidade limítrofe. Dados dos tumores benignos serão colocados em tabelas 2 X 2, onde constam os resultados do exame anatomopatológico de congelação (positivo ou negativo para lesões benignas) e do teste final em parafina (tumores benignos *vs* malignos ou tumores de malignidade limítrofe). Tabelas de contingência serão construídas para comparar o diagnóstico anatomopatológico (positivo para tumores de malignidade limítrofe *vs* lesões benignas ou malignas) e o exame em parafina (tumores com malignidade limítrofe *vs* benignos ou malignos) assim como para o diagnóstico de tumores malignos comparando com o anatomopatológico de congelação (lesões malignas e benignas) e exame final em parafina (lesões malignas *vs* benignas).

Síntese dos dados e análise estatística

Para avaliar a concordância entre a elegibilidade e a qualidade metodológica assim como a concordância entre o exame anatomopatológico de congelação e o exame final histológico em parafina, serão calculadas as percentagens de concordância e o coeficiente κ (Altman DG, 1999). Em cada estudo será construída uma tabela 2 x 2 na qual as biópsias por exame anatomopatológico de congelação serão classificadas como benignas, de malignidade limítrofe ou malignas. Serão calculados os índices dos valores positivo-verdadeiros (sensibilidade) e índice falso-positivos (1-especificidade), as razões de verossimilhança positiva ou negativa e a probabilidade pós-teste para cada estudo com seus respectivos intervalos de confiança (ICs). Quando em alguma tabela 2 x 2 houver o valor de 0, será acrescentado o valor de 0,5, que será também adicionado aos outros valores das demais caselas.

Metanálises para produzir o somatório estimado da sensibilidade e da especificidade serão realizadas de forma independente. A associação entre sensibilidade e especificidade será calculada pelo teste de Sperman de correlação entre duas variáveis contínuas para os três tipos de lesões (lesões benignas, de malignidade limítrofe ou malignas) (Altman DG, 1999, Altman DG, 1999). Se não ocorrer correlação positiva, serão estimativas a sensibilidade e a especificidade porque se terá dois tipos de categoria – os resultados com desfecho positivo ou negativo do teste – e não existe variabilidade no limiar do teste diagnóstico (Deeks JJ, 2001). Razões de verossimilhança poderão ser estimadas a partir do sumário das estimativas de sensibilidade e especificidade usando-se as seguintes fórmulas: para razão de verossimilhança positiva, sensibilidade/1-especificidade e para razão de verossimilhança negativa, 1-sensibilidade/especificidade (Irwig *et al.*, 1994). Razão de verossimilhança indica que quanto maior o valor dado para o exame anatopatológico de congelação maior ou menor é a probabilidade do resultado final de o diagnóstico ser lesão benigna, de malignidade limítrofe ou maligna (Irwig *et al.*, 1994). Em adição, será calculado o somatório da probabilidade pós-teste que será calculada multiplicando o valor da probabilidade pré-teste (prevalência) pelo somatório das razões de verossimilhança (Deeks JJ, 2001). A heterogeneidade da sensibilidade e especificidade será calculada usando-se usando-se o teste estatístico *Cochran's Q* onde o valor do P será obtido comparando-o com o teste estatístico da distribuição do qui-quadrado com N-1 graus de liberdade. Caso a sensibilidade e a especificidade forem homogêneas, serão usados modelos de efeitos fixos para os cálculos do somatório geral, também com intervalo de confiança de 95% (Deeks JJ, 2001,Sutton AJ, 2000). A análise estatística será realizada com *software* Meta-DiSc® (versão Beta 1.1.1) (Zamora J *et al.*, 2006)

Análise de sensibilidade

Para verificar se a qualidade do estudo afeta o teste de acurácia diagnóstica do exame anatomopatológico de congelação, serão excluídos os estudos com menos de 50% dos critérios de qualidade e também os com níveis de evidência 3 com base no instrumento de análise do Centro de Evidência de Oxford (*Oxford Centre for Evidence-Based Medicine Level of Evidence Grades of Recomendations*) (Phillips B, 2007, Sutton AJ, 2000). Para análise do viés de publicação será adotado o gráfico do funil invertido, usando-se o log da razão de chance contra o tamanho da amostra (Deekes JJ, 2001, Sutton AJ, 2000). Para verificar a robustez dos resultados, serão repetidas as análises com modelos de efeito randômico, para ver se existe muita diferença nos cálculos com modelos fixos e randômicos. Caso os resultados encontrados sejam muitos díspares, será escolhido o modelo de efeitos randômicos para os cálculos da metanálise, por ser um tipo mais conservador e por conter o cálculo da heterogeneidade de cada um dos estudos selecionados para metanálise (Deekes JJ, 2001, Sutton AJ, 2000).

ARTIGO 3

Acurácia da ultra-sonografia com Doppler colorido no diagnóstico das tumorações ovarianas: uma revisão sistemática

Introdução

Cistos benignos e funcionais de ovário são muito freqüentemente encontrados em mulheres na idade reprodutiva, sendo responsáveis pela quinta causa de hospitalização por doenças ginecológicas nos Estados Unidos e na Inglaterra (Westhoff and Beral, 1984). Segundo dados do *National Hospital Discharge Surbey*, entre 1988 e 1990, a média anual de hospitalização por cistos benignos ovarianos foi de 32, 7 (IC 95% de 28,8-36,6) /10.000 mulheres na idade reprodutiva, sendo que 68% destas pacientes foram submetidas a procedimento cirúrgico (Velebil *et al.*, 1995).

É necessário dispor-se de procedimentos não invasivos para efetivamente discriminar as lesões ovarianas benignas das malignas (Tailor *et al.*, 1997). O uso da ultra-sonografia transvaginal com Doppler colorido tem facilitado o estudo do fluxo vascular no interior das lesões ovarianas, para diagnosticar áreas de angiogênese. Tumores malignos exibem uma vasta rede vascular com novos vasos em formação. Baixa impedância do fluxo intratumoral tem sido observada em casos de malignidade, com o índice de resistência abaixo de 0,4 em alguns casos (Bourne, 1991). Em adição, também é utilizado o índice de pulsatilidade (PI), que é inferior a 1 nas lesões com malignidade limítrofe ou malignas(Bourne *et al.*, 1989). Para evitar resultados falso-positivos de angiogênese do corpo lúteo, as mulheres no período pré-menopausa devem ser entre o 3º e o 8º dia após o ciclo menstrual. Kurjak *et al.* já utilizaram a ultra-sonografia com Doppler colorido para detectar lesões ovarianas malignas em seus estágios iniciais pela FIGO (Ia, Ib e Ic) (Kurjak *et al.*, 1993). Por tais razões, será conduzido um estudo de revisão sistemática para avaliar a acurácia diagnóstica da ultra-sonografia com Doppler colorido no diagnóstico de câncer ovariano e para explorar os motivos de certas controvérsias.

MATERIAL E MÉTODOS

Uma busca em diversos bancos de dados como o MEDLINE (PUBMED), o CANCERLIT, o LILACS, a COHRANE LIBRARY e o EMBASE, será feita entre janeiro de 1990 a dezembro de 2007. Os termos usados para a busca nos bancos de dados serão: "*ovarian neoplasm*" e "*transvaginal ultrasound with color Doppler*", combinados com expressões médicas de diagnóstico "*sensitivity" e "specificity*". A pesquisa será limitada a humanos, mas sem restrição de idioma. Serão também utilizadas listas de referências de todos os estudos primários quando considerada relevante.

Estratégias de busca da Ultrasonografia com Doppler no PUBMED

1. "sensitivity and specificity" [all fields]
2. "sensitivity and specificity/standards" [all fields]
3. "specificity" [all fields]
4. "screening" [all fields]
5. "false positive" [all fields]
6. "false negative" [all fields]
7. #1 OR #2 OR #3 OR #4 OR #5 OR #6
8. "accuracy" [all fields]
9. "predictive value" [all fields]
10. "predictive value of tests" [all fields]
11. "reference value" [all fields]
12. "reference values" [all fields]
13. "reference values/standards" [all fields]
14. #8 OR #9 OR #10 OR #11 OR #12 OR #13
15. "roc" [all fields]
16. "roc analysis" [all fields]
17. "roc and" [all fields]
18. "roc area" [all fields]
19. "roc auc" [all fields]
20. "roc characteristics" [all fields]
21. "roc curve" [all fields]
22. "roc curve method" [all fields]
23. "roc curves" [all fields]
24. "roc estimated" [all fields]
25. "roc evaluation" [all fields]

26. “likelihood ratio” [all fields]
27. #16 OR #17 OR #18 OR #19 OR #20 OR #21 OR #22 OR #23 OR #24 OR #25 OR #26
28. #14 OR #27
29. "color"[mh] AND Doppler [all fields]
30. “color” [tw] AND Doppler [all fields]
31. “colour” [tw] AND Doppler [all fields]
32. "ultrasonography"[TIAB] AND Doppler [all fields]
33. "ultrasonography"[mh] AND Doppler [all fields]
34. sonography [tw] AND Doppler [all fields]
35. "ultrasonography"[sh] AND transvaginal[All Fields]
36. "ultrasonography"[mh] AND transvaginal[All Fields]
37. “ultrasonography” [tw] AND transvaginal[All Fields]
38. "color"[mh] AND transvaginal[All Fields]
39. flow[All Fields] AND Doppler[All Fields]
40. transvaginosonography [tw]
41. #29 OR #30 OR #31 OR #32 OR #33 OR #34 OR #35 OR #36 OR #37 OR #38 OR #39 OR #40
42. Ovarian cysts [mh]
43. ovar*[tw] AND tumo*[tw]
44. ovar*[tw] AND cancer [tw]
45. Adnexal diseases [mh]
46. Ovarian neoplasms [mh]
47. Pelvic*[tw] AND tumo* [tw]
48. Pelvic*[tw] AND masses*[tw]
49. #42 OR #43 OR #44 OR #45 OR #46 OR #47 OR #48
50. #28 AND #41
51. #49 AND #50

Estratégias de busca para ultra-sonografia com Doppler colorido no EMBASE

1. (sensitiv$ adj5 specific$). af.
2. (sensitive$ adj5 specificity/standards). af.
3. “specificity”. af.
4. “screening”. af.
5. “false positive”. af.
6. “false negative”. af.
7. or/1-6
8. “accuracy”. af.
9. “predictive value”. af.
10. “predictive value of tests”. af.
11. “reference value”. af.
12. “reference values” . af.
13. “reference values/standards”. af.
14. “roc” . af.
15. “roc analysis”. af.
16. “roc adj”. af.
17. “roc area”.af.
18. “roc auc”. af.
19. “roc characteristics”. af.
20. “roc curve”. af.
21. “roc curve method”. af.
22. “roc curves”. af.
23. “roc estimated”. af.
24. “roc evaluation”. af.
25. “likelihood ratio”. af.
26. OR/8-25
27. #7 OR #26
28. (color adj5 Doppler$).ti,ab,hw,tn,sh,mf.
29. (colour adj5 Doppler$).ti,ab,hw,tn,sh,mf.
30. (ultrasonography$ adj5 Doppler$).ti,ab,hw,tn,sh,mf.

31. (sonography$ adj5 Doppler$).ti,ab,hw,tn,sh,mf.
32. (ultrasonograph$ adj5 transvaginal).ti,ab,hw,tn,sh,mf.
33. (color adj5 transvagin$).ti,ab,hw,tn,sh,mf.
34. (flow adj5 Doppler$).ti,ab,hw,tn,sh,mf.
35. transvaginosonography.ti,ab,hw,tn,sh,mf.
36. OR/28-35
37. #27 AND #36
38. nonhuman/
39. animal/ not (human/and animal/)
40. or/38-39
41. 37 not 40
42. exp Ovary Cyst/
43. exp Ovarian neoplasms/
44. (ovar$ adj5 tumor?r).tw.
45. (ovar$ adj5 tumo$).tw.
46. (ovar$ adj5 cancer).tw.
47. (ovar$ adj5 neoplasm$).tw.
48. exp Adnexa Disease/
49. (Pelvic$ adj5 tumo$). tw.
50. (Pelvic$ adj5 mass$).tw
51. OR/42-50
52. #41 AND #51

Através da CENTRAL da *Cochrane Library*, estudos do *The National Research Register* (NRR) e pesquisas clínicas registradas serão procurados em todos os campos usando como palavras- chave: *ovarian cancer*, *ultra-sonografia, Doppler, ovarian tumor* e *ovarian tumour*.

Listas de citações de relevantes publicações, resumos de jornadas e de congressos científicos serão pesquisados através de busca manual, procurando-se, quando possível, manter contato com os autores das publicações para dirimir dúvidas. A busca manual será realizada em fontes da área de ginecologia oncológica como: *Gynecologic Oncology; International Journal of Gynecological Cancer; Bristish Journal of Cancer; Bristish Cancer Research Meeting; Annual Meeting of the International Gynecologic Cancer Society; Annual Meeting of the American Society of Gynecologic Oncologist; Annual Meeting of The European Society of Medical Oncology (ESMO).*

Critérios de seleção

Esta revisão sistemática tem como foco estudos observacionais nos quais foram avaliadas, através de ultra-sonografia com Doppler colorido com 5 MHz, pacientes portadoras de tumores ovarianos, sendo esse resultado comparado com o teste padrão-ouro, que é exame diagnóstico de patologia em parafina em pacientes submetidas à cirurgia por apresentarem tumoração ovariana. Há suspeita de malignidade suspeita quando o índice de resistência encontrado for ≤ 0,5. Os casos do estudo serão as mulheres com tumor ovariano tratadas cirurgicamente. O teste diagnóstico será a ultra-sonografia com Doppler, e o teste considerado padrão-ouro será o histológico em parafina. O resultado da ultra-sonografia com Doppler colorido será considerada correto se não diferir do resultado histológico em parafina. Para serem incluídos os estudos devem conter exame de histológico em parafina com designação final de lesões benignas, de malignidade limítrofe ou malignas. Serão excluídos estudos que comparem a acurácia diagnóstica da ultra-sonografia com Doppler colorido ou que descrevam somente um dos tipos de patologia. Também serão excluídos estudos em que seja impossível construir tabela 2 x 2.

O diagnóstico final será comparado somente de uma maneira: lesões malignas ou com malignidade limítrofe *versus* lesões benignas. Casos em que houver dúvida serão excluídos do cálculo de acurácia final. O desfecho primário será mensurar a acurácia diagnóstica da ultra-sonografia com Doppler colorido para lesões benignas, de malignidades limítrofes e malignas. O desfecho secundário a ser pesquisado será a distribuição do tipo histológico de acordo com o exame histopatológico em parafina. Os estudos serão identificados de forma independente por três pesquisadores (L.R.M, D.D.R e M.I.R). A inclusão ou exclusão final dos artigos será baseada nos critérios de seleção contidos em um *checklist*. Discordâncias sobre inclusão e exclusão serão inicialmente resolvidas por consenso e, quando não for possível, um quarto revisor será consultado (M.C.B.) As concordâncias entre os revisores serão computadas.

Avaliação da qualidade metodológica dos estudos

Todos artigos selecionados terão sua qualidade metodológica avaliada. Essa avaliação consiste em analisar o tipo de delineamento do estudo, os critérios relevantes da população, a presença do teste diagnóstico (ultra-sonografia com Doppler colorido) e o teste considerado padrão-ouro, que deve ser o histopatológico em parafina. A avaliação dos dados levará em consideração a forma como foram coletados, a seleção de pacientes, a descrição do método da ultra-sonografia com Doppler colorido e do exame de parafina, bem como a presença de viés (Bossuyt *et al.*, 2003, Jaeschke *et al.*, 1994). Os estudos incluídos terão seus dados avaliados através de um instrumento de avaliação de qualidade metodológica denominado de *Scoring of study quality* (Lijmer *et al.*, 1999). Os estudos também terão sua qualidade julgada através de outro instrumento para avaliação de qualidade metodológica conhecido como *Oxford Centre for Evidence-Based Medicine Level of Evidence Grades of Recomendantions rubric.* Somente estudos com níveis de evidência entre 1 e 3 serão considerados com alta qualidade, enquanto os classificados entre os níveis 4 e 5 serão excluídos da avaliação (Phillips B, 2007).

Extração dos dados

A extração dos dados será realizada por três investigadores de forma independente (L.R.M, D.D.R e M.I.R) com base nos seguintes dados: prevalência da lesões benignas, malignidade limítrofe e lesões ovarianas malignas, bem como os dados de sensibilidade (verdadeiros positivos), especificidade (verdadeiros negativos, ou também 1- falso-positivos) dos estudos primários de ultra-sonografia com Doppler colorido. A avaliação dos artigos em inglês será feita por dois revisores (L.R.M., D.D.R). Os artigos publicados entre outros idiomas que não o inglês terão seus dados extraídos também de forma independente por dois outros revisores (M.I.R, M.C.B.), quando se fizer necessária a tradução. Qualquer discordância será resolvida por consenso tanto para os artigos publicados em inglês como para os artigos publica dos em outros idiomas. Serão considerados três tipos de desfechos para as lesões ovarianas (benignas, com malignidade limítrofe e lesões malignas). As benignas serão consideradas como desfecho primário, seguidas pelas de baixa incidência, como as lesões malignas e de malignidade limítrofe. Dados dos tumores benignos e malignos ou malignidade limítrofe serão colocados em tabelas 2 X 2, onde constam os resultados do exame de ultra-sonográfico com Doppler colorido (IR inferior ou igual a 0,5 ou superior 0,5) e do diagnóstico final em parafina (tumores benignos *vs* malignos ou tumores de malignidade limítrofe).

Síntese dos dados e análise estatística

Para avaliar a concordância entre a elegibilidade e a qualidade metodológica, assim como a concordância entre o exame anatomopatológico de congelação e o exame final histológico em parafina serão realizadas percentagens de concordância e o coeficiente κ (Altman DG, 1999). Em cada estudo será construída uma tabela de 2 x 2 na qual os resultados da ultra-sonografia com Doppler colorido serão classificados como: lesões benignas, de malignidade limítrofe ou malignas. Serão calculados os índices dos valores positivo-verdadeiros (sensibilidade) e índice falso-positivos (1-especificidade) para cada um dos estudos primários de ultra-sonografia com Doppler colorido. Quando em alguma tabela 2 x 2 houver o valor de 0 será acrescentado o valor de 0,5, que também será adicionado aos outros valores das demais caselas.

Metanálises para produzir o somatório estimado da sensibilidade e da especificidade serão realizadas de forma independente. Para realizar esse sumário, a sensibilidade e especificidade de cada estudo serão multiplicadas pelo número de sujeitos do estudo, sendo esse valor divido pelo número total dos sujeitos de todos os estudos. A associação entre sensibilidade e especificidade será calculada pelo teste de Sperman de correlação entre duas variáveis contínuas para os três tipos de lesões (benignas, de malignidade limítrofe ou malignas) (Altman DG, 1999,Altman DG, 1999). Se não ocorrer correlação positiva, se estimará a sensibilidade e a especificidade porque se terá dois tipos de categoria – os resultados com desfecho positivo ou negativo do teste – e não existe variabilidade no limiar do teste diagnóstico (Deeks JJ, 2001). No caso de ocorrer correlação positiva de Sperman, será apresentado o sumário da curva ROC que será calculado usando-se dados de todos os limiares, através do médodo de Littenberg e Moses (Irwig *et al.*, 1994, Suttom AJ *et al.*, 2000).

A DOR (*diagnostic odds ratio*) pode mostrar as diferentes combinações de sensibilidade e especificidade. Ela descreve a chance de um resultado de teste positivo em pacientes portadores da doença, comparando a chance de quem tem a doença com a de que não tem a doença. Um simples cálculo da DOR corresponde ao valor de limiar entre sensibilidade e especificidade apesar do sumário da curva ROC. Isso pode mudar de acordo com o limiar, e a curva ROC pode ser usada para definir um resultado definido como anormal ao exame. O sumário da curva ROC é um balanceamento entre sensibilidade e especificidade. Também a área sob a curva (AUC) pode sumarizar os resultados

inerentes do teste discriminatório de quem tem a doença e de quem não a tem. Testes considerados ideais possuem AUCs próximo ao valor 1, e testes considerados pobres para diagnóstico usualmente possuem valor próximo a 0,5.

A heterogeneidade da sensibilidade e especificidade será calculada usando-se usando-se o teste estatístico *Cochran's Q* onde o valor do P será obtido comparando-o com o teste estatístico da distribuição do qui-quadrado com N-1 graus de liberdade. Se a sensibilidade e a especificidade forem homogêneas, serão usados modelos de efeitos fixos para os cálculos do somatório geral, adotando-se também o intervalo de confiança de 95% (Deeks JJ, 2001, Sutton AJ, 2000). A análise estatística será realizada com *software* Meta-DiSc® (versão Beta 1.1.1) (Zamora J *et al.*, 2006)

Análise de sensibilidade

Para verificar se a qualidade do estudo afeta o teste de acurácia diagnóstica da ultra-sonografia com Doppler colorido, serão excluídos os estudos com menos de 50% dos critérios de qualidade e também com níveis de evidência 3, com base no instrumento de análise do Centro de Evidência de Oxford (*Oxford Centre for Evidence-Based Medicine Level of Evidence Grades of Recomendations*) (Phillips B, 2007, Sutton AJ, 2000). Para análise do viés de publicação, será adotado o gráfico do funil invertido usando-se o log da razão de chance contra o tamanho da amostra (Deekes JJ, 2001, Sutton AJ, 2000). Para verificar a robustez dos resultados, serão repetidas as análises com modelos de efeito randômico, para ver se existe muita diferença nos cálculos com modelos fixos e randômicos. Caso os resultados encontrados forem muitos díspares, será escolhido o modelo de efeitos randômicos para os cálculos da metanálise, por ser um tipo mais conservador e por conter o cálculo da heterogeneidade de cada um dos estudos selecionados para metanálise (Deekes JJ, 2001, Sutton AJ, 2000).

ARTIGO 4

Acurácia do nível sérico do CA 125 no diagnóstico das tumorações ovarianas: uma revisão sistemática

Introdução

Tumores malignos são responsáveis por 4% de todos os tipos de câncer que afetam a mulher, sendo a segunda causa mais comum de morte por câncer ginecológico e a quarta causa mais comum de mortes por todos os tipos de câncer nas mulheres (Yancik, 1993). Lesões benignas, com malignidade limítrofe ou malignas podem ser encontradas em uma única lesão (Scully RE, 1995). Entretanto, a freqüência e a velocidade de evolução de uma lesão benigna para displasia ou câncer não são conhecidas (Scully RE, 1999).

Tumores ovarianos suspeitos são comuns na prática clínica ginecológica. Embora a grande maioria das tumorações ovarianas seja de origem benigna, o objetivo primário da conduta clínica é a exclusão do diagnóstico de malignidade. Os métodos usados na elucidação diagnóstica de tumorações ovarianas suspeitas são o exame físico, a ultra-sonografia e a determinação do nível sérico de CA 125 (Alcázar JL *et al.*, 1999). O CA 125 é uma glicoproteína, determinante antigênico de alto peso molecular, reconhecida pelo anticorpo monoclonal (OC 125). Essa glicoproteína é expressa por tumores epiteliais assim como por tecidos de origem mülleriana (peritônio, pleura e pericárdio) (Alcázar JL *et al.*, 1999, Jacobs and Bast, Jr., 1989). O nível sérico do CA 125 ≥ 35 U/ml sugere suspeita de malignidade (Alcázar JL *et al.*, 1999, Jacobs and Bast, Jr., 1989). Em geral, o CA 125 está elevado em 85% das mulheres com tumor de origem epitelial e em 65% das pacientes com carcinoma mucinoso de ovário (Alcázar JL *et al.*, 1999, Jacobs and Bast, Jr., 1989). A acurácia diagnóstica é significativamente elevada quando se combinam o valor do nível sérico do CA 125 com os resultados da ultra-sonografia com Doppler (Maggino *et al.*, 1987).

Por tais razões, decidiu-se realizar uma revisão sistemática quantitativa para avaliar a acurácia diagnóstica dos níveis séricos do CA 125 no diagnóstico do câncer ovariano e explorar os motivos de determinadas controvérsias.

MATERIAL E MÉTODOS

Uma busca em diversos bancos de dados como o MEDLINE (PUBMED), o CANCERLIT, o LILACS,a COHRANE LIBRARY e o EMBASE será feita entre janeiro de 1990 e dezembro de 2007. As expressões usadas para a busca serão: "*ovarian neoplasm*", "*transvaginal ultrasound with color Doppler*", combinados com termos médicos de diagnóstico "*sensitivity" e "specificity*". A pesquisa será limitada a humanos, mas sem restrição de idioma. Serão também utilizadas as listas de referências de todos os estudos primários quando contiverem citações relevantes.

Estratégias de busca do CA 125 no PUBMED

1. "sensitivity and specificity" [all fields]
2. "sensitivity and specificity/standards" [all fields]
3. "specificity" [all fields]
4. "screening" [all fields]
5. "false positive" [all fields]
6. "false negative" [all fields]
7. #1 OR #2 OR #3 OR #4 OR #5 OR #6
8. "accuracy" [all fields]
9. "predictive value" [all fields]
10. "predictive value of tests" [all fields]
11. "reference value" [all fields]
12. "reference values" [all fields]
13. "reference values/standards" [all fields]
14. #8 OR #9 OR #10 OR #11 OR #12 OR #13
15. "roc" [all fields]
16. "roc analysis" [all fields]
17. "roc and" [all fields]
18. "roc area" [all fields]
19. "roc auc" [all fields]
20. "roc characteristics" [all fields]

21. “roc curve” [all fields]
22. “roc curve method” [all fields]
23. “roc curves” [all fields]
24. “roc estimated” [all fields]
25. “roc evaluation” [all fields]
26. “likelihood ratio” [all fields]
27. #16 OR #17 OR #18 OR #19 OR #20 OR #21 OR #22 OR #23 OR #24 OR #25 OR #26
28. #7 OR #15 OR #27
29. “CA 125” [all fields]
30. “monoclonal antibody” [mh] AND CA 125 [all fields]
31. “OC 125” [mh]
32. “Serum CA125” [all fields]
33. #29 OR #30 OR #31 OR #32
34. Ovarian cysts [mh]
35. ovar*[tw] AND tumo*[tw]
36. ovar*[tw] AND cancer [tw]
37. Adnexal diseases [mh]
38. Ovarian neoplasms [mh]
39. Pelvic*[tw] AND tumo* [tw]
40. Pelvic*[tw] AND masses*[tw]
41. #34 OR #35 OR #36 OR #37 OR #38 OR #39 OR #40
42. #28 AND #33
43. #41 AND #42

Estratégias de busca do CA 125 no EMBASE

1. (sensitiv$ adj5 specific$). af.
2. (sensitive$ adj5 specificity/standards). af.
3. “specificity”. af.
4. “screening”. af.
5. “false positive”. af.
6. “false negative”. af.
7. or/1-6
8. “accuracy”. af.
9. “predictive value”. af.
10. “predictive value of tests”. af.
11. “reference value”. af.
12. “reference values” . af.
13. “reference values/standards”. af.
14. “roc” . af.
15. “roc analysis”. af.
16. “roc adj”. af.
17. “roc area”.af.
18. “roc auc”. af.
19. “roc characteristics”. af.
20. “roc curve”. af.
21. “roc curve method”. af.
22. “roc curves”. af.
23. “roc estimated”. af.
24. “roc evaluation”. af.
25. “likelihood ratio”. af.
26. OR/8-25
27. #7 OR #26
28. “CA 125”. ti,ab,hw,tn,sh,mf.
29. (monoclonal antibody adj5 CA 125). ti,ab,hw,tn,sh,mf.
30. “OC 125” . ti,ab,hw,tn,sh,mf.
31. “Serum CA 125”. ti,ab,hw,tn,sh,mf.
32. OR/38-31
33. nonhuman/
34. animal/ not (human/and animal/)
35. or/33-34

36. 32 not 35
37. exp Ovary Cyst/
38. exp Ovarian neoplasms/
39. (ovar$ adj5 tumor?r).tw.
40. (ovar$ adj5 tumo$).tw.
41. (ovar$ adj5 cancer).tw.
42. (ovar$ adj5 neoplasm$).tw.
43. exp Adnexa Disease/
44. (Pelvic$ adj5 tumo$). tw.
45. (Pelvic$ adj5 mass$).tw
46. OR/37-45
47. #36 AND #46

Através da CENTRAL da *Cochrane Library*, estudos do *The National Research Register* (NRR) e pesquisas clínicas registradas serão procuradas em todos os campos usando como palavras-chave: *ovarian cancer*, Ca 125, *ovarian tumor* e *ovariam tumour*.

Listas de citações de relevantes publicações, resumos de jornadas e de congressos científicos serão pesquisados através de busca manual,procurando-se, quando possível, manter contato com os autores das publicações para dirimir dúvidas. A busca manual será realizada em fontes da área de ginecologia oncológica como: *Gynecologic Oncology; International Journal of Gynecological Cancer; Bristish Journal of Cancer; Bristish Cancer Research Meeting; Annual Meeting of the International Gynecologic Cancer Society; Annual Meeting of the American Society of Gynecologic Oncologist; Annual Meeting of The European Society of Medical Oncology (ESMO).*

Critérios de seleção

Esta revisão sistemática tem como foco estudos observacionais com dosagem sérica do CA 125 em pacientes portadoras de tumores ovarianos, sendo esses resultados comparados com o teste padrão- ouro, que é exame diagnóstico de patologia em parafina em pacientes submetidas a cirurgia por apresentarem tumoração ovariana. Há suspeita de malignidade se o nível sérico for ≥35 U/ml. Os casos do estudo serão as mulheres com tumor ovariano tratadas cirurgicamente. O teste diagnóstico é a dosagem sérica do CA 125, e o teste considerado padrão-ouro será o histológico em parafina. O resultado do nível sérico do CA 125 será considerado correto para malignidade limítrofe ou malignidade se não diferir do resultado histológico em parafina. Para serem incluídos, os estudos devem ter exame histológico em parafina com designação final de lesões benignas, de malignidade limítrofe ou malignas. Serão excluídos trabalhos que comparem a acurácia diagnóstica do CA 125 ou que descrevam somente um tipo das três patologias. Também não farão parte da revisão os estudos em que seja impossível construir uma tabela 2 x 2.

O diagnóstico final será comparado somente de uma maneira: lesões malignas ou com malignidade limítrofe *versus* lesões benignas. Casos em que houver dúvida serão excluídos do cálculo de acurácia final. O desfecho primário será mensurar a acurácia diagnóstica do nível sérico do CA 125 diagnosticando lesões benignas, de malignidade limítrofe e malignas. O desfecho secundário a ser pesquisado será a distribuição do tipo histológico de acordo com exame histopatológico em parafina. Os estudos serão identificados de forma independente por três pesquisadores (L.R.M, D.D.R e M.I.R) A inclusão ou exclusão final dos artigos será baseada nos critérios de seleção contidos em um *checklist*. Discordâncias sobre inclusão e exclusão serão inicialmente resolvidas por consenso e, quando não for possível, um quarto revisor será consultado (M.C.B.) A concordância entre os revisores serão computadas.

Avaliação da qualidade metodológica dos estudos

Todos os artigos selecionados terão sua qualidade metodológica avaliada. Essa avaliação consiste em analisar o tipo de delineamento do estudo, os critérios relevantes da população, a presença do teste diagnóstico (dosagem sérica do CA 125) e o teste considerado padrão-ouro, que deve ser o histopatológico em parafina. A avaliação dos dados levará em consideração a forma como foram coletados, a seleção de pacientes, a descrição da dosagem sérica do CA 125 e do exame de parafina e a presença de viés (Bossuyt *et al.*, 2003, Jaeschke *et al.*, 1994). Os estudos incluídos terão seus dados avaliados através de um instrumento de avaliação de qualidade metodológica denominado *Scoring of study quality* (Lijmer *et al.*, 1999). Os estudos também terão sua qualidade julgada através de outro instrumento para avaliação de qualidade metodológica conhecido como *Oxford Centre for Evidence-Based Medicine Level of Evidence Grades of Recomendantions rubric.*

Somente estudos com níveis de evidência entre 1 e 3 serão considerados com alta qualidade, enquanto os classificados entre os níveis 4 e 5 serão excluídos da avaliação (Phillips B, 2007).

Extração dos dados

A extração dos dados será realizada por quatro investigadores independentes (L.R.M., D.D.R, M.I.R e M.C.B) com base nos seguintes dados: prevalência da lesões benignas, malignidade limítrofe e lesões ovarianas malignas, bem como os dados de sensibilidade (verdadeiros positivos) e especificidade (verdadeiros negativos, ou também 1- falso-positivos) dos estudos primários com dosagem sérica do CA 125. A avaliação dos artigos em inglês será feita por dois revisores (L.R.M e D.D.R). Os artigos publicados entre outros idiomas que não o inglês terão seus dados extraídos também de forma independente por dois outros revisores (M.I.R e M.C.B.) quando se fizer necessária a tradução. Qualquer discordância será resolvida por consenso tanto para os artigos publicados em inglês como para os publicados em outros idiomas. Serão considerados três tipos de desfechos para as diferentes lesões ovarianas (benignas, com malignidade limítrofe e lesões malignas). As lesões benignas serão consideradas como desfecho primário, seguidas pelas de baixa incidência, como as lesões malignas e as de malignidade limítrofe. Dados dos tumores benignos, e malignos ou com malignidade limítrofe serão colocadas em tabelas 2 X 2, onde constam os resultados da dosagem sérica do CA 125 (< 35 U/ml ou ≥ 35U/ml) e o diagnóstico final em parafina (tumores benignos *vs* malignos ou com malignidade limítrofe).

Síntese dos dados e análise estatística

Para evoluir a concordância entre a elegibilidade e a qualidade metodológica, assim como a concordância entre anatomopatológico de congelação e exame final histológico em parafina serão realizadas percentagens de concordância e o coeficiente κ (Altman DG, 1999). Em cada estudo será construída uma tabela 2 x 2 na qual, de acordo com os resultados da dosagem sérica do CA 125, as lesões serão classificadas como benignas, de malignidade limítrofe ou malignas. Serão calculados os índices do valores verdadeiros positivos (sensibilidade) e índice falso-positivo (1-especificidade) para cada um dos estudos primários de ultra-sonografia com Doppler colorido. Quando em alguma tabela 2 x 2 houver valor de 0, será acrescentado o valor de 0,5, que será também adicionado aos outros valores das demais caselas.

Metanálises para produzir o somatório estimado da sensibilidade e da especificidade serão realizadas de forma independente. Para realizar esse sumário, a sensibilidade e a especificidade de cada estudo será multiplicada pelo número dos sujeitos do estudo, sendo esse valor divido pelo número total dos sujeitos de todos os estudos. A associação entre sensibilidade e especificidade será calculada através do teste de Sperman de correlação entre duas variáveis contínuas, tanto para as lesões benignas, como com malignidade limítrofe ou malignas (Altman DG, 1999,Altman DG, 1999). Se não ocorrer correlação positiva, serão estimadas a sensibilidade e a especificidade porque se terá dois tipos de categoria – os resultados com desfecho positivo ou negativo do teste – e não existe variabilidade no limiar do teste diagnóstico (Deeks JJ, 2001). No caso de ocorrer correlação positiva de Sperman, será apresentado o sumário da curva ROC que será calculada usando-se dados de todos os limiares, através do médodo de Littenberg e Moses (Irwig *et al.*, 1994, Suttom AJ *et al.*, 2000).

A DOR (*diagnostic odds ratio*) pode mostrar as diferentes combinações de sensibilidade e especificidade. Ela descreve a chance de um resultado de teste positivo em pacientes portadores da doença, comparando a chance de quem tem a doença com a de que não tem a doença. Um simples cálculo da DOR corresponde ao valor de limiar entre sensibilidade e especificidade apesar do sumário da curva ROC. Isso pode mudar de acordo com o limiar, e a curva ROC pode ser usada para definir um resultado definido como anormal ao exame. O sumário da curva ROC é um balanceamento entre sensibilidade e especificidade. Também a área sob a curva (AUC) pode sumarizar os resultados inerentes do teste discriminatório de quem tem a doença e de quem não a tem. Testes considerados ideais possuem AUCs próximo ao valor 1, e testes considerados pobres para diagnóstico usualmente possuem valor próximo a 0,5.

A heterogeneidade da sensibilidade e especificidade será calculada usando-se usando-se o teste estatístico *Cochran's Q* onde o valor do P será obtido comparando-o com o teste estatístico da distribuição do qui-quadrado com N-1 graus de liberdade. Se a sensibilidade e a especificidade forem homogêneas, serão usados modelos de efeitos fixos para os cálculos do somatório geral, adotando-se também o intervalo de confiança de 95% (Deeks JJ, 2001, Sutton AJ, 2000). A análise estatística será realizada com *software* Meta-DiSc® (versão Beta 1.1.1) (Zamora J *et al.*, 2006)

Análise de sensibilidade

Para verificar se a qualidade do estudo afeta o teste de acurácia diagnóstica da ultra-sonografia com Doppler colorido, serão excluídos os estudos com menos de 50% dos critérios de qualidade e também com níveis de evidência 3, com base no instrumento de análise do Centro de Evidência de Oxford (*Oxford Centre for Evidence-Based Medicine Level of Evidence Grades of Recomendations*) (Phillips B, 2007, Sutton AJ, 2000). Para análise do viés de publicação, será adotado o gráfico do funil invertido usando-se o log da razão de chance contra o tamanho da amostra (Deekes JJ, 2001, Sutton AJ, 2000). Para verificar a robustez dos resultados, serão repetidas as análises com modelos de efeito randômico, para ver se existe muita diferença nos cálculos com modelos fixos e randômicos. Caso os resultados encontrados forem muitos díspares, será escolhido o modelo de efeitos randômicos para os cálculos da metanálise, por ser um tipo mais conservador e por conter o cálculo da heterogeneidade de cada um dos estudos selecionados para metanálise (Deekes JJ, 2001, Sutton AJ, 2000).

Referências Bibliográficas

Alcázar JL, Errasti T, Zornosa A, Mínguez JA, Galán MJ. Transvaginal color Doppler ultrasonogrphy and CA 125 in suspicious adnexal masses. Int J Gynecol Obstet 1999; (66): 255-261.

Altman DG. Relation between two continuous variables. In: Practical statistics for medical research. (Ed.Altman DG). London: Chapman & Hall, 1999; 277-299.

Altman DG. Some common problems in medical research. In: Practical statistics for medical research. (Ed.Altman DG). London: Chapman & Hall, 1999; 403-409.

Aslam N, Banerjee S, Carr J V, Savvas M, Hooper R, Jurkovic D. Prospective evaluation of logistic regression models for the diagnosis of ovarian cancer. Obstet Gynecol 2000; (96): 75-80.

Barber H R. Ovarian cancer: diagnosis and management. Am J Obstet Gynecol 1984; (150): 910-916.

Benedet J L, Bender H, Jones H, III, Ngan H Y, Pecorelli S. FIGO staging classifications and clinical practice guidelines in the management of gynecologic cancers. FIGO Committee on Gynecologic Oncology. Int J Gynaecol Obstet 2000; (70): 209-262.

Berchuck A, Schildkraut J M, Marks J R, Futreal P A. Managing hereditary ovarian cancer risk. Cancer 1999; (86): 2517-2524.

Bossuyt P M, Reitsma J B, Bruns D E, Gatsonis C A, Glasziou P P, Irwig L M, Lijmer J G, Moher D, Rennie D, de Vet H C. Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD Initiative. Ann Intern Med 2003; (138): 40-44.

Bourne T H. Transvaginal color Doppler in gynecology. Ultrasound Obstet Gynecol 1991; (1): 359-373.

Bourne T, Campbell S, Steer C, Whitehead M I, Collins W P. Transvaginal colour flow imaging: a possible new screening technique for ovarian cancer. BMJ 1989; (299): 1367-1370.

Brioschi P A, Irion O, Bischof P, Bader M, Forni M, Krauer F. Serum CA 125 in epithelial ovarian cancer. A longitudinal study. Br J Obstet Gynaecol 1987; (94): 196-201.

Brown D L, Doubilet P M, Miller F H, Frates M C, Laing F C, DiSalvo D N, Benson C B, Lerner M H. Benign and malignant ovarian masses: selection of the most discriminating gray-scale and Doppler sonographic features. Radiology 1998; (208): 103-110.

Burger C W, Prinssen H M, Baak J P, Wagenaar N, Kenemans P. The management of borderline epithelial tumors of the ovary. Int J Gynecol Cancer 2000; (10): 181-197.

Canis M, Mage G, Pouly J L, Wattiez A, Manhes H, Bruhat M A. Laparoscopic diagnosis of adnexal cystic masses: a 12-year experience with long-term follow-up. Obstet Gynecol 1994; (83): 707-712.

Canis M, Pouly JL, Walttiez A, Mage G, Manhes H, Bruhat RS. Laparoscopic management of adnexal masses suspicious at ultrasound. Obstet Gynecol 1997; (89): 679-83.

Canis M, Rabischong B, Botchorishvili R, Tamburro S, Watiez A, Mage G, Pouly JL, *et al.* Risk of spread of ovarian cancer afher laparoscopic surgery. Curr Opin Obstet Gynecol 2001; (13): 9-14.

Canis M, Rabischong B, Houlle C B, Jarson K, Safi A W A, *et al.* Laparoscopic management of adnexal masses: a gold standart? Curr Opin Obstet Gynecol 2002; (14): 423-8.

Canis M, Watiez A, Mage G, Pouly JL, Raiga J, Goff B M B, Manhes H, Bruhat MA. Laparoscopic management of adnexal masses. Baillieres Best Pract Res Clin Obstet Gynaecol 1994; (8): 723-34.

Chen D X, Schwartz P E, Li X G, Yang Z. Evaluation of CA 125 levels in differentiating malignant from benign tumors in patients with pelvic masses. Obstet Gynecol 1988; (72): 23-27.

Chi D S, Curtin J P. Gynecologic cancer and laparoscopy. Obstet Gynecol Clin North Am 1999; (26): 201-215.

Childers J M, Aqua K A, Surwit E A, Hallum A V, Hatch K D. Abdominal-wall tumor implantation after laparoscopy for malignant conditions. Obstet Gynecol 1994; (84): 765-769.

Childers JM, Nasseri A, Surwit EA. Laparoscopic management of suspious adnexal masses. Am J Obstet Gynecol 1996; (175): 1451-9.

Childers JM. Operative laparoscopy in gynaecological oncology. Baillieres Best Pract Res Clin Obstet Gynaecol 1994; (8): 831-47.

Cochrane Methods Working Group on Systematic Reviews of Screening and Diagnostic Tests. Recommended methods. Cochrane Library 2004 January 25 [cited 2004 Jan 25];Available from: URL: http://www.cochrane.org/cochrane/sadtdoc1.htm

Concato J, Shah N, Horwitz R I. Randomized, controlled trials, observational studies, and the hierarchy of research designs. N Engl J Med 2000; (342): 1887-1892.

Cook D J, Mulrow C D, Haynes R B. Systematic reviews: synthesis of best evidence for clinical decisions. Ann Intern Med 1997; (126): 376-380.

Crayford T J, Campbell S, Bourne T H, Rawson H J, Collins W P. Benign ovarian cysts and ovarian cancer: a cohort study with implications for screening. Lancet 2000; (355): 1060-1063.

Cuello MF, Galleguillos GL, Zárate CR. Biopsia rápida por congelación en el diagnóstico de tumores de ovario: correlación diagnóstica según diámetro y peso en tumores de origem epitelial. Rev Méd Chile 1999; (127): 1199-1205.

Cuesta SR, Goff BA, Fuller AF, Nikrui N, Eichhorn JH, Rice LW. Prognostic importance of intraoperative rupture of malignant ovarian epithelial neoplasm. Obstet Gynecol 1994; (84): 1-7.

Curtin JP. Management of the adnexal mass. Gynecol Oncol 1994; (55): S42-S46.

da Cunha B A, Salvatore C A, Faria R M. Frozen section biopsy of ovarian neoplasms. Int J Gynaecol Obstet 1983; (21): 103-110.

Deekes JJ A D B M. Statistical methods for examining heterogeneity and combining results from several studies in meta-analysis. In: Systematic Reviews in Health care: Meta-analysis in context. (Ed.Egger M SGAD). London: BMJ Publishing, 2001; 285-312.

Deeks J J. Systematic reviews in health care: Systematic reviews of evaluations of diagnostic and screening tests. BMJ 2001; (323): 157-162.

Deeks JJ H J A D e. Analysing and presenting results. In: Cochrane Reviewers' Handbook 4.2.2 [updated March 2004]. (Ed.In: Alderson P GSHJe). Oxford: Cochrane Library, 2004.

Deeks JJ, Altman DG, Bradburn MJ. Statistical methods for examining heterogeneity and combining results from several studies in meta-analysis. In: Systematic Reviews in Health Care - Meta-analysis in context. (Eds.Egger M, Smit DG, Altman DG). London: BMJ Publishing Group, 2001; 2nd: 285-312.

Deeks JJ, Higgins JPT, Altman DJ. Analysing and presenting results. Alderson P, Green S, Higgins J, and Higgins J. [Cochrane Reviewers' Handbook 4.2.2 [updated March 2004]]. 2003. Ref Type: Computer Program

Deeks JJ. Systematic reviwes of evaluations of diagnostic and screening tests. In: Systematic reviews in health care - Meta-analysis in context. (Eds.Egger M, Smit DG, Altman DG). London: BMJ Publishing Group, 2001; 248-282.

Dembo A J, Davy M, Stenwig A E, Berle E J, Bush R S, Kjorstad K. Prognostic factors in patients with stage I epithelial ovarian cancer. Obstet Gynecol 1990; (75): 263-273.

Dembo AJ D M, Stenwing A, Berle EJ, Bush RS, Kjorstad K. Prognostic factors in patients with stage I epithelial ovarian cancer. Obstet Gynecol 2007; (75): 263-73D.

DePriest P D, Varner E, Powell J, Fried A, Puls L, Higgins R, Shenson D, Kryscio R, Hunter J E, Andrews S J, . The efficacy of a sonographic morphology index in identifying ovarian cancer: a multi-institutional investigation. Gynecol Oncol 1994; (55): 174-178.

DiSaia PJ, Creasman WT. The adnexal mass and early ovarian cancer. In: Clinical gynecologic oncology. (Eds.DiSaia PJ, Creasman WT). St Louis: Mosby, 2002; 6th edn: 259-288.

Donattino PR, Tobias DH, Beddoe AM, Golden AL, Cohen CJ. Laparoscopic lymphadenectomy for gynecologic malignancies. Gynecol Oncol 1999; (73): 383-388.

Dottino P, Levine DA, Ripley D C C. Laparoscopic Managment of adnexal Masses in premenopausal and Postmenopausal women. Obstet Gynecol 1999; (93): 223-7.

Dottino P, Tobias DH, Beddoe AM, Golden AL, Cohen CJ. Laparoscopic managment of bordeline ovarian tumors: results of a French multicenter study. Ann Oncol 2005; (16): 403-10.

Duffy M J, Bonfrer J M, Kulpa J, Rustin G J, Soletormos G, Torre G C, Tuxen M K, Zwirner M. CA125 in ovarian cancer: European Group on Tumor Markers guidelines for clinical use. Int J Gynecol Cancer 2005; (15): 679-691.

Editorial. Early ovarian cancer - time for a retthink on stage? Gynecol Oncol 2003; (90): 253-257.

Edmondson RJ M J. The epidemiology of ovarian cancer. Int J Gynecol Cancer 2001; (11): 423-429.

Egger M, Ebrahim S, Smith G D. Where now for meta-analysis? Int J Epidemiol 2002; (31): 1-5.

Egger M, Schneider M, Davey S G. Spurious precision? Meta-analysis of observational studies. BMJ 1998; (316): 140-144.

Egger M, Smith G D, Phillips A N. Meta-analysis: principles and procedures. BMJ 1997; (315): 1533-1537.

Egger M, Smith G D, Sterne J A. Uses and abuses of meta-analysis. Clin Med 2001; (1): 478-484.

Egger M, Smith G D. Meta-Analysis. Potentials and promise. BMJ 1997; (315): 1371-1374.

Egger M, Smith G D. Misleading meta-analysis. BMJ 1995; (310): 752-754.

Egger M, Smith GD, Schneider M. Sytematic reviews of observational studies. In: Systematic reviews in Health Care - Meta-analysis in context. (Eds.Egger M, Smith GD, Schneider M). London: BMJ Publishing, 2001; 2nd: 211-227.

Einhorn N, Bast R C, Jr., Knapp R C, Tjernberg B, Zurawski V R, Jr. Preoperative evaluation of serum CA 125 levels in patients with primary epithelial ovarian cancer. Obstet Gynecol 1986; (67): 414-416.

Einhorn N, Bast R, Knapp R, Nilsson B, Zurawski V, Sjövall K. Long term follo-up of the Stockholm screening study on ovarian cancer. Gynecol Oncol 2007; (79): 466-70.

Einhorn N, Knapp R C, Bast R C, Zurawski V R, Jr. CA 125 assay used in conjunction with CA 15-3 and TAG-72 assays for discrimination between malignant and non-malignant diseases of the ovary. Acta Oncol 1989; (28): 655-657.

Ekerhovd E, Wienerroith H, Staudach A, Granberg S. Preoperative assessment of unilocular adnexal cysts by transvaginal ultrasonography: a comparison between ultrasonographic morphologic imaging and histopathologic diagnosis. Am J Obstet Gynecol 2001; (184): 48-54.

Eleftheriadis E, Kotzampassi K P M H M S K. Gut Ischemia, oxidative stress, and bacterial translocation in elevated abdominal pressure in rats. World J Surg 1996; (20): 11-16.

Elit L, Chambers A, Fyles A, Covens A, Carey M, Fung M F. Systematic review of adjuvant care for women with Stage I ovarian carcinoma. Cancer 2004; (101): 1926-1935.

Eltabbakh G H, Natarajan N, Piver M S, Mettlin C J. Epidemiologic differences between women with borderline ovarian tumors and women with epithelial ovarian cancer. Gynecol Oncol 1999; (74): 103-107.

Erdogan N, Ozcelik B, Serin I S, Akgun M, Ozturk F. Doppler ultrasound assessment and serum cancer antigen 125 in the diagnosis of ovarian tumors. Int J Gynaecol Obstet 2005; (91): 146-150.

Ghezzi F, Cromi A U S, Bergamini V, Tomera S, Franchi M, Bolis P. Laparoscopy versus laparotomy for the surgical management of apparent early stage ovarian cancer. Gynecol Oncol 2007; (105): 409-13.

Gleeson NC, Nicosia SV, Mark JE, Hoffman MS, Cavanagh D. Adominal wall metastases from ovarian cancer after laparoscopy. Am J Obstet Gynecol 1993; (169): 522-3.

Go AS. Refinando a probabilidade:introdução à solicitação de exames complementares. In: Medicina baseada em evidências. (Eds.Friedland DJ, Go AS, Davoren JB, Shlipak MG, Bent SW, Subak LL, Mendelson T). Rio de Janeiro: Editora Guanabara Koogan, 2001; 1 ed: 12-31.

Gol M, Baloglu A, Yigit S, Dogan M, Aydin C, Yensel U. Accuracy of frozen section diagnosis in ovarian tumors: Is there a change in the course of time? Int J Gynecol Cancer 2003; (13): 593-597.

Granberg S. Ultrasound in the diagnosis and treatment of ovarian tumors. Acta Obstet Gynecol Scand 1991; (70): 385-386.

Greene FL. Principles of cancer biology in relayion to minimal acess surgical techniques. Semin Laparosc Surg 1995; (2): 155-157.

Greenhalgh T. How to read a paper. Papers that report diagnostic or screening tests. BMJ 1997; (315): 540-543.

Greenhalgh T. Papers that summarise other papers (systematic reviews and meta-analyses). BMJ 1997; (315): 672-675.

Guerriero S, Ajossa S, Lai M P, Alcazar J L, Paoletti A M, Marisa O, Melis G B. The diagnosis of functional ovarian cysts using transvaginal ultrasound combined with clinical parameters, CA125 determinations, and color Doppler. Eur J Obstet Gynecol Reprod Biol 2003; (110): 83-88.

Hamed F, Badia J, Chuaqui R, Wild R, Barrena N, Oyarzun E, Mayerson D. [Role of frozen section biopsy in the diagnosis of adnexal neoplasms]. Rev Chil Obstet Ginecol 1993; (58): 361-364.

Hand R, Fremgen A, Chmiel J S, Recant W, Berk R, Sylvester J, Sener S. Staging procedures, clinical management, and survival outcome for ovarian carcinoma. JAMA 1993; (269): 1119-1122.

Hata K, Hata T, Manabe A, Sugimura K, Kitao M. A critical evaluation of transvaginal Doppler studies, transvaginal sonography, magnetic resonance imaging, and CA 125 in detecting ovarian cancer. Obstet Gynecol 1992; (80): 922-926.

Higgins J P, Thompson S G, Deeks J J, Altman D G. Measuring inconsistency in meta-analyses. BMJ 2003; (327): 557-560.

Hua KQ, Jim FM, Xu H Z Z L J F Y. [Evaluation of laparoscopic surgery in the early stage malignant tumor of ovary with lower risk]. Zhonghua Yi Xue Za Zhi 2005; (85): 169-72.

INCA. Incidênia de câncer no Brasil em mulheres. Internet . 2007. Ref Type: Electronic Citation

Irwig L, Macaskill P, Glasziou P, Fahey M. Meta-analytic methods for diagnostic test accuracy. J Clin Epidemiol 1995; (48): 119-130.

Irwig L, Tosteson A N, Gatsonis C, Lau J, Colditz G, Chalmers T C, Mosteller F. Guidelines for meta-analyses evaluating diagnostic tests. Ann Intern Med 1994; (120): 667-676.

Irwig L, Tosteson A N, Gatsonis C, Lau J, Colditz G, Chalmers T C, Mosteller F. Guidelines for meta-analyses evaluating diagnostic tests. Ann Intern Med 1994; (120): 667-676.

Jacobs I, Bast R C, Jr. The CA 125 tumour-associated antigen: a review of the literature. Hum Reprod 1989; (4): 1-12.

Jadad A R, Moore R A, Carroll D, Jenkinson C, Reynolds D J, Gavaghan D J, McQuay H J. Assessing the quality of reports of randomized clinical trials: is blinding necessary? Control Clin Trials 1996; (17): 1-12.

Jaeschke R, Guyatt G, Sackett D L. Users' guides to the medical literature. III. How to use an article about a diagnostic test. A. Are the results of the study valid? Evidence-Based Medicine Working Group. JAMA 1994; (271): 389-391.

Juni P, Witschi A, Bloch R, Egger M. The hazards of scoring the quality of clinical trials for meta-analysis. JAMA 1999; (282): 1054-1060.

Kabawat S E, Bast R C, Welch W R, Knapp R C, Colvin R B. Immunopathologic characterization of a monoclonal antibody that recognizes common surface antigens of human ovarian tumors of serous, endometrioid, and clear cell types. Am J Clin Pathol 1983; (79): 98-104.

Kadar N. Laparoscopic surgery for gynaecological malignancies in women age 65 years or more. Gynaecological Endoscopy 1995; (4): 173-81.

Kadar N. Laparoscopy managment of gynecological malignancies. Curr Opin Obstet Gynecol 1997; (9): 247-255.

Kadar N. Port site recurrence following laparoscopic operations for gynaecological malignancies. Br J Obstet Gynaecol 1997; (104): 1308-13.

Kiu-Kwong C, Fang-Ping C, Shuenn-Dyh. Laparoscopic surgical procedures for early ovarian cancer. Acta Obstet Gynecol Scand 1995; (74): 391-401.

Klidermann G, Massen V, Kuhn W. [Laparoscopic preliminary surgery of ovarian malignancies. Experiences from 127 German gynecologic]. Geburtsshife Frauenheilkd 1995; (55): 687-94.

Köhler C, Klemm P, Schau A, Possover M, Krause N, Tozzi R, Schneider A. Introduction of transperitoneal lymphaadenectomy in a gynecologic oncology center: analysis of 650 laparoscopic pelvic and/or paraaortic transperitoneal lymphadenectomies. Gynecol Oncol 2004; (95): 52-61.

Koonings P P, Campbell K, Mishell D R, Jr., Grimes D A. Relative frequency of primary ovarian neoplasms: a 10-year review. Obstet Gynecol 1989; (74): 921-926.

Kosary C L. FIGO stage, histology, histologic grade, age and race as prognostic factors in determining survival for cancers of the female gynecological system: an analysis of 1973-87 SEER cases of cancers of the endometrium, cervix, ovary, vulva, and vagina. Semin Surg Oncol 1994; (10): 31-46.

Kreder H J. Evidence-based surgical practice: what is it and do we need it? World J Surg 1999; (23): 1232-1235.

Kurjak A, Kupesic S. Transvaginal color Doppler and pelvic tumor vascularity: lessons learned and future challenges. Ultrasound Obstet Gynecol 1995; (6): 145-159.

Kurjak A, Predanic M. New scoring system for prediction of ovarian malignancy based on transvaginal color Doppler sonography. J Ultrasound Med 1992; (11): 631-638.

Kurjak A, Schulman H, Sosic A, Zalud I, Shalan H. Transvaginal ultrasound, color flow, and Doppler waveform of the postmenopausal adnexal mass. Obstet Gynecol 1992; (80): 917-921.

Kurjak A, Shalan H, Kupesic S, Predanic M, Zalud I, Breyer B, Jukic S. Transvaginal color Doppler sonography in the assessment of pelvic tumor vascularity. Ultrasound Obstet Gynecol 1993; (3): 137-154.

Kurjak A, Shalan H, Matijevic R, Predanic M, Kupesic-Urek S. Stage I ovarian cancer by transvaginal color Doppler sonography: a report of 18 cases. Ultrasound Obstet Gynecol 1993; (3): 195-198.

Kurjak A, Zalud I, Jurkovic D, Alfirevic Z, Miljan M. Transvaginal color Doppler for the assessment of pelvic circulation. Acta Obstet Gynecol Scand 1989; (68): 131-135.

Kurtz A B, Tsimikas J V, Tempany C M, Hamper U M, Arger P H, Bree R L, Wechsler R J, Francis I R, Kuhlman J E, Siegelman E S, Mitchell D G, Silverman S G, Brown D L, Sheth S, Coleman B G, Ellis J H, Kurman R J, Caudry D J, McNeil B J. Diagnosis and staging of ovarian cancer: comparative values of Doppler and conventional US, CT, and MR imaging correlated with surgery and histopathologic analysis--report of the Radiology Diagnostic Oncology Group. Radiology 1999; (212): 19-27.

Lécuru F, Desfeux P, Camatte S, Bissery A, Robin F, Blanc B, Querleu D. Stage I ovarian cancer:comparison of laparoscopy and laparotomy on stanging and survival. Eur J Gynaec Oncol 2004;571-6.

Lécuru F, Taurelle F. Transperitoneal laparoscopic pelvic lypmphadenectomy for gynecologic malignancies. Sur Endosc 1998; (12): 97-100.

Lefebvre C, Clarke MJ. Identifying randomised trials. In: Systematic reviews in health care - Meta-analysis in context. (Eds.Egger M, Smith GD, Altman DG). London: BMJ Pushishing Group, 2001; 69-86.

Lehner R, Wenzl R, Heinzl H, Husslein P, Sevelda P. Influence of delayed statging laparotomy after laparoscopy removal of ovariam masses later found malignant. Obstet Gynecol 1998; (92): 967-71.

Leminen A, Mage G. Spread of ovarian cancer after laparoscopic surgery report of eight cases. Gynecol Oncol 1999; (75): 387-390.

Lerman R I, Pitcock J A. Frozen section experience in 3,249 specimens. Surg Gynecol Obstet 1972; (135): 930-932.

Lijmer J G, Mol B W, Heisterkamp S, Bonsel G J, Prins M H, van der Meulen J H, Bossuyt P M. Empirical evidence of design-related bias in studies of diagnostic tests. JAMA 1999; (282): 1061-1066.

Lim F K, Yeoh C L, Chong S M, Arulkumaran S. Pre and intraoperative diagnosis of ovarian tumours: how accurate are we? Aust N Z J Obstet Gynaecol 1997; (37): 223-227.

Littenberg B, Moses L E. Estimating diagnostic accuracy from multiple conflicting reports: a new meta-analytic method. Med Decis Making 1993; (13): 313-321.

Littenberg B, Moses L E. Estimating diagnostic accuracy from multiple conflicting reports: a new meta-analytic method. Med Decis Making 1993; (13): 313-321.

Lorenz W, Troidl H, Solomkin J S, Nies C, Sitter H, Koller M, Krack W, Roizen M F. Second step: testing-outcome measurements. World J Surg 1999; (23): 768-780.

Maggino T, Gadducci A, D'Addario V, Pecorelli S, Lissoni A, Stella M, Romagnolo C, Federghini M, Zucca S, Trio D, . Prospective multicenter study on CA 125 in postmenopausal pelvic masses. Gynecol Oncol 1994; (54): 117-123.

Maggino T, Sopracordevole F, Matarese M, Di P C, Tambuscio G. CA-125 serum level in the diagnosis of pelvic masses: comparison with other methods. Eur J Gynaecol Oncol 1987; (8): 590-595.

Malkasian G D, Jr., Knapp R C, Lavin P T, Zurawski V R, Jr., Podratz K C, Stanhope C R, Mortel R, Berek J S, Bast R C, Jr., Ritts R E. Preoperative evaluation of serum CA 125 levels in premenopausal and postmenopausal patients with pelvic masses: discrimination of benign from malignant disease. Am J Obstet Gynecol 1988; (159): 341-346.

Mancuso A, De V A, Triolo O, Irato S. The role of transvaginal ultrasonography and serum CA 125 assay combined with age and hormonal state in the differential diagnosis of pelvic masses. Eur J Gynaecol Oncol 2004; (25): 207-210.

Manolitsas TP, Folwer JM. Role of laparoscopy in the management of the adnexal mass and staging of gynecologic cancers. Clin Obstet Gynecol 2001; (44): 495-521.

Medeiros L R, Rosa D D, Edelweiss M I, Stein A T, Bozzetti M C, Zelmanowicz A, Pohlmann P R, Meurer L, Carballo M T. Accuracy of frozen-section analysis in the diagnosis of ovarian tumors: a systematic quantitative review. Int J Gynecol Cancer 2005; (15): 192-202.

Midgette A S, Stukel T A, Littenberg B. A meta-analytic method for summarizing diagnostic test performances: receiver-operating-characteristic-summary point estimates. Med Decis Making 1993; (13): 253-257.

Mogensen O, Mogensen B, Jakobsen A. CA 125 in the diagnosis of pelvic masses. Eur J Cancer Clin Oncol 1989; (25): 1187-1190.

Moher D, Cook D J, Eastwood S, Olkin I, Rennie D, Stroup D F. Improving the Quality of Reports of Meta-Analyses of Randomised Controlled Trials: The QUOROM Statement. Onkologie 2000; (23): 597-602.

Moher D, Schulz K F, Altman D. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. JAMA 2001; (285): 1987-1991.

Müller JM, Schwenk W, Jacobi CA, Böhm B. Endoscopy surgery: fit for malignancy? World J Surg 1999; (23): 808-15.

Mulrow C D, Cook D J, Davidoff F. Systematic reviews: critical links in the great chain of evidence. Ann Intern Med 1997; (126): 389-391.

Mulrow C D. Rationale for systematic reviews. BMJ 1994; (309): 597-599.

National Institutes of Health Consensus Development Conference Statement. Ovarian cancer: screening, treatment, and follow-up. Gynecol Oncol 1994; (55): S4-14.

Nelson L, Ekbom A, Gerdin E. Ovarian cancer in young women in Sweden, 1989-1991. Gynecol Oncol 1999; (74): 472-476.

Obermair A, Fuller A L-V E v G T, Vergote I, Eaton L F, *et al.* A new prognostic model for FIGO stage 1 epithelial ovarian cancer. Gynecol Oncol 2007; (104): 607-11.

Obiakor I, Maiman M, Mittal K, Awobuluyi M, DiMaio T, Demopoulos R. The accuracy of frozen section in the diagnosis of ovarian neoplasms. Gynecol Oncol 1991; (43): 61-63.

Osmers R G, Osmers M, von M B, Wagner B, Kuhn W. Evaluation of ovarian tumors in postmenopausal women by transvaginal sonography. Eur J Obstet Gynecol Reprod Biol 1998; (77): 81-88.

Oxman A D, Cook D J, Guyatt G H. Users' guides to the medical literature. VI. How to use an overview. Evidence-Based Medicine Working Group. JAMA 1994; (272): 1367-1371.

Parmar M K, Torri V, Stewart L. Extracting summary statistics to perform meta-analyses of the published literature for survival endpoints. Stat Med 1998; (17): 2815-2834.

Paspalicchio JC, Fristachi CE., Castanho PROL, Kue CM, Piatto S, Bacarat FF. Epidemiologia do câncer de ovário no Brasil. Revista da Sociedade Brasileira de Cancerologia 11. 2000. Ref Type: Electronic Citation

Phillips B, Ball C, Sackett D *et al.* Oxford Centre for evidence-based Medicine Level of evidence Grades of Recomendations (may 2001). Oxford Centre for evidence-based Medicine . 2007. Ref Type: Electronic Citation

Pinto P B, Andrade L A, Derchain S F. Accuracy of intraoperative frozen section diagnosis of ovarian tumors. Gynecol Oncol 2001; (81): 230-232.

Pomel C, Provencher D D J, Gauthier P, Le Bouedec G, Drouuin P, Audet-Lapointe P, *et al.* Laparoscopic staging of early ovarian cancer. Gynecol Oncol 1995; (58): 301-306.

Predanic M, Vlahos N, Pennisi J A, Moukhtar M, Aleem F A. Color and pulsed Doppler sonography, gray-scale imaging, and serum CA 125 in the assessment of adnexal disease. Obstet Gynecol 1996; (88): 283-288.

Puls L, Heidtman E, Hunter J E, Crane M, Stafford J. The accuracy of frozen section by tumor weight for ovarian epithelial neoplasms. Gynecol Oncol 1997; (67): 16-19.

Querleu D, Leblanc E, Cartron G N F F G, Martel P. Audit of preoperative and early complications lymph node dissection in 1000 gynecologic cancer patients. Am J Obstet Gynecol 2006; (195): 1287-92.

Querleu D, Leblanc E, Ferron G, Narducci F. Laparoscopy surgery in gynaecological oncology. EJSO 2006; (32): 853-58.

Ramirez P T, Frumovitz M, Wolf J K, Levenback C. Laparoscopic port-site metastases in patients with gynecological malignancies. Int J Gynecol Cancer 2004; (14): 1070-1077.

Reeves B. Health-technology assessment in surgery. Lancet 1999; (353 Suppl 1): SI3-SI5.

Richardson W S, Wilson M C, Guyatt G H, Cook D J, Nishikawa J. Users' guides to the medical literature: XV. How to use an article about disease probability for differential diagnosis. Evidence-Based Medicine Working Group. JAMA 1999; (281): 1214-1219.

Romagnolo C, Gadducci A, Sartori E, Zola P, Maggino T. Management of bordeline ovarian tumors:results of an Italian mulcenter study. Gynecol Oncol 2006; (101): 255-260.

Roman L D, Muderspach L I, Stein S M, Laifer-Narin S, Groshen S, Morrow C P. Pelvic examination, tumor marker level, and gray-scale and Doppler sonography in the prediction of pelvic cancer. Obstet Gynecol 1997; (89): 493-500.

Rose P G, Rubin R B, Nelson B E, Hunter R E, Reale F R. Accuracy of frozen-section (intraoperative consultation) diagnosis of ovarian tumors. Am J Obstet Gynecol 1994; (171): 823-826.

Sackett D L, Rosenberg W M. On the need for evidence-based medicine. J Public Health Med 1995; (17): 330-334.

Sackett D L. Clinical diagnosis and the clinical laboratory. Clin Invest Med 1978; (1): 37-43.

Sackett D L. Clinical epidemiology. Am J Epidemiol 1969; (89): 125-128.

Sackett D L. Clinical epidemiology. what, who, and whither. J Clin Epidemiol 2002; (55): 1161-1166.

Sassone A M, Timor-Tritsch I E, Artner A, Westhoff C, Warren W B. Transvaginal sonographic characterization of ovarian disease: evaluation of a new scoring system to predict ovarian malignancy. Obstet Gynecol 1991; (78): 70-76.

Sauerland S, Lefering R, Neugebauer E A. The pros and cons of evidence-based surgery. Langenbecks Arch Surg 1999; (384): 423-431.

Scully R E. Early de novo ovarian cancer and cancer developing in benign ovarian lesions. Int J Gynaecol Obstet 1995; (49 Suppl): S9-15.

Scully R E. Influence of origin of ovarian cancer on efficacy of screening. Lancet 2000; (355): 1028-1029.

Scully RE. Early de novo ovarian cancer and cancer developing in benign ovarial lesions. Int J Gynecol Cancer 1995; (49): S9-S15.

Scully RE. Histological typing of ovarian tumours - World Health Organization International histological classification of tumors. Springer-Verlag, Berlin 1999.

Scully RE. Influence of origin of ovarian cancer on efficacy of screening. Lancet 2000; (355): 1028-9.

Seltzer V. Laparoscopic surgery for ovarian lesions: potential pitfalls. Clin Obstet Gynecol 1993; (36): 402-412.

Sevelda P, Vavra N, Schemper M, Salzer H. Prognostic factors for survival in stage I epithelial ovarin carcinoma. Cancer 1990; (65): 2349-52.

Sijmos EA, van Lankveld AL, Witteveen PO, Peeters PHM, Kooot VCM, van Leeuwen JS. Compliance to clinical guidelines for early-stage epithelial ovarian cancer in relation to patients outcome. Eur J Obstet Gynecol Reprod Biol 2007; (131): -203.

Sjovall K, Nilsson B, Einhorn N. Different types of rupture of the tumor capsule and the impact on survival in early ovarian carcinoma. Int J Gynecol Cancer 1994; (4): 333-336.

Slavutin L, Rotterdam HZ. Frozen section diagnosis of serous epithelial tumors of the ovary. Am J Diagnostic Gynecol Obstet 1979; (1): 89-94.

Smidt V J, Singh D M, Hurteau J A, Hurd W W. Effect of carbon dioxide on human ovarian carcinoma cell growth. Am J Obstet Gynecol 2001; (185): 1314-1317.

Smith-Bindman R, Kerlikowske K, Feldstein V A, Subak L, Scheidler J, Segal M, Brand R, Grady D. Endovaginal ultrasound to exclude endometrial cancer and other endometrial abnormalities. JAMA 1998; (280): 1510-1517.

Soper J T, Hunter V J, Daly L, Tanner M, Creasman W T, Bast R C, Jr. Preoperative serum tumor-associated antigen levels in women with pelvic masses. Obstet Gynecol 1990; (75): 249-254.

Sövall K, Nilsson B, Einhorn N. Different types of rupture of the tumor capsule and the impact on survival in early ovarian carcinoma. Int J Gynecol Cancer 1994; (4): 333-6.

Spann C O, Kennedy J E, Musoke E. Intraoperative consultation of ovarian neoplasms. J Natl Med Assoc 1994; (86): 141-144.

Spanos W J. Preoperative hormonal therapy of cystic adnexal masses. Trans Pac Coast Obstet Gynecol Soc 1972; (40): 111-116.

Stroup D F, Berlin J A, Morton S C, Olkin I, Williamson G D, Rennie D, Moher D, Becker B J, Sipe T A, Thacker S B. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. JAMA 2000; (283): 2008-2012.

Sugarbaker TA, Chang D, Koslowe P S P. Pathobiology of peritoneal carcinomatosis from ovarian malignancy. Cancer Treat Res 1996; (81): 63-74.

Suttom AJ A K J D, Sheldon TA, Song F. Randon effects methods for combining study estimates. In: Methods for Meta-analysis in Medical Research. (Eds.Suttom AJ AKJD, Sheldon TA, Song F). Chichester: John Wiley, 2000; 1nd ed: 73-86.

Sutton AJ, Abrams KR, Jones DR, Sheldon TA, Song F. Meta-analysis of different types of data. In: Methods for Meta-Analysis in Medical Research. (Eds.Sutton AJ, Abrams KR, Jones DR, Sheldon TA, Song F). Chichester: John Wiley & Sons, 2000; 1: 205-228.

Sutton AJ. Randon effects methods for combining study estimates. In: Methods for Meta-Analysis in Medical Research. (Ed.Sutton AJ AKJDSTSF). Chichester: John Wiley, 2000; 1: 73-86.

Tailor A, Jurkovic D, Bourne T H, Collins W P, Campbell S. Sonographic prediction of malignancy in adnexal masses using multivariate logistic regression analysis. Ultrasound Obstet Gynecol 1997; (10): 41-47.

Torres JP, Suso JP, Perea E, Tafur L, Agudelo M. Tumores ováricos:correlación entre los informes de estudios solicitados por congelacion y la histopatologia definitiva. Hospital Universitário del Valle 1994-1997. Rev Colomb Obstet Gynecol 1998; (49): 149-151.

Tosteson A N, Begg C B. A general regression methodology for ROC curve estimation. Med Decis Making 1988; (8): 204-215.

Tozzi R, Köhler C, Ferrara A, Schneider A. Laparoscopic treatment of early ovarian cancer: surgical and survival outcomes. Gynecol Oncol 2004; (93): 199-203.

Tozzi R, Schneider A. Laparoscopy treatment of early ovarian cancer. Curr Opin Obstet Gynecol 2005.

Trimbos J B, Parmar M, Vergote I, Guthrie D, Bolis G, Colombo N, Vermorken J B, Torri V, Mangioni C, Pecorelli S, Lissoni A, Swart A M. International Collaborative Ovarian Neoplasm trial 1 and Adjuvant ChemoTherapy In Ovarian Neoplasm trial: two parallel randomized phase III trials of adjuvant chemotherapy in patients with early-stage ovarian carcinoma. J Natl Cancer Inst 2003; (95): 105-112.

Trimbos J B, van der Burg M E. [Adjuvant chemotherapy in patients operated on for early ovarian carcinoma]. Ned Tijdschr Geneeskd 2004; (148): 874-878.

Troidl H. Endoscopic surgery: innovation versus evaluation-introduction. World J Surg 1999; (23): 743-744.

Twaalfhoven F C, Peters A A, Trimbos J B, Hermans J, Fleuren G J. The accuracy of frozen section diagnosis of ovarian tumors. Gynecol Oncol 1991; (41): 189-192.

Twickler D M, Forte T B, Santos-Ramos R, McIntire D, Harris P, Scott D. The Ovarian Tumor Index predicts risk for malignancy. Cancer 1999; (86): 2280-2290.

Usubutun A, Altinok G, Kucukali T. The value of intraoperative consultation (frozen section) in the diagnosis of ovarian neoplasms. Acta Obstet Gynecol Scand 1998; (77): 1013-1016.

Vamvakas E C. Meta-analyses of studies of the diagnostic accuracy of laboratory tests: a review of the concepts and methods. Arch Pathol Lab Med 1998; (122): 675-686.

van N J, Jr., DePriest P D, Reedy M B, Gallion H H, Ueland F R, Pavlik E J, Kryscio R J. The efficacy of transvaginal sonographic screening in asymptomatic women at risk for ovarian cancer. Gynecol Oncol 2000; (77): 350-356.

Vasilev S A, Schlaerth J B, Campeau J, Morrow C P. Serum CA 125 levels in preoperative evaluation of pelvic masses. Obstet Gynecol 1988; (71): 751-756.

Velebil P, Wingo P A, Xia Z, Wilcox L S, Peterson H B. Rate of hospitalization for gynecologic disorders among reproductive-age women in the United States. Obstet Gynecol 1995; (86): 764-769.

Vergote I B, Amant F. Early ovarian cancer--time for a rethink on stage? Gynecol Oncol 2004; (94): 607-608.

Vergote I B, Bormer O P, Abeler V M. Evaluation of serum CA 125 levels in the monitoring of ovarian cancer. Am J Obstet Gynecol 1987; (157): 88-92.

Vergote I, De B J, Fyles A, Bertelsen K, Einhorn N, Sevelda P, Gore M E, Kaern J, Verrelst H, Sjovall K, Timmerman D, Vandewalle J, Van G M, Trope C G. Prognostic importance of degree of differentiation and cyst rupture in stage I invasive epithelial ovarian carcinoma. Lancet 2001; (357): 176-182.

Vergote I, Trimbos JB. Treatment of patients with early epithelial ovarian cancer. Curr Opin Oncol 2003; (15): 452-55.

Vergote I. Role of surgery in ovarian cancer: an update. Acta Chir Belg 2004; (104): 246-256.

Vessey M, Metcalfe A, Wells C, McPherson K, Westhoff C, Yeates D. Ovarian neoplasms, functional ovarian cysts, and oral contraceptives. Br Med J (Clin Res Ed) 1987; (294): 1518-1520.

Vinatier D, Dufour P, Cosson M, Querleu D. Laparoscopy in gynaecological cancer. Surg Oncol 1996; (5): 211-220.

Volz J, Koster S, Schaeff B, Paolucci V. Laparoscopic surgery: the effects of insufflation gas on tumor-induced lethality in nude mice. Am J Obstet Gynecol 1998; (178): 793-795.

Volz J, Köster S, Schaeff B. Laparoscopy managment of gynaecological malignancies: time to hesite. Gynaecological Endoscopy 1997; (6): 145-146.

Volz J, Koster S, Spacek Z, Paweletz N. The influence of pneumoperitoneum used in laparoscopic surgery on an intraabdominal tumor growth. Cancer 1999; (86): 770-774.

Volz J, Köster S. Laparoscopy:to inflate or lift. Cancer 1999; (86): 749-50.

von E E, Altman D G, Egger M, Pocock S J, Gotzsche P C, Vandenbroucke J P. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet 2007; (370): 1453-1457.

Wakahara F, Kikkawa F, Nawa A, Tamakoshi K, Ino K, Maeda O, Kawai M, Mizutani S. Diagnostic efficacy of tumor markers, sonography, and intraoperative frozen section for ovarian tumors. Gynecol Obstet Invest 2001; (52): 147-152.

Walter S D, Jadad A R. Meta-analysis of screening data: a survey of the literature. Stat Med 1999; (18): 3409-3424.

Wang P H, Yuan C C, Lin G, Ng H T, Chao H T. Risk factors contributing to early occurrence of port site metastases of laparoscopic surgery for malignancy. Gynecol Oncol 1999; (72): 38-44.

Weiner Z, Thaler I, Beck D, Rottem S, Deutsch M, Brandes J M. Differentiating malignant from benign ovarian tumors with transvaginal color flow imaging. Obstet Gynecol 1992; (79): 159-162.

Wells B, Shea B, O'coneell, Peterson J, Welch V, Losos M, Tugwell P. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analysis. Internet . 2007. Ref Type: Electronic Citation

Wenzl R, Lehner R, Husslein P, Sevelda P. Laparoscopic surgery in cases of ovarian malignancies: an Austria-wide survey. Gynecol Oncol 1996; (63): 57-61.

Westhoff C L, Beral V. Patterns of ovarian cyst hospital discharge rates in England and Wales, 1962-79. Br Med J (Clin Res Ed) 1984; (289): 1348-1349.

Wu C C, Lee C N, Chen T M, Shyu M K, Hsieh C Y, Chen H Y, Hsieh F J. Incremental angiogenesis assessed by color Doppler ultrasound in the tumorigenesis of ovarian neoplasms. Cancer 1994; (73): 1251-1256.

Yancik R. Ovarian cancer. Age contrasts in incidence, histology, disease stage at diagnosis, and mortality. Cancer 1993; (71): 517-523.

Yedema C, Massuger L, Hilgers J, Servaas J, Poels L, Thomas C, Kenemans P. Pre-operative discrimination between benign and malignant ovarian tumors using a combination of CA125 and CA15.3 serum assays. Int J Cancer Suppl 1988; (3): 61-67.

Yeo E L, Yu K M, Poddar N C, Hui P K, Tang L C. The accuracy of intraoperative frozen section in the diagnosis of ovarian tumors. J Obstet Gynaecol Res 1998; (24): 189-195.

Zamora J AV, Muriel A, Khan KS, Coomarasamy A. Meta-DiSc: a software for meta-analysis of test accuracy data. BMC Medical Research Methodology 6, 31. 2006.
Ref Type: Electronic Citation

Zamora J, Abraira V, Muriel A, Khan KS, Coomarasamy A. Meta-DiSc: a software for meta-analysis of test accuracy data. In: BMC Medical Research Methodology, 2006; 31.

ANEXO B

B.1 - Formulário para avaliação dos critérios de inclusão

B.2 - Avaliação da qualidade metodológica dos estudos

B.1 - Formulário para avaliação dos critérios de inclusão

Avaliação inicial dos estudos

Título resumido: _____ Data: ____ / ___ /_____

Referência(s): _____

Tipos de estudos

- O tratamento foi alocado aleatoriamente?
  Sim
  Não
  Indeterminado ☐

Tipos de participantes

- Os participantes eram apropriados para a resposta da pergunta da pesquisa?
  Sim
  Não
  Indeterminado ☐

Participantes	Diagnóstico

Tipos de intervenções

- A intervenção é claramente definida?
  Sim
  Não
  Indeterminado ☐

- Quais são os grupos de comparações?

Grupo tratamento (experimental)	Grupo Controle

Geração do sigilo da alocação

Categoria	Geração de sigilo
A	Significa que o processo de alocação foi adequadamente relatado (aleatorização centralizada por um escritório central ou farmácia; administração seqüencial de pacotes pré-codificados ou numerados aos pacientes selecionados para o estudo; sistema computadorizado on-line; dados gerados por um programa de computador contendo a distribuição codificada; envelopes seriados opacos e numerados; outras maneiras que pareçam oferecer uma alocação adequada, combinadas com o fato de que a pessoa que fez a alocação não esteja envolvida na sua utilização.
B	Significa: que o processo de alocação é incerto, isto é não foi descrito, mas é mencionado no texto que o estudo é aleatório (lista ou tabelas utilizadas; envelopes, mas não quantificando o tipo; uma alocação aparentemente adequada, mas com nenhuma outra informação no estudo).
C	Significa que o processo de alocação foi inadequadamente relatado (alternância; números de prontuários; data de nascimento; dias da semana)

B.2 - Avaliação da qualidade metodológica dos estudos

Índice para cada item deve ser:

Informa de forma correta – **A**
Incerto – **B**
Expõe claramente que não feito/ou tomado o cuidado – **C**

Validade Interna

Índice	Validade Interna de cada estudo
	O grupo tratamento (experimental) foi adequadamente cegado antes da alocação?
	Foram os desfechos dos pacientes "perdidos" ou excluídos após a alocação descritos e incluídos na avaliação final? ("intenção de tratar")
	Foram os desfechos avaliados cegamente para o status do estudo?
	Foram comparados entre si o grupo tratamento e o controle?
	Foram os sujeitos cegados para avaliação status do seguindo-se a alocação?
	Foram os pacientes do grupo tratamento alertados quando ao cegamento do status do estudo?
	Foram cuidadosamente programados, que outra opção de estudo, idêntica?
	As perdas foram <10% na população do estudo

Validade Externa

Índice	Validade Externa de cada estudo
	Foram os critérios de inclusão e exclusão claramente definidos entre si?
	Os desfechos mensurados foram claramente definidos?
	Foram a acurácia, a precisão e a variação de cada desfecho mensurada adequadamente?
	O tempo para avaliação do desfecho foi mensurado de forma apropriada?

ANEXO C

Níveis de evidências das publicações científicas (maio-2001)

Níveis de evidências das publicações científicas (maio-2001)

Nível	Terapia/Prevenção Etiologia/Dano	Prognóstico	Diagnóstico	Diagnóstico diferencial / estudos de prevalência	Econômico / Decisão de Análise
1a A	Estudos de RS (homogeneidade*) ou estudos ensaios clínicos randomizados (ECR)	Estudos de RS* (homogeneidade*) de estudos de coorte com controle desde o inicio dos casos; NDC† com validade em diferentes populações	Estudos de RS (homogeneidade*) de nível 1 em estudos diagnósticos; NDC† de estudos 1b de diferentes centros clínicos	Estudos de RS (homogeneidade*) ou estudos de coorte prospectivos	Estudos de RS* (homogeneidade*) ou nível 1 de estudos com enfoque econômico
1b A	Estudos individuais randomizados e controlados com estreito intervalo de confiança	Estudos individuais de coorte com > 80% de seguimento; NDC† validado em grupo populacional	Validação** de estudos de coorte com bom††† padrão de referência; NCD† testado em um único centro	Estudos de coorte prospectivos com bom seguimento****	Análise baseada em custos clínicos ou alternativas de custos; RS* de evidências; incluindo análises de sensibilidades de várias alternativas
1c A	Todos morriam antes do tratamento e alguns sobreviviam após início do tratamento, mas nenhum morria na vigência do tratamento§	Todos morriam antes do tratamento e alguns sobreviviam após início do tratamento, mas nenhum morria na vigência do tratamento§	Absoluta sensibilidade†† Absoluta especificidade††	Todos ou nenhum da série de casos	Estimativa de análise com absoluta estimativa de melhora ou piora††††
2a B	RS (homogeneidade*) de estudos de coorte	RS (homogeneidade*) ou outros estudos de coorte retrospectivo ou grupo controle dos grupos de estudos clínicos randomizados	RS (homogeneidade*) estudos diagnósticos de nível 2 ou com melhores níveis de evidência	RS (homogeneidade*) de 2b e estudos com melhores níveis de evidência	RS (homogeneidade*) de estudos com enfoque econômico com nível de evidência 2 ou com melhores níveis de evidência
2b B	Estudos de coorte individual (incluindo estudos randomizados de baixa qualidade, isto é, < 80% de seguimento)	Estudos de coorte retrospectivos ou de seguimento de grupo controle de pacientes tratados por estudos clínicos randomizados; derivados de NDC† que utilizam análise de regressão§§§	Estudos de coorte exploratórios** com bom††† padrão de referência (ouro); derivados de NDC† com análise de regressão dos dados§§§	Estudos de coorte retrospectivos, ou com pobre seguimento	Análise baseada em custos ou limitadas alternativas de revisão de evidências de estudos simples incluindo análise de sensibilidade de várias alternativas.
2c B	Estudos de desfechos; estudos ecológicos	Estudos de desfechos		Estudos ecológicos	Estudos de desfecho ou de auditoria
3a B	RS (homogeneidade*) de estudos de casos e controles		RS (homogeneidade*) de 3b	RS (homogeneidade*) de 3b	RS (homogeneidade*)de 3b
3b B	Estudos individuais de casos e controles		Estudos não consecutivos ou sem aplicar padrão-ouro de referência	Estudo de coorte não consecutivo, ou população muito limitada	Análise baseada em alternativas limitadas de custo, dados de estimativas muito pobres, mas incorporando análise de sensibilidade
4 C	Série de casos (ou estudos de coorte com pobre qualidade ou estudos de casos e controles §§)	Série de casos (com pobre qualidade de prognóstico estudos de coorte***	Estudos de casos e controles que dependem de padrão-ouro	Série de casos ou estudos que substituem o padrão-ouro	Análise de decisão com análise de sensibilidade
5 D	Opinião de especialista, sem explicitar uma avaliação crítica ou baseada em estudos de fisiolo-	Opinião de especialista, sem explicitar uma avaliação crítica ou baseada em estudos de fisiolo-	Opinião de especialista, sem explicitar uma avaliação crítica ou baseada em estudos de fisiologia ou	Opinião de especialista, sem explicitar uma avaliação crítica ou baseada em	Opinião de especialista, sem explicitar uma avaliação crítica ou baseada em estudos de fisiologia ou em estudos de

	gia ou em estudos de princípios iniciais	gia ou em estudos de princípios iniciais	em estudos de princípios iniciais	estudos de fisiologia ou em estudos de princípios iniciais	princípios iniciais

Fonte: PHILLIPS *et al*. (2001)

Níveis de evidências das publicações científicas (maio-2001)

NOTAS

*	RS = artigos de revisão sistemática, com homogeneidade entre os artigos, significa estar livre de artigos com variação heterogênea.
†	Normas de decisão clínica (NDC) representadas graficamente por algoritmos ou sistema de escores o qual fornece estimativa de diagnóstico ou prognóstico.
§	Quando todos os pacientes morriam antes de fazer o tratamento, mas agora alguns sobrevivem com início da terapêutica, ou quando alguns pacientes morrem antes de o tratamento tornar-se disponível, entretanto nenhum morre em vigência do tratamento.
§§	Estudos de coorte (com pobre qualidade) - falharam em definir a comparação entre os grupos e/ou falharam em mensurar exposição e desfecho (preferencialmente deveriam ser cegados); falharam em identificar grupo controle e fatores de confusão; o seguimento não foi suficientemente longo para avaliar desfecho, o seguimento dos pacientes não foi completo. Estudos de casos e controles (com pobre qualidade) – falham em definir claramente a comparação entre os grupos, falham em mensurar exposição e desfecho (preferencialmente deveriam ser cegados), falham em identificar grupo controle e fatores de confusão.
§§§	Estudos de validação testam a qualidade de um teste diagnóstico específico, com base em evidências prévias. Um estudo exploratório coleta informações e utiliza a análise de regressão para identificar fatores que sejam significativos.
††	Há absoluta especificidade (resultado negativo) quando exclui o diagnóstico. Há absoluta sensibilidade (teste positivo) quando o teste define o diagnóstico.
†††	Bom padrão é como se chama o “padrão-ouro”, são testes independentes e aplicadas às cegas objetivamente em todos os pacientes.
††††	Estimativa de tratamento de melhor valor são claramente os que possuem baixo custo. Estimativa de tratamento de menor valor pode ser uma opção boa, mas mais cara, também pode ser uma opção ruim com igual custo ou ainda mais cara.
**	Validando estudos de testes diagnósticos específicos, baseados anteriormente em evidências. Estudo de coletas de informações e análise de dados (utiliza análise de regressão) para encontrar fatores que possam ser considerados significantes.
***	Estudos de coorte, com enfoque de prognóstico, são considerados de pobre qualidade quando ocorre viés na seleção da amostra; mensuração do desfecho ocorre somente < 80% dos pacientes que concluem o estudo; quando os desfechos são determinados, mas não blindados e não há objetividade nem correção dos fatores de confusão.
****	Bom seguimento (> 80%) em estudos com diagnóstico diferencial, com adequado tempo de acompanhamento: em quadro agudo (1-6 meses) e em quadro crônico (1-5 anos)

Graus de recomendação (maio de 2001)

A	Consiste em estudos de nível 1. Estudo com forte recomendação na escolha são excelentes os níveis de evidência para recomendar rotineiramente a conduta. Os benefícios possuem peso maior que o dano. Há boas evidências para apoiar a recomendação.
B	Consiste em estudos do nível 2 e 3 ou generalização de estudos de nível 1. Estudo que recomenda a ação e são encontradas evidências importantes no desfecho e conclusão é de que há benefício na escolha da ação em relação aos riscos do dano. Há evidências razoáveis para apoiar a recomendação.
C	Consiste em estudos de nível 4 ou generalização de estudos de nível 2 ou 3. Encontra mínimas evidências satisfatórias na análise dos desfechos, mas conclui que os benefícios e os riscos do procedimento não justificam a generalização da recomendação. Há evidências insuficientes, contra ou a favor, mas as recomendações podem ter outras bases.
D	Consiste em estudos de nível 5 ou qualquer estudo inconclusivo. Estudos com pobre qualidade. Há evidências para descartar a recomendação.

Fonte: Adaptado de PHILLIPS *et al.* (2001)

ANEXO D

Accurate Reporting of Studies of Diagnostic Acuracy: The STARD Iniciative

Empirical Evidence of Design-Related Bias in Studies of Diagnostic Tests

Annals of Internal Medicine

Towards Complete and Accurate Reporting of Studies of Diagnostic Accuracy: The STARD Initiative

ISSN: 0003-4819

Accession: 00000605-200301070-00010

Author(s): Bossuyt, Patrick M.; Reitsma, Johannes B.; Bruns, David E.; Gatsonis, Constantine A.; Glasziou, Paul P.; Irwig, Les M.; Lijmer, Jeroen G.; Moher, David; Rennie, Drummond; de Vet, Henrica C.W.; for the STARD Group

Issue: Volume 138(1), 7 January 2003, pp 40-44

Abstract

Background: To comprehend the results of diagnostic accuracy studies, readers must understand the design, conduct, analysis, and results of such studies. That goal can be achieved only through complete transparency from authors.

Objective: To improve the accuracy and completeness of reporting of studies of diagnostic accuracy in order to allow readers to assess the potential for bias in the study and to evaluate its generalizability.

Methods: The Standards for Reporting of Diagnostic Accuracy (STARD) steering committee searched the literature to identify publications on the appropriate conduct and reporting of diagnostic studies and extracted potential items into an extensive list. Researchers, editors, methodologists and statisticians, and members of professional organizations shortened this list during a 2-day consensus meeting with the goal of developing a checklist and a generic flow diagram for studies of diagnostic accuracy.

Results: The search for published guidelines on diagnostic research yielded 33 previously published checklists, from which we extracted a list of 75 potential items. The consensus meeting shortened the list to 25 items, using evidence on bias whenever available. A prototypical flow diagram provides information about the method of patient recruitment, the order of test execution, and the numbers of patients undergoing the test under evaluation, the reference standard, or both.

Conclusions: Evaluation of research depends on complete and accurate reporting. If medical journals adopt the checklist and the flow diagram, the quality of reporting of studies of diagnostic accuracy should improve to the advantage of the clinicians, researchers, reviewers, journals, and the public.

The world of diagnostic tests is highly dynamic. New tests are developed at a fast rate and the technology of existing tests is continuously being improved. Exaggerated and biased results from poorly designed and reported diagnostic studies can trigger their premature dissemination and lead physicians into making incorrect treatment decisions. A rigorous evaluation process of diagnostic tests before introduction into clinical practice could not only reduce the number of unwanted clinical consequences related to misleading estimates of test accuracy, but also limit health care costs by preventing unnecessary testing. Studies to determine the diagnostic accuracy of a test are a vital part in this evaluation process (1-3).

In studies of diagnostic accuracy, the outcomes from one or more tests under evaluation are compared with outcomes from the reference standard, both measured in subjects who are suspected of having the condition of interest. The term test refers to any method for obtaining additional information on a patient's health status. It includes information from history and physical examination, laboratory tests, imaging tests, function tests, and histopathology. The condition of interest or target condition can refer to a particular disease or to any other identifiable condition that may prompt clinical actions, such as further diagnostic testing, or the initiation, modification, or termination of treatment. In

this framework, the reference standard is considered to be the best available method for establishing the presence or absence of the condition of interest. The reference standard can be a single method, or a combination of methods, to establish the presence of the target condition. It can include laboratory tests, imaging tests, and pathology, but also dedicated clinical follow-up of subjects. The term accuracy refers to the amount of agreement between the information from the test under evaluation, referred to as the index test, and the reference standard. Diagnostic accuracy can be expressed in many ways, including sensitivity and specificity, likelihood ratios, diagnostic odds ratio, and the area under a receiver-operator characteristic (ROC) curve (4-6).

There are several potential threats to the internal and external validity of a study of diagnostic accuracy. A survey of studies of diagnostic accuracy published in four major medical journals between 1978 and 1993 revealed that the methodological quality was mediocre at best (7). However, evaluations were hampered because many reports lacked information on key elements of design, conduct, and analysis of diagnostic studies (7). The absence of critical information about the design and conduct of diagnostic studies has been confirmed by authors of meta-analyses (8, 9). As in any other type of research, flaws in study design can lead to biased results. One report showed that diagnostic studies with specific design features are associated with biased, optimistic estimates of diagnostic accuracy compared to studies without such deficiencies (10).

At the 1999 Cochrane Colloquium meeting in Rome, the Cochrane Diagnostic and Screening Test Methods Working Group discussed the low methodological quality and substandard reporting of diagnostic test evaluations. The Working Group felt that the first step to correct these problems was to improve the quality of reporting of diagnostic studies. Following the successful CONSORT (Consolidated Standards of Reporting Trials) initiative (11-13), the Working Group aimed at the development of a checklist of items that should be included in the report of a study of diagnostic accuracy.

The objective of the Standards for Reporting of Diagnostic Accuracy (STARD) initiative is to improve the quality of reporting of studies of diagnostic accuracy. Complete and accurate reporting allows the reader to detect the potential for bias in the study (internal validity) and to assess the generalizability and applicability of the results (external validity).

Methods

The STARD steering committee (see Appendix for membership and details) started with an extensive search to identify publications on the conduct and reporting of diagnostic studies. This search included MEDLINE, EMBASE, BIOSIS, and the methodological database from the Cochrane Collaboration up to July 2000. In addition, the steering committee members examined reference lists of retrieved articles, searched personal files, and contacted other experts in the field of diagnostic research. They reviewed all relevant publications and extracted an extended list of potential checklist items.

Subsequently, the STARD steering committee convened a 2-day consensus meeting for invited experts from the following interest groups: researchers, editors, methodologists, and professional organizations. The aim of the conference was to reduce the extended list of potential items, where appropriate, and to discuss the optimal format and phrasing of the checklist. The selection of items to retain was based on evidence whenever possible.

The meeting format consisted of a mixture of small group sessions and plenary sessions. Each small group focused on a group of related items of the list. The suggestions of the small groups were then discussed in plenary sessions. Overnight, a first draft of the STARD checklist was assembled based on the suggestions from the small group and the additional remarks from the plenary sessions. All meeting attendees discussed this version the next day and made additional changes. The members of the STARD group could suggest further changes through a later round of comments by electronic mail.

Potential users field-tested the conference version of the checklist and flow diagram and additional comments were collected. This version was placed on the CONSORT Web site with a call for

comments. The STARD steering committee discussed all comments and assembled the final checklist.

Results

The search for published guidelines for diagnostic research yielded 33 lists. Based on these published guidelines and on input of steering and STARD group members, the steering group assembled a list of 75 items. During the consensus meeting on 16 and 17 September 2000, participants consolidated and eliminated items to form the 25-item checklist. Conference members made major revisions to the phrasing and format of the checklist.

The STARD group received valuable comments and remarks during the various stages of evaluation after the conference, which resulted in the version of the STARD checklist that appears in the Table.

[Help with image viewing]
[Email Jumpstart To Image]

Table. STARD Checklist for the Reporting of Studies of Diagnostic Accuracy

The flow diagram provides information about the method of patient recruitment (e.g., based on a consecutive series of patients with specific symptoms, case–control), the order of test execution, and the number of patients undergoing the test under evaluation (index test) and the reference test (Figure). We provide one prototypical flow chart that reflects the most commonly employed design in diagnostic research. Examples that reflect other designs are on the STARD Web site (see www.consort-statement.org\stardstatement.htm) .

Figure. Prototypical flow diagram of a diagnostic accuracy study.

[Help with image viewing]
[Email Jumpstart To Image]

Discussion

The purpose of the STARD initiative is to improve the quality of the reporting of diagnostic studies. The items in the checklist and the flow chart can help authors in describing essential elements of the design and conduct of the study, the execution of tests, and the results.

We arranged the items under the usual headings of a medical research article but this is not intended to dictate the order in which they have to appear within an article.

The guiding principle in the development of the STARD checklist was to select items that would help readers to judge the potential for bias in the study and to appraise the applicability of the findings. Two other general considerations shaped the content and format of the checklist. First, the STARD group believes that one general checklist for studies of diagnostic accuracy, rather than different checklists for each field, is likely to be more widely disseminated and perhaps accepted by authors, peer reviewers, and journal editors. Although the evaluation of imaging tests differs from that of tests in the laboratory, we felt that these differences were more in degree than of kind. The second consideration was the development of a checklist specifically aimed at studies of diagnostic accuracy. We did not include general issues in the reporting of research findings, like the recommendations contained in the Uniform Requirements for Manuscripts Submitted to Biomedical Journals (14).

Wherever possible, the STARD group based the decision to include an item on evidence linking the item to biased estimates (internal validity) or to variation in measures of diagnostic accuracy (external validity). The evidence varied from narrative articles explaining theoretical principles and papers presenting results from statistical modeling to empirical evidence derived from diagnostic studies. For several items, the evidence is rather limited.

A separate background document, available at http://www.annals.org , explains the meaning and rationale of each item and briefly summarizes the type and amount of evidence (15). This background document should enhance the use, understanding, and dissemination of the STARD checklist.

The STARD group put considerable effort into the development of a flow diagram for diagnostic studies. A flow diagram has the potential to communicate vital information about the design of a study and the flow of participants in a transparent manner (16). A comparable flow diagram has become an essential element in the CONSORT standards for reporting of randomized trials (12, 16). The flow diagram could be even more essential in diagnostic studies, given the variety of designs employed in diagnostic research. Flow diagrams in the reports of diagnostic accuracy studies indicate the process of sampling and selecting participants (external validity), the flow of participants in relation to the timing and outcomes of tests, the number of subjects who fail to receive either the index test and/or the reference standard (potential for verification bias [17-19]), and the number of patients at each stage of the study, thus providing the correct denominator for proportions (internal consistency).

The STARD group plans to measure the impact of the statement on the quality of published reports on diagnostic accuracy using a before-and-after evaluation (13). Updates of STARD will be provided when new evidence on sources of bias or variability becomes available. We welcome any comments, whether on content or form, to improve the current version.

Appendix

Members of the STARD Steering Committee

Patrick Bossuyt, Academic Medical Center, Department of Clinical Epidemiology, Amsterdam, the Netherlands; David Bruns, Clinical Chemistry, Washington, D.C., United States of America; Constantine Gatsonis, Brown University, Centre for Statistical Sciences, Providence, Rhode Island, United States of America; Paul Glasziou, Mayne Medical School, Department of Social and Preventive Medicine, Herston, Australia; Les Irwig, University of Sydney, Department of Public Health and Community Medicine, Sydney, Australia; Jeroen Lijmer, Academic Medical Center, Department of Clinical Epidemiology, Amsterdam, the Netherlands; David Moher, Chalmers Research Group, Ottowa, Ontario, Canada; Drummond Rennie, Journal of the American Medical Association, Chicago, Illinois, United States of America; and Riekie de Vet, Free University, Institute for Research in Extramural Medicine, Amsterdam, the Netherlands.

Members of the STARD Group

Doug Altman, Institute of Health Sciences, Centre for Statistics in Medicine, Oxford, United Kingdom; Stuart Barton, British Medical Journal, BMA House, London, United Kingdom; Colin Begg, Memorial Sloan-Kettering Cancer Center, Department of Epidemiology and Biostatistics, New York, New York, United States of America; William Black, Dartmouth-Hitchcock Medical Center, Department of Radiology, Lebanon, New Hampshire, United States of America; Harry Büller, Academic Medical Center, Department of Vascular Medicine, Amsterdam, the Netherlands; Gregory Campbell, U.S. Food and Drug Administration, Center for Devices and Radiological Health, Rockville, Maryland, United States of America; Frank Davidoff, Annals of Internal Medicine, Philadelphia, Pennsylvania, United States of America; Jon Deeks, Institute of Health Sciences, Centre for Statistics in Medicine, Old Road, United Kingdom; Paul Dieppe, Department of Social Medicine, University of Bristol, Bristol, United Kingdom; Kenneth Fleming, John Radcliffe Hospital, Oxford, United Kingdom; Rijk van Ginkel, Academic Medical Center, Department of Clinical Epidemiology, Amsterdam, the Netherlands; Afina Glas, Academic Medical Center, Department of Clinical Epidemiology, Amsterdam, the Netherlands; Gordon Guyatt, McMaster University, Clinical Epidemiology and Biostatistics, Hamilton, Canada; James Hanley, McGill University, Department of Epidemiology and Biostatistics, Montreal, Canada; Richard Horton, The Lancet, London, United Kingdom; Myriam Hunink, Erasmus Medical Center, Department of Epidemiology and Biostatistics, Rotterdam, the Netherlands; Jos Kleijnen, National Health Services Centre for Reviews and Dissemination, York, United Kingdom; Andre Knottnerus, Maastricht University, Netherlands School of Primary Care Research, Maastricht, the Netherlands; Erik Magid, Amager Hospital, Department of Clinical Biochemistry, Copenhagen, Denmark; Barbara McNeil, Harvard Medical School, Department of Health Care Policy, Boston, Massachusetts, United States of America; Matthew McQueen, Hamilton Civic Hospitals, Department of Laboratory Medicine, Hamilton, Canada; Andrew Onderdonk, Channing Laboratory, Boston, Massachusetts, United States of America; John Overbeke, Nederlands Tijdschrift voor Geneeskunde, Amsterdam, the Netherlands; Christopher Price, St. Bartholomew's—Royal London School of Medicine and Dentistry, London, United Kingdom; Anthony Proto, Radiology Editorial Office, Richmond, United States of America; Hans Reitsma, Academic Medical Center, Department of Clinical Epidemiology, Amsterdam, the Netherlands; David Sackett, Trout Research and Education Centre, Irish Lake, Ontario, Canada; Gerard Sanders, Academic Medical Center, Department of Clinical Chemistry, Amsterdam, the Netherlands; Harold Sox, Annals of Internal Medicine, Philadelphia, Pennsylvania, United States of America; Sharon Straus, Mt. Sinai Hospital, Toronto, Canada; and Stephan Walter, McMaster University, Clinical Epidemiology and Biostatistics, Hamilton, Canada. [Context Link]

References

1.Guyatt GH, Tugwell PX, Feeny DH, Haynes RB, Drummond M. A framework for clinical evaluation of diagnostic technologies. CMAJ. 1986;134:587-94. [PMID: 3512062] Bibliographic Links [Context Link]

2.Fryback DG, Thornbury JR. The efficacy of diagnostic imaging. Med Decis Making. 1991;11:88-94. [PMID: 1907710] Bibliographic Links [Context Link]

3.Kent DL, Larson EB. Disease, level of impact, and quality of research methods. Three dimensions of clinical efficacy assessment applied to magnetic resonance imaging. Invest Radiol. 1992;27:245-54. [PMID: 1551777] Buy Now Bibliographic Links [Context Link]

4.Griner PF, Mayewski RJ, Mushlin AI, Greenland P. Selection and interpretation of diagnostic tests and procedures. Principles and applications. Ann Intern Med. 1981;94:557-92. [PMID: 6452080] Bibliographic Links [Context Link]

5.Sackett DL, Haynes RB, Guyatt GH, Tugwell P. The selection of diagnostic tests. In: Sackett D, ed. Clinical Epidemiology. 2nd ed. Boston/Toronto/London: Little, Brown; 1991:47-57. [Context Link]

6.Metz CE. Basic principles of ROC analysis. Semin Nucl Med. 1978;8:283-98. [PMID: 112681] Bibliographic Links [Context Link]

7.Reid MC, Lachs MS, Feinstein AR. Use of methodological standards in diagnostic test research. Getting better but still not good. JAMA. 1995;274:645-51. [PMID: 7637146] Bibliographic Links [Context Link]

8.Nelemans PJ, Leiner T, de Vet HC, van Engelshoven JM. Peripheral arterial disease: meta-analysis of the diagnostic performance of MR angiography. Radiology. 2000;217:105-14. [PMID: 11012430] Bibliographic Links [Context Link]

9.de Vries SO, Hunink MG, Polak JF. Summary receiver operating characteristic curves as a technique for meta-analysis of the diagnostic performance of duplex ultrasonography in peripheral arterial disease. Acad Radiol. 1996;3:361-9. [PMID: 8796687] Bibliographic Links [Context Link]

10.Lijmer JG, Mol BW, Heisterkamp S, Bonsel GJ, Prins MH, van der Meulen JH, et al. Empirical evidence of design-related bias in studies of diagnostic tests. JAMA. 1999;282:1061-6. [PMID: 10493205] [Context Link]

11.Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, et al. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. JAMA. 1996;276:637-9. [PMID: 8773637] [Context Link]

12.Moher D, Schulz KF, Altman D. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. JAMA. 2001;285:1987-91. [PMID: 11308435] Bibliographic Links [Context Link]

13.Moher D, Jones A, Lepage L. Use of the CONSORT statement and quality of reports of randomized trials: a comparative before-and-after evaluation. JAMA. 2001;285:1992-5. [PMID: 11308436] Bibliographic Links [Context Link]

14.Uniform requirements for manuscripts submitted to biomedical journals. International Committee of Medical Journal Editors. JAMA. 1997;277:927-34. [PMID: 9062335] Also available at http://www.acponline.org. [Context Link]

15.Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. Clin Chem. 2003;49:7-18. [Context Link]

16.Egger M, Jüni P, Bartlett C. Value of flow diagrams in reports of randomized controlled trials. JAMA. 2001;285:1996-9. [PMID: 11308437] Bibliographic Links [Context Link]

17.Knottnerus JA. The effects of disease verification and referral on the relationship between symptoms and diseases. Med Decis Making. 1987;7:139-48. [PMID: 3613914] Bibliographic Links [Context Link]

18.Panzer RJ, Suchman AL, Griner PF. Workup bias in prediction research. Med Decis Making. 1987;7:115-9. [PMID: 3574021] Bibliographic Links [Context Link]

19.Begg CB. Biases in the assessment of diagnostic tests. Stat Med. 1987;6:411-23. [PMID: 3114858] [Context Link]

Empirical Evidence of Design-Related Bias in Studies of Diagnostic Tests

Jeroen G. Lijmer, MD; Ben Willem Mol, MD, PhD; Siem Heisterkamp, PhD; Gouke J. Bonsel, MD, PhD; Martin H. Prins, MD, PhD; Jan H. P. van der Meulen, MD, PhD; Patrick M. M. Bossuyt, PhD

JAMA. 1999;282:1061-1066.

ABSTRACT

Context The literature contains a large number of potential biases in the evaluation of diagnostic tests. Strict application of appropriate methodological criteria would invalidate the clinical application of most study results.

Objective To **empirical**ly determine the quantitative effect of study design shortcomings on estimates of diagnostic accuracy.

Design and Setting Observational study of the methodological features of 184 original studies evaluating 218 diagnostic tests. Meta-analyses on diagnostic tests were identified through a systematic search of the literature using MEDLINE, EMBASE, and DARE databases and the Cochrane Library (1996-1997). Associations between study characteristics and estimates of diagnostic accuracy were evaluated with a regression model.

Main Outcome Measures Relative diagnostic odds ratio (RDOR), which compared the diagnostic odds ratios of studies of a given test that lacked a particular methodological feature with those without the corresponding shortcomings in design.

Results Fifteen (6.8%) of 218 evaluations met all 8 criteria; 64 (30%) met 6 or more. Studies evaluating tests in a diseased population and a separate control group overestimated the diagnostic performance compared with studies that used a clinical population (RDOR, 3.0; 95% confidence interval [CI], 2.0-4.5). Studies in which different reference tests were used for positive and negative results of the test under study overestimated the diagnostic performance compared with studies using a single reference test for all patients (RDOR, 2.2; 95% CI, 1.5-3.3). Diagnostic performance was also overestimated when the reference test was interpreted with knowledge of the test result (RDOR, 1.3; 95% CI, 1.0-1.9), when no criteria for the test were described (RDOR, 1.7; 95% CI, 1.1-2.5), and when no description of the population under study was provided (RDOR, 1.4; 95% CI, 1.1-1.7).

Conclusion These data provide **empirical** evidence that diagnostic studies with methodological shortcomings may overestimate the accuracy of a diagnostic test, particularly those including nonrepresentative patients or applying different reference standards.

INTRODUCTION

During recent decades, the number of available diagnostic tests has been rapidly increasing. As for all new medical technologies, new diagnostic tests should be thoroughly evaluated prior to their introduction into daily practice. The number of test evaluations in the literature is increasing but the methodological quality of these studies is on average poor. A survey of the diagnostic literature (1990-1993) showed that only 18% of the studies satisfied 5 of the 7 methodological standards examined.[1] Different guidelines have been written to help physicians with the critical appraisal of the diagnostic literature consisting of lists of criteria for the assessment of study quality.[2-4] Criteria enable readers to check whether studies fulfill methodological criteria on study design, data collection, and methods of reporting the results.

As few diagnostic studies meet all of the methodological criteria, physicians and reviewers are faced with a difficult choice. Strict application of the methodological criteria would imply that only a small minority of the available data can be used in clinical practice. Alternatively, inclusion of a wider range of imperfect studies would require weighting of the evidence according to the relative importance of the criteria that such studies failed to satisfy. One article has reported such weights for methodological criteria. Unfortunately, these weights were established through a consensus procedure in a general internal medicine division at an academic medical center, rather than on **empirical** data.[5] A data-driven approach had been previously used by Schulz et al[6] when evaluating the influence of study design features on estimates of treatment effects in randomized controlled trials.

The purpose of our study was to assess **empirical**ly the impact of shortcomings in design, data collection, and reporting on the estimates of diagnostic accuracy. We compared estimates of diagnostic accuracy for a given test reported in studies with lower quality with estimates for the same test from studies without these shortcomings. We hypothesized that estimates of diagnostic accuracy would be exaggerated in studies that failed to meet methodological standards.

METHODS

Data Sources and Data Extraction

An electronic search of the literature was performed to identify meta-analyses summarizing the accuracy of diagnostic tests. We focused on meta-analyses because they enabled us to identify a large number of studies on a single diagnostic problem. We concentrated on recent meta-analyses as we expected these meta-analyses to include both older studies, using suboptimal designs, as well as recent studies, applying a more up-to-date approach that lives up to current methodological standards. To be included, a meta-analysis had to be based on a systematic search of the literature, had to include at least 5 studies, and had to report sensitivities and specificities of included studies. The latter criterion was introduced to assure that from each reviewed study sensitivity and specificity were available and to allow for easy replication of our work.

The MEDLINE and EMBASE databases were searched (January 1996 to December 1997) using combinations of the words *meta-analysis; diagnostic imaging; diagnostic tests, routine*; *sensitivity and specificity*; and *review,* publication type. In addition, the Cochrane Library and the DARE database of the NHS Centre for Reviews and Dissemination were examined for relevant abstracts.

We retrieved 26 articles that included 5 or more studies. Fifteen articles had to be excluded: 7 were not based on a systematic literature search and 8 reported no list of sensitivities and specificities. A list of excluded articles is available from the authors.

For the 11 remaining articles, all original papers included in the analyses were retrieved (Table 1). The characteristics of these studies were extracted on a standard form by 1 of the authors (J.G.L.). The set of characteristics on the standard form was based on a synthesis from different lists of criteria for study quality.[2, 4-5] All studies were independently scored a second time by a second reviewer (B.W.M., P.M.M.B.). Disagreement was resolved by consensus, if necessary the judgment of a third reviewer was decisive.

View this table: [in this window] [in a new window]	**Table 1.** Diagnostic Problems, Tests, Number of Studies, and Search Period of the Meta-Analyses

Assessment of Study Quality

The optimal design for assessing the accuracy of a diagnostic test is considered to be a prospective blind comparison of the test and the reference test in a consecutive series of patients from a relevant clinical population.[2, 7] A relevant clinical population is a group of patients covering the spectrum of disease that is likely to be encountered in the current or future use of the test. There are several threats to the validity of a diagnostic study. Diagnostic accuracy can be overestimated if the test is evaluated in a group of patients already known to have the disease and a separate group of normal patients, rather than in a relevant clinical population.[8] This will be referred to as a case-control study.

Selection bias can be present when not all patients presenting with the relevant condition are included in order of entry (consecutive) into the study, and when this selection is not random. If it was not clear from the text that a consecutive series of patients was included or a random subset, the corresponding study was scored as nonconsecutive.

Verification bias looms if the decision to perform the reference test is based on the result of the test under examination. In many diagnostic studies with an invasive reference test, most of the positive test results and only a small part of the negative test results are verified. Alternatively, negative test results are verified by a different, often less thorough, standard, for example follow-up. We will refer to these 2 forms of verification bias as partial verification bias and differential reference standard bias, respectively. In cases in which more than 10% of the study group was not subjected to the reference test, the study was scored as applying partial verification; in cases in which different reference tests were used, the study was scored as differential reference standard. All other cases were scored as complete verification.

Interpreting the reference test with knowledge of the results of the test under study can lead to an overestimation of a test's accuracy, especially if the reference test is open to subjective interpretation. If the sequence of testing is reversed, it is important that the results of the test under study are interpreted without knowledge of the reference test. If it was not clear from the text that the interpretation of both tests was done while investigators were blinded the study was scored as not blinded.

In addition to characteristics of the study design, we also looked at methods of data collection and reporting. The data collection was categorized as either prospective or retrospective. In case of doubt, the method of data collection was scored as unknown. The reference test, the test under study, and the study population should be described with sufficient detail to allow for replication, validation, and generalization of the study.[2] Descriptions of the tests were scored as sufficient if clear definitions of positive and negative test results were mentioned in the text. Description of the study population was sufficient if 2 of the following characteristics were described: age of participants, female to male ratio, or distribution of symptoms.

Statistical Analysis

The results of an individual study on diagnostic accuracy can be summarized in a 2x2 table. From this table, frequently used measures such as sensitivity, specificity, and predictive values can easily be calculated.[9-10] Another measure for the diagnostic accuracy of a test is the diagnostic odds ratio (DOR), the odds for a positive test result in diseased persons relative to the odds of a positive result in nondiseased persons.[11-12] The DOR is a single statistic of the results in a 2x2 table, incorporating sensitivity as well as specificity. Expressed in terms of sensitivity and specificity the formula is:

$$\mathrm{DOR} = \frac{\dfrac{\text{Sensitivity}}{(1 - \text{Sensitivity})}}{\dfrac{(1 - \text{Sensitivity})}{\text{Specificity}}}$$

The effect of study characteristics was examined with a regression model that is adapted from the summary receiver operating characteristic curve model, developed for meta-analyses of diagnostic tests.[13-15] The basic model contains the logarithm of the diagnostic odds ratio computed for a single study as a dependent variable and 2 explaining parameters, 1 for the intercept and 1 for the slope of

the curve, for each meta-analysis. The intercept can be interpreted as the common DOR of the corresponding test and the parameter for the slope expresses variation of the DOR across individual studies due to threshold differences.

We added covariates to this model to examine whether, on average, studies that failed to meet the methodological criteria yielded different DORs. The resulting parameter estimates of the covariates can be interpreted after antilogarithm transformation as relative DORs (RDORs). They indicate the diagnostic performance of a test in studies failing to satisfy the methodological criterion, relative to its performance in studies with the corresponding feature. If the RDOR is larger than 1, studies not satisfying the criterion yield larger estimates of the DOR than studies with this corresponding feature.

In summary, the dependent variable of the model was the logarithm DOR. Explaining variables were 2 parameters for each meta-analysis (the common DOR and the threshold parameter) and 9 covariates to examine the effect of the different study characteristics, 1 for each feature. All study characteristics were evaluated simultaneously in a multivariate model.

A weighted linear regression analysis was used, with weights proportional to the reciprocal of the variance of the log DOR. This weighted linear regression assumes fixed effects. In case of zero entries, the DOR is not defined. This problem was solved by adding 0.5 to all cells of the 2x2 table for all studies in a meta-analysis.[14-16] The model was fitted using maximum likelihood estimation, and programmed using statistical software (S-plus 4.5, Mathsoft Inc, Cambridge, Mass).

RESULTS

The subjects of the 11 included articles and the reviewed tests are summarized in Table 1. Two articles[17-18] reviewed 3 tests, 3 articles[19-21] reviewed 2 tests, and 6 articles[22-27] reviewed 1 test. This resulted in a total of 18 separate meta-analyses for this analysis. These 18 meta-analyses summarized the results of 193 published studies. Nine studies could not be used in the final analysis because only abstracts were available for 4 and 2x2 table calculations were not possible for 5. Of the 184 studies remaining, some evaluated multiple diagnostic tests. A total of 218 diagnostic test evaluations were available for analysis.

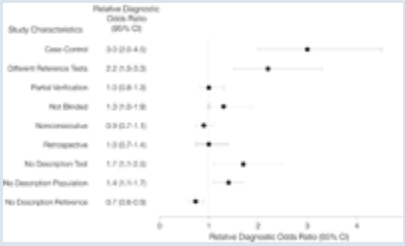
The overall results of the quality assessment of the included studies are listed in Table 2. Most studies used a clinical cohort and described the cut-off that the test evaluated (98% and 89%, respectively). Only 15 (6.8%) of the 218 studies satisfied all 8 criteria used. Sixty-four (30%) of the 218 studies satisfied 6 or more criteria.

View this table: [in this window] [in a new window]	**Table 2.** Results of the Scoring of Study Quality (N = 218)

Table 2. Results of the Scoring of Study Quality (N = 218)

Study Characteristic	Score	No. of Subjects	
Spectrum	Clinical population	213	
	Case-control	5	
Verification	Complete	106	
	Different reference tests	48	
	Partial	54	
Interpretation of test results	Blinded	70	
	Not blinded	148	
Patient selection	Consecutive	97	Value
	Nonconsecutive	121	
collection	Prospective	98	Data
	Unknown	92	
	Retrospective	28	
ls test	Sufficient	195	Detai
	Insufficient	23	
ls reference test	Sufficient	138	Detai
	Insufficient	80	
ls population	Sufficient	132	Detai
	Insufficient	86	

The results from the regression analysis are presented in Figure 1. Studies using a case-control design tended to overestimate the DOR 3-fold compared with studies with a clinical cohort (RDOR, 3.0). Studies using different reference tests for positive and negative test results had an RDOR of 2.2, showing approximately a 2-fold overestimation of the DOR compared with studies that used 1 reference test. Studies verifying only part of the population had on average the same DOR as studies that subjected all patients to the reference test (RDOR, 1.0). There were no studies verifying only part of the population and in conjunction using different reference standards. Interpretation of the reference test with knowledge of the outcomes of the test under study resulted in a RDOR of 1.3, causing an overestimation of the DOR by approximately 30% compared with studies with adequate blinding. Selective inclusion of patients into the study did not change the estimation of diagnostic accuracy significantly.



View larger version (14K):
[in this window]
[in a new window]

Figure. Relative Diagnostic Odds Ratios and 95% Confidence Intervals (CIs) of the 9 Study Characteristics Examined With a Multivariate Regression Analysis

Retrospective data collection was not associated with an overestimation or underestimation of diagnostic accuracy in comparison with studies with prospective or unknown data collection. In a

univariate analysis (data not shown) the RDOR of studies with unknown data collection was nearest to that of prospective studies, we therefore collapsed these 2 categories. The DORs in articles without a sufficient description of the test under study or the study population were, respectively, about 70% and 40% higher than estimates in articles reporting sufficient details. Studies reporting no details of the cut-off of the reference test had DORs that were approximately 30% smaller than studies reporting the details of the reference tests.

COMMENT

This study describes the quantitative effects of characteristics of study design on estimates of diagnostic accuracy. By collecting data from studies in published diagnostic meta-analyses, we were able to examine the effect of study characteristics on diagnostic accuracy. Our analysis shows that studies of lower methodological quality, particularly those including nonrepresentative patients or applying different reference standards, tend to overestimate the diagnostic performance of a test.

The largest effect on the estimation of diagnostic accuracy was generated by studies using cases and controls, also labeled as spectrum bias.[28] Often, mild cases that are difficult to diagnose are omitted from case-control studies, causing an overestimation of sensitivity as well as specificity. Another large effect was seen in studies that used different reference tests for the verification of positive and negative test results. The effect of this differential verification depends on the quality of the different reference tests used. Using a "gold" reference test for the positive test results and a poor reference test for the negative results can lead to an overestimation of both sensitivity and specificity of a test.[29] For example, some studies evaluating the diagnostic performance of C-reactive protein (CRP) for the diagnosis of acute appendicitis used surgery and pathology as a reference test for patients with a high CRP. Patients with a low CRP were not operated on and clinical follow-up determined whether they were classified as having acute appendicitis. As low-grade infections with low CRPs can resolve spontaneously, this verification strategy fails to identify all false-negative test results. This way the diagnostic performance of CRP will be overestimated. If the poor reference standard fails to identify true-negative test results, the use of different reference standards can lead to an underestimation of the diagnostic performance of a test.

The terms *verification bias* or *workup bias* are sometimes used when not all patients are subjected to the reference test.[29-30] We prefer the term *partial verification bias* to differentiate this situation from the situation in which different reference standards were used for verification. In theory, verifying more positive test results than negative test results will lead to an overestimation of sensitivity and an underestimation of specificity, resulting in a either an increase or a decrease of the DOR.[29] In the analysis reported here, partial verification resulted in DORs comparable with those from studies with complete verification. The absence of an association with estimation of diagnostic accuracy could be caused by the definition we used. Partial verification will only lead to bias if systematically more abnormal than normal (or more normal than abnormal) test results are subjected to the reference standard. In many studies it was not clear why some patients were not subjected to the reference standard and if this was related to the test under study. Therefore, we scored studies as partial verification when 10% of the patients were not verified, including studies in which patients were not verified due to a random error. This could weaken the possible effect of partial verification in our analysis. Some other studies also have shown no effect of partial verification on the overall diagnostic accuracy of a test. In these studies only a shift of threshold values along the receiver operating characteristic curve was observed.[31-32]

The average effect of inappropriate blinding was small. The studies included in this analysis used many different reference standards ranging from diagnostic imaging to histology. In case the reference standard is objective, no effect is to be expected. Thus, in clinical situations with a subjective reference standard, the effect of not blinding could be larger.

Given the distinction made between case-control and cohort studies, we did not observe an influence of nonconsecutive sampling of patients. Retrospective collection did not generate different results than studies with prospective data collection (when corrected for all other methodological flaws).

When looking at the criteria for the methods of reporting, we found a sufficient description of the test and a sufficient description of the population associated with an overestimation of diagnostic accuracy in case of their absence. As these criteria are not directly related to the study design, it is unclear how they lead to an overestimation of diagnostic accuracy. Somehow they seem to be predictors of methodological flaws in studies. In contrast with these findings, studies with an insufficient description of the reference standard generated less optimistic results compared with studies with an adequate description. An explanation could be that these studies possibly had a large variation in the interpretation of the reference standard. Large interobserver variation is associated with a poor diagnostic accuracy, leading to an underclassification of diseased persons.[33] However, the same argument could be used for lack of description of the test under study. Many times we had difficulties in deciding whether the reference test was described with enough detail. How much detail is needed if the reference standard is histology? The definition was largely dependent on the clinical situation under study. The scores of 20 studies changed after the consensus reading in comparison with the first reading.

Four meta-analyses have examined the quantitative effect of study characteristics on diagnostic performance. These analyses were published before 1996 and, hence, not included in our sample. All were limited to a single test.[34-37] Some of these studies found that partial verification and absence of blinding affected the estimates of diagnostic accuracy,[33-34,36] while others found no effect of these characteristics.[35] As in the first 3 meta-analyses, we found an overestimation of diagnostic accuracy in studies without appropriate blinding, as in the latter we found no effect of partial verification. One of these meta-analyses also looked at the reporting method of the test and the reference test in combination with other criteria and also found that insufficient description was associated with overestimation of the diagnostic accuracy.[37]

The decision to limit our analysis to data from recent meta-analyses could have affected our results. Extending our sample to older meta-analyses would most likely change the relative frequency of the study characteristics, but not necessarily the relative size of their effects on diagnostic accuracy.

Publication bias also has to be taken into account since only diagnostic studies published in scientific journals were included in the analysis.[38] One can speculate that studies have a higher likelihood of being published when they are either of good quality or when they show encouraging results. Such a selective publication policy could lead to an inflation of the associations we found. It is difficult to examine the effect of publication bias since there is no registration of unpublished diagnostic studies. For future research in this field and for reviewers of diagnostic tests, such a central registration of diagnostic research protocols would be useful.

When reading the results of a single diagnostic study, it is difficult to weigh the methodological flaws against the available evidence. How large is the possible overestimation and will it have clinical consequences? In a study with different reference standards, without blinding and lacking a description of the test, the DOR would on average be overestimated by 5-fold based on the results of our analysis. This is equal to reporting a sensitivity and specificity of about 84% when in fact both should be 70%. Differences will be smaller if sensitivity and specificity are higher and if only a few minor criteria are not fulfilled.

Our results stress the importance of adequate methodology and the need for complete and reliable reporting of research. Assessment of quality is only feasible in the light of complete clarity on the methodology. Authors should therefore describe explicitly their methods of patient selection, methods of disease verification, and criteria for interpretation of the test and the reference test.

This study shows that shortcomings in design, data collection, and reporting affect estimates of diagnostic accuracy. Investigators should be aware of this when designing their studies and readers should be aware of this when interpreting the results. Our results can be of help in determining the merits of the available evidence when appraising literature. Greater editorial vigilance could help make researchers aware of current methodological standards and thereby decrease the potential for bias in future diagnostic studies.

AUTHOR INFORMATION

cknowledgment: The authors gratefully acknowledge Iain Chalmers, MSc, for his review of and comments on the article.

Corresponding Author and Reprints: J.G. Lijmer, MD, Department of Clinical Epidemiology and Biostatistics, Academic Medical Center, University of Amsterdam, PO Box 22700, Amsterdam, the Netherlands 1100 DE (e-mail: j.g.lijmer@amc.uva.nl).

Author Affiliations: Department of Clinical Epidemiology and Biostatistics, Academic Medical Center, University of Amsterdam, Amsterdam, the Netherlands.

REFERENCES

1. Reid MC, Lachs MS, Feinstein AR. Use of methodological standards in diagnostic test research: getting better but still not good. *JAMA.* 1995;274:645-651. ABSTRACT

2. Jaeschke R, Guyatt G, Sackett DL. Users' guides to the medical literature, III: how to use an article about a diagnostic test, A: are the results of the study valid? *JAMA.* 1994;271:389-391. FULL TEXT | ISI | PUBMED

3. Jaeschke R, Guyatt GH, Sackett DL. Users' guides to the medical literature, III: how to use an article about a diagnostic test, B: what are the results and will they help me in caring for my patients? *JAMA.* 1994;271:703-707. FULL TEXT | ISI | PUBMED

4. Greenhalgh T. How to read a paper: papers that report diagnostic or screening tests. *BMJ.* 1997;315:540-543. **FREE** FULL TEXT

5. Mulrow CD, Linn WD, Gaul MK, Pugh JA. Assessing quality of a diagnostic test evaluation. *J Gen Intern Med.* 1989;4:288-295. ISI | PUBMED

6. Schulz KF, Chalmers I, Hayes RJ, Altman DG. **Empirical** evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA.* 1995;273:408-412. ABSTRACT

7. Feinstein AR. Diagnostic and spectral markers. *Clinical Epimediology: The Architecture of Clinical Research.* Philadelphia, Pa: WB Saunders Co; 1985:597-631.

8. van der Schouw YT, Verbeek AL, Ruijs SH. Guidelines for the assessment of new diagnostic tests. *Invest Radiol.* 1995;30:334-340. FULL TEXT | ISI | PUBMED

9. Griner PF, Mayewski RJ, Mushlin AI, Greenland P. Selection and interpretation of diagnostic tests and procedures: principles and applications. *Ann Intern Med.* 1981;94:557-592. PUBMED

10. Sackett DL, Haynes RB, Guyatt GH, Tugwell P. The selection of diagnostic tests. In: Sackett D, ed. *Clinical Epidemiology.* Boston, Mass: Little Brown & Co; 1991:47-57.

11. Kraemer HC. *Evaluating Medical Tests: Objective and Quantitative Guidelines.* Newbury Park, Calif: SAGE Publications Inc; 1992:103-113.

12. Stoffers HE, Kester AD, Kaiser V, Rinkens PE, Kitslaar PJ, Knottnerus JA. The diagnostic value of the measurement of the ankle-brachial systolic pressure index in primary health care. *J Clin Epidemiol.* 1996;49:1401-1405. FULL TEXT | ISI | PUBMED

13. Moses LE, Shapiro D, Littenberg B. Combining independent studies of a diagnostic test into a summary ROC curve: data-analytic approaches and some additional considerations. *Stat Med.* 1993;12:1293-1316. ISI | PUBMED

14. Littenberg B, Moses LE. Estimating diagnostic accuracy from multiple conflicting reports: a new meta-analytic method. *Med Decis Making.* 1993;13:313-321. **FREE** FULL TEXT

15. Irwig L, Macaskill P, Glasziou P, Fahey M. Meta-analytic methods for diagnostic test accuracy. *J Clin Epidemiol.* 1995;48:119-130. FULL TEXT | ISI | PUBMED

16. Haldane JBS. The estimation and significance of the logarithm of a ratio of frequencies. *Ann Hum Genet.* 1955;20:309-314. PUBMED

17. Scheidler J, Hricak H, Yu KK, Subak L, Segal MR. Radiological evaluation of lymph node metastases in patients with cervical cancer: a meta-analysis. *JAMA.* 1997;278:1096-1101. ABSTRACT

18. Siegman-Igra Y, Anglim AM, Shapiro DE, Adal KA, Strain BA, Farr BM. Diagnosis of vascular catheter-related bloodstream infection: a meta-analysis. *J Clin Microbiol.* 1997;35:928-936. ABSTRACT

19. Smith ER, Petersen J, Okorodudu AO, Bissell MG. Does the addition of unconjugated estriol in maternal serum screening improve the detection of trisomy 21? a meta-analysis. *Clin Lab Manage Rev.* 1996;10:176-181.

20. Becker DM, Philbrick JT, Bachhuber TL, Humphries JE. D-dimer testing and acute venous thromboembolism: a shortcut to accurate diagnosis? *Arch Intern Med.* 1996;156:939-946. ABSTRACT

21. De Vries SO, Hunink MGM, Polak JF. Summary receiver operating characteristic curves as a technique for meta-analysis of the diagnostic performance of duplex ultrasonography in peripheral arterial disease. *Acad Radiol.* 1996;3:361-369. FULL TEXT | ISI | PUBMED

22. Bonis PA, Ioannidis JP, Cappelleri JC, Kaplan MM, Lau J. Correlation of biochemical response to interferon alfa with histological improvement in hepatitis C: a meta-analysis of diagnostic test characteristics. *Hepatology.* 1997;26:1035-1044. FULL TEXT | ISI

23. Hallan S, Asberg A. The accuracy of C-reactive protein in diagnosing acute appendicitis: a meta-analysis. *Scand J Clin Lab Invest.* 1997;57:373-380. ISI | PUBMED

24. Huicho L, Campos M, Rivera J, Guerrant RL. Fecal screening tests in the approach to acute infectious diarrhea: a scientific overview. *Pediatr Infect Dis J.* 1996;15:486-494. FULL TEXT | ISI | PUBMED

25. Reed WW, Byrd GS, Gates RH Jr, Howard RS, Weaver MJ. Sputum Gram's stain in community-acquired pneumococcal pneumonia: a meta-analysis. *West J Med.* 1996;165:197-204. ISI | PUBMED

26. Mol BWJ, Dijkman AB, Wertheim P, Lijmer JG, Van der Veen F, Bossuyt PMM. The accuracy of serum chlamydial antibodies in the diagnosis of tubal pathology: a meta-analysis. *Fertil Steril.* 1997;67:1031-1037. FULL TEXT | ISI | PUBMED

27. Mirvis SE, Shanmuganathan K, Miller BH, White CS, Turney SZ. Traumatic aortic injury: diagnosis with contrast-enhanced thoracic CT—five-year experience at a major trauma center. *Radiology.* 1996;200:413-422. ABSTRACT

28. Ransohoff DF, Feinstein AR. Problems of spectrum and bias in evaluating the efficacy of diagnostic tests. *N Engl J Med.* 1978;299:926-930. ABSTRACT

29. Panzer RJ, Suchman AL, Griner PF. Workup bias in prediction research. *Med Decis Making.* 1987;7:115-119. **FREE** FULL TEXT

30. Begg CB, Greenes RA. Assessment of diagnostic tests when disease verification is subject to selection bias. *Biometrics.* 1983;39:207-215. FULL TEXT | ISI | PUBMED

31. Hunink MGM, Richardson D, Doubiley PM, Begg CB. Testing for fetal pulmonary maturity: an ROC analysis involving covariates, verification bias and combination testing. *Med Decis Making.* 1990;10:201-211. **FREE** FULL TEXT

32. Lijmer JG, Hunink MGM, van den Dungen JJ, Loonstra J, Smit AJ. ROC analysis of noninvasive tests for peripheral arterial disease. *Ultrasound Med Biol.* 1996;22:391-398. FULL TEXT | ISI | PUBMED

33. Quinn MF. Relation of observer agreement to accuracy according to a two-receiver signal detection model of diagnosis. *Med Decis Making.* 1989;9:196-206. **FREE** FULL TEXT

34. Fahey MT, Irwig L, Macaskill P. Meta-analysis of pap test accuracy. *Am J Epidemiol.* 1995;141:680-689. **FREE** FULL TEXT

35. Detrano R, Janosi A, Lyons KP, Marcondes G, Abbassi N, Froelicher VF. Factors affecting sensitivity and specificity of a diagnostic test: the exercise thallium scintigram. *Am J Med.* 1988;84:699-710. FULL TEXT | ISI | PUBMED

36. Detrano R, Gianrossi R, Froelicher V. The diagnostic accuracy of the exercise electrocardiogram: a meta-analysis of 22 years of research. *Prog Cardiovasc Dis.* 1989;32:173-206. FULL TEXT | ISI | PUBMED

37. Wells PS, Lensing AW, Davidson BL, Prins MH, Hirsh J. Accuracy of ultrasound for the diagnosis of deep venous thrombosis in asymptomatic patients after orthopedic surgery: a meta-analysis. *Ann Intern Med.* 1995;122:47-53. **FREE** FULL TEXT

38. Begg CB, Berlin JA. Publication bias and dissemination of clinical research. *J Natl Cancer Inst.* 1989;81:107-115. **FREE** FULL TEXT

ANEXO E

E.1 - Newcastle Ottawa Quality

E.2 - Newcastle - Ottawa Quality Assessment Scale

E.1 - Newcastle Ottawa Quality

E.1.1 - Case Control Studies

Note: A study can be awarded a maximum of one star for each numbered item within the Selection and Exposure categories. A maximum of two stars can be given for Comparability.

Selection

1) Is the case definition adequate?
   a) yes, with independent validation *****
   b) yes, eg record linkage or based on self reports
   c) no description

2) Representativeness of the cases
   a) consecutive or obviously representative series of cases *****
   b) potential for selection biases or not stated

3) Selection of Controls
   a) community controls *****
   b) hospital controls
   c) no description

4) Definition of Controls
   a) no history of disease (endpoint) *****
   b) no description of source

Comparability

1) Comparability of cases and controls on the basis of the design or analysis
   a) study controls for _____ (Select the most important factor.) *
   b) study controls for any additional factor (This criteria could be modified to indicate specific control for a second important factor.)

Exposure

1) Ascertainment of exposure
   a) secure record (eg surgical records) *****
   b) structured interview where blind to case/control status *****
   c) interview not blinded to case/control status
   d) written self report or medical record only
   e) no description

2) Same method of ascertainment for cases and controls
   a) yes *****
   b) no

3) Non-Response rate
   a) same rate for both groups *****
   b) non respondents described
   c) rate different and no designation

E.2 - Newcastle - Ottawa Quality Assessment Scale

E.2.2 - Cohort Studies

Note: A study can be awarded a maximum of one star for each numbered item within the Selection and Outcome categories. A maximum of two stars can be given for Comparability

Selection

1) Representativeness of the exposed cohort
   a) truly representative of the average _____ (describe) in the community *****
   b) somewhat representative of the average _____ in the community *****
   c) selected group of users eg nurses, volunteers
   d) no description of the derivation of the cohort
2) Selection of the non exposed cohort
   a) drawn from the same community as the exposed cohort *****
   b) drawn from a different source
   c) no description of the derivation of the non exposed cohort
3) Ascertainment of exposure
   a) secure record (eg surgical records) *****
   b) structured interview *****
   c) written self report
   d) no description
4) Demonstration that outcome of interest was not present at start of studya) a) yes ***** b) no

Comparability

1) Comparability of cohorts on the basis of the design or analysis
   a) study controls for _____ (select the most important factor)
   b) study controls for any additional factor ***** (This criteria could be modified to indicate specific control for a second important factor.)

Outcome

1) Assessment of outcome
   a) independent blind assessment *****
   b) record linkage *****
   c) self report
   d) no description
2) Was follow-up long enough for outcomes to occur
   a) yes (select an adequate follow up period for outcome of interest) *****
   b) no
3) Adequacy of follow up of cohorts
   a) complete follow up - all subjects accounted for *****
   b) subjects lost to follow up unlikely to introduce bias - small number lost ≥ ____ % (select an adequate %) follow up, or description provided of those lost)
   c) follow up rate < ____% (select an adequate %) and no description of those lost
   d) no statement

E.2.1 - Coding Manual for Case-Control Studies

Selection

1) Is the Case Definition Adequate?
   a) Requires some independent validation (e.g. >1 person/record/time/process to extract information, or reference to primary record source such as x-rays or medical/hospital records)
   b) Record linkage (e.g. ICD codes in database) or self-report with no reference to primary record ☆
   c) No description

2) Representativeness of the Cases
   a) All eligible cases with outcome of interest over a defined period of time, all cases in a defined catchment area, all cases in a defined hospital or clinic, group of hospitals, health maintenance organisation, or an appropriate sample of those cases (e.g. random sample)
   b) Not satisfying requirements in part (a), or not stated ☆

3) Selection of Controls
   This item assesses whether the control series used in the study is derived from the same population as the cases and essentially would have been cases had the outcome been present

   a) Community controls (i.e. same community as cases and would be cases if had outcome)
   b) Hospital controls, within same community as cases (i.e. not another city) but derived from a hospitalised population ☆
   c) No description

4) Definition of Controls
   a) If cases are first occurrence of outcome, then it must explicitly state that controls have no history of this outcome. If cases have new (not necessarily first) occurrence of outcome, then controls with previous occurrences of outcome of interest should not be excluded
   b) No mention of history of outcome ☆

Comparability

1) Comparability of Cases and Controls on the Basis of the Design or Analysis
   A maximum of 2 stars can be allotted in this category
   Either cases and controls must be matched in the design and/or confounders must be adjusted for in the analysis. Statements of no differences between groups or that differences were not statistically significant are not sufficient for establishing comparability. Note: If the odds ratio for the exposure of interest is adjusted for the confounders listed, then the groups will be considered to be comparable on each variable used in the adjustment

   There may be multiple ratings for this item for different categories of exposure (e.g. ever vs. never, current vs. previous or never)

   Age = ☆ , Other controlled factors = ☆

Exposure

1) Ascertainment of Exposure
   Allocation of stars as per rating sheet

2) Non-Response Rate
   Allocation of stars as per rating sheet

E.2.2 - Coding Manual for Cohort Studies

Selection

1) Representativeness of the Exposed Cohort
   Item is assessing the representativeness of exposed individuals in the community, not the representativeness of the sample of women from some general population. For example, subjects derived from groups likely to contain middle class, better educated, health oriented women are likely to be representative of postmenopausal estrogen users while they are not representative of all women (e.g. members of a health maintenance organisation (HMO) will be a representative sample of estrogen users. While the HMO may have an under-representation of ethnic groups, the poor, and poorly educated, these excluded groups are not the predominant users users of estrogen).
   Allocation of stars as per rating sheet

2) Selection of the Non-Exposed Cohort
   Allocation of stars as per rating sheet

3) Ascertainment of Exposure
   Allocation of stars as per rating sheet

4) Demonstration That Outcome of Interest Was Not Present at Start of Study
   In the case of mortality studies, outcome of interest is still the presence of a disease/ incident, rather than death. That is to say that a statement of no history of disease or incident earns a star.

Comparability

1) Comparability of Cohorts on the Basis of the Design or Analysis
   A maximum of 2 stars can be allotted in this category
   Either exposed and non-exposed individuals must be matched in the design and/or confounders must be adjusted for in the analysis. Statements of no differences between groups or that differences were not statistically significant are not sufficient for establishing comparability. Note: If the relative risk for the exposure of interest is adjusted for the confounders listed, then the groups will be considered to be comparable on each variable used in the adjustment.
   There may be multiple ratings for this item for different categories of exposure (e.g. ever vs. never, current vs. previous or never)
   Age = ☆ Other controlled factors = ☆

Outcome

1) Assessment of Outcome
   For some outcomes (e.g. fractured hip), reference to the medical record is sufficient to satisfy the requirement for confirmation of the fracture. This would not be adequate for vertebral fracture outcomes where reference to x-rays would be required.
   a) Independent or blind assessment stated in the paper, or confirmation of the outcome by reference to secure records (x-rays, medical records, etc.)
   b) Record linkage (e.g. identified through ICD codes on database records) ☆
   c) Self-report (i.e. no reference to original medical records or x-rays to confirm the outcome) ☆
   d) No description.

2) Was Follow-Up Long Enough for Outcomes to Occur
   An acceptable length of time should be decided before quality assessment begins (e.g. 5 yrs. for exposure to breast implants)

3) Adequacy of Follow Up of Cohorts
   This item assesses the follow-up of the exposed and non-exposed cohorts to ensure that losses are not related to either the exposure or the outcome.
   Allocation of stars as per rating sheet

ANEXO F

The Strengthening the Reporting of Observational Studies (Strobe)

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No	Recommendation
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found
Introduction		
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported
Objectives	3	State specific objectives, including any prespecified hypotheses
Methods		
Study design	4	Present key elements of study design early in the paper
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group
Bias	9	Describe any efforts to address potential sources of bias
Study size	10	Explain how the study size was arrived at
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why

Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding
		(*b*) Describe any methods used to examine subgroups and interactions
		(*c*) Explain how missing data were addressed
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy
		(*e*) Describe any sensitivity analyses
Results		
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed
		(b) Give reasons for non-participation at each stage
		(c) Consider use of a flow diagram
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders
		(b) Indicate number of participants with missing data for each variable of interest
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time
		Case-control study—Report numbers in each exposure category, or summary measures of exposure
		Cross-sectional study—Report numbers of outcome events or summary measures
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included
		(*b*) Report category boundaries when continuous variables were categorized
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period

Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses
Discussion		
Key results	18	Summarise key results with reference to study objectives
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence
Generalisability	21	Discuss the generalisability (external validity) of the study results
Other information		
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

ANEXO G

Meta-analysis of Observational Studies in Epidemiology MOOSE

Table. A Proposed Reporting Checklist for Authors, Editors, and Reviewers of Meta-analyses of Observational Studies

Reporting of background should include
Problem definition
Hypothesis statement
Description of study outcome(s)
Type of exposure or intervention used
Type of study designs used
Study population
Reporting of search strategy should include
Qualifications of searchers (eg, librarians and investigators)
Search strategy, including time period included in the synthesis and keywords
Effort to include all available studies, including contact with authors
Databases and registries searched
Search software used, name and version, including special features used (eg, explosion)
Use of hand searching (eg, reference lists of obtained articles)
List of citations located and those excluded, including justification
Method of addressing articles published in languages other than English
Method of handling abstracts and unpublished studies
Description of any contact with authors
Reporting of methods should include
Description of relevance or appropriateness of studies assembled for assessing the hypothesis to be tested
Rationale for the selection and coding of data (eg, sound clinical principles or convenience)
Documentation of how data were classified and coded (eg, multiple raters, blinding, and interrater reliability)
Assessment of confounding (eg, comparability of cases and controls in studies where appropriate)
Assessment of study quality, including blinding of quality assessors; stratification or regression on possible predictors of study results
Assessment of heterogeneity
Description of statistical methods (eg, complete description of fixed or random effects models, justification of whether the chosen models account for predictors of study results, dose-response models, or cumulative meta-analysis) in sufficient detail to be replicated
Provision of appropriate tables and graphics
Reporting of results should include
Graphic summarizing individual study estimates and overall estimate
Table giving descriptive information for each study included
Results of sensitivity testing (eg, subgroup analysis)
Indication of statistical uncertainty of findings
Reporting of discussion should include
Quantitative assessment of bias (eg, publication bias)
Justification for exclusion (eg, exclusion of non–English-language citations)
Assessment of quality of included studies
Reporting of conclusions should include
Consideration of alternative explanations for observed results
Generalization of the conclusions (ie, appropriate for the data presented and within the domain of the literature review)
Guidelines for future research
Disclosure of funding source

ANEXO H

Estadiamento do Câncer de Ovário

Estadiamento do Câncer de Ovário

Estádio I	I - Tumor limitado aos ovários Ia – tumor limitado a um ovário, sem ascite de células tumorais, sem tumoração na superfície externa e com cápsula intacta Ib – tumor em ambos os ovários, sem ascite de células tumorais, sem tumoração na superfície externa e com cápsula(s) intacta Ic – tumor nos estádios Ia ou Ib, mas com tumor na superfície de um ou de ambos os ovários, ou com ruptura da cápsula (s), ou com ascite presente contento células malignas, ou com lavado peritoneal positivo para células malignas
Estádio II	II - Tumor envolvendo um ou dois ovários com extensão a órgãos pélvicos IIa – extensão ou metástases para útero e/ou tubas IIb – extensão para outros órgãos pélvicos IIc – tumor estádio IIa ou IIb com ascite presente ou citologia peritoneal positiva e /ou com ruptura da cápsula (s)
Estádio III	III - Tumor envolvendo um ou ambos os ovários com metástases peritoneal fora da pelve e/ou linfonodos retroperitoneais positivos ou inguinais; tumor limitado à verdadeira pelve com extensão histológica para intestino delgado e/ou omento. Metástase na superfície do fígado IIIa – tumor grosseiramente limitado à pequena pelve, com linfonodos negativos, mas com implantes microscópicos confirmados por histologia na superfície abdominal peritoneal ou em intestino ou mesentério IIIb – tumor em um ou ambos os ovários, com implantes peritoneais com menos de 2 cm em superfície peritoneal com linfonodos negativos IIIc – implantes peritoneais com mais de 2 cm de diâmetro e /ou linfonodos retroperitoneais ou inguinais
Estádio IV	IV - Tumor envolvendo um ou ambos os ovários com metástases à distância, se houver derrame pleural e/ou citologia positiva considera-se estádio IV. Metástases parenquimatosa hepática

Fonte: Benedet *et al.,* 2000

ANEXO I

Resumo da Classificação Histológica dos Tumores Ovarianos Conforme as Subseções do Texto

Resumo da Classificação Histológica dos Tumores Ovarianos Conforme as Subseções do Texto

Classificação histológica	Tipos de tumores	Características
Tumores epiteliais	– seroso – mucinoso – endometrióide – de células claras – de células transicionais – de células escamosas – epitelial misto – carcinoma indiferenciado	• correspondem a 65-70% dos tumores ovarianos • 50-60% são benignos • 25% são malignos • 10-20% possuem baixo potencial de malignidade • podem ocorrer em mulheres entre 20 e 60 anos, com média idade 55 anos • na grande maioria são císticos
Tumores do estroma do cordão sexual	– de células da granulosa e da teca do ovário – da célula de Leydig-Sertoli	• correspondem a 5-10% dos tumores ovarianos • 5% ocorrem na idade pré-puberal • 55% são diagnosticados na menacma e 40% na pós-menopausa • são predominantemente sólidos • os da granulosa são potencialmente malignos com estimativa de recorrência variando entre 5% e 25% • medem cerca de 12 cm de diâmetro • associados a hiperplasia endometrial em 10% casos • dos tumores da teça o fibrotecoma é mais comum ocorrendo em 4% dos casos e são benignos
Tumores de células germinativas	– teratoma – disgerminoma	• correspondem a 15-20% dos tumores de ovário • mais comum é o teratoma • 10% dos teratomas são bilaterais • 96,7% dos teratomas são benignos • o disgerminoma corresponde a 2% da tumorações malignas de ovário • o disgerminoma é freqüente em crianças, adolescentes e mulheres jovens
Lesões semelhantes a tumores	– cistos funcionais – endometriose	• são as tumorações mais comuns • ocorrem no menacma • o cisto funcional em 80% casos regride com uso de anticoncepcional • endometriose com pico máximo entre 20 e 29 anos e 40 e 44 anos

Fonte: Scully (1999)

ANEXO J

Characteristics of Excluded Studies do Artigo 1

Characteristics of Excluded Studies

Study	Reason for exclusion
Amara 1996	Series of cases.
Berman 2003	Narrative review.
Bristow 2000	Narrative review.
Canis 1994	Narrative review.
Canis 1997	Not give the stage of ovarian cancer (Ia, Ib or Ic). They wrote about 10 cases of low malignant potencial tumour and 15 cases of cancer, but without stage.
Canis 2000	Not give the stage of ovarian cancer (Ia, Ib or Ic). Only related about 28 cases of cancer and bordeline tumor.
Chapron 1998	Narrative review.
Childers 1995	Casol control trial, but with second look laparoscopy for evaluate both intraperitoneal cavity and retroperitoneal lymph nodes.
Childers 1996	Not give the stage of ovarian cancer (Ia, Ib or Ic). Only related about 19 cases of cancer.
Darai1998	Retrospective trial.
Dottino 1999	Another ovarian disease, wrote about ovarian cancer, they gave stage IIC, IIIa abd IV for ovarian cancer.
Fauvet 2005	Retrospective study.
Goff 2006	Narrative review.
Kadar 1995	Not randomised.Other kind of cancer (endometrial, cervical, ovarian).
Klindermann 1995	A questionary was mailed to 237 German Departament Gyn/obs. A response rate 46% (127 hospital) concerning the Endoscopicall tecnique used for cancer operation
Leblanc 2004	Coorte with other types of cancer (fallopium tube carcinoma) , and in patientes that inadequadeteky stage at the time of initial surgery for inavsive ovarian carcinoma.
Leblanc 2006	Narrative review.
Lécuru 2004	Retrospetive and mulcentric study.
Maiman 1991	Members and candidate members of the Society of Gynecologic Oncologists responded to a survey concerning the "laparoscopy managmente of ovarian neoplasm subsequently found be malignant".
Malik 1998	Retrospective study.
Maneo 2004	Criteria fo exclusion: 62 patients had ferlity -sparinf after surgery.
Manolitsas 2001	Narrative review.
Mehra 2004	To describe experinece of laparoscopy extraperitoneal paraaortic in 32 patients with cervical, ovarian and endometrial carcinomas.
Nezhat 1992	Series cases.
Parker 1990	Only benign ovarian cysts.
Pomel 1995	Patients with I ovarian carcinoma undefwente a laparoscopic procedure to complete their staging.
Poncheville 2001	Retrospective study.
Querleu 2003	Retrospective study.
Querleu 2006	Many types of tumors (cervical, vaginal, endometrial, and ovarian carcinoma.
Querleu 2006	Narrartive review.
Reich 1990	Series of cases.
Romagnolo 2006	Pactientes affected by ovarian masss suspected of bordeline ovarian tumor are opered on primary laparoscopic, but 46,9% have notive for exclusion of SR, because a fertility-

	sparing surgical treatment was chosen.
Rouzier 2005	Narrative review.
Spirtos 2005	Other kinds of gynecological cancer.
Tozzi 2005	Narrative review.
Tropé 2006	Narrative review.
Vaisbuch 2005	Narrative review.
Vergote 2003	Narrative review.
Vinatier 1996	Narrative review.
Volz 1996	Narrative Review.
Wenzl 1996	A questionary was sent to all 97 Departaments of Gynecology in Austria was to determine the frequency of discovering a malignant ovarian mass when laparoscopy is uses to manage an adnexal mass.

ANEXO L

Artigo de HUA *et al*. em Chinês

·腔镜技术临床应用·

腹腔镜手术治疗早期低危型卵巢恶性肿瘤的价值

华克勤 金福明 徐焕 朱芝玲 林金芳 丰有吉

【摘要】 **目的** 探讨腹腔镜手术治疗早期低危型卵巢恶性肿瘤的价值。**方法** 对 10 例卵巢恶性肿瘤患者，在腹腔镜下行全子宫切除＋双附件切除＋卵巢动静脉高位结扎＋盆腔淋巴结清扫＋大网膜切除＋阑尾切除，并将同期 11 例施行经腹同类手术的卵巢恶性肿瘤患者作为对照组（开腹组）。就两组手术的手术时间、术中出血量、术后恢复情况和切除的盆腔淋巴结的数目进行比较。**结果** 手术时间腹腔镜组为 298 min ± 60 min，开腹组为 182 min ± 43 min，两者比较 $P < 0.05$；术中出血量腹腔镜组为 280 ml ± 156 ml，开腹组为 346 ml ± 170 ml，两者比较 $P < 0.05$；切除淋巴数目腹腔镜组为 25 个 ± 5 个，开腹组为 27 个 ± 6 个，两者比较 $P > 0.05$；术后病率腹腔镜组为 20.0%，开腹组为 72.7%，两者比较 $P < 0.01$；48 h 下床活动腹腔镜组为 7 例，开腹组为 1 例，两者比较 $P < 0.05$；腹腔镜下手术发生 1 例右侧闭孔神经损伤，在腹腔镜下缝合成功。**结论** 腹腔镜下可施行早期低危型卵巢恶性肿瘤手术包括全子宫＋双附件切除＋卵巢动静脉高位结扎＋盆腔淋巴结清扫＋大网膜切除＋阑尾切除的全部过程，具有术中出血少、术后病率低、术后恢复快的优点。

【关键词】 腹腔镜检查； 卵巢肿瘤； 低危型

Evaluation of laparoscopic surgery in the early stage malignant tumor of ovary with lower risk

HUA Ke-qin, JIN Fu-ming, XU Huan, ZHU Zhi-ling, LIN Jin-fang, FENG You-ji. Department of Gynecology, Fudan University, Gynecology and Obstetric hospital, Shanghai 200011, China

【**Abstract**】 **Objective** To evaluate the laparoscopic operation for early ovarian malignant tumor with low risk. **Methods** Ten patients with ovarian malignant tumor who underwent laparoscopic total hysterectomy, pelvic lymph nodes dissection, bilateral adnexectomy, ovarian aortic and vein high ligation, omentectomy, and additional appendectomy. Eleven patients with the same diagnosis who underwent operation by laparotomy were served as control group. The operation time, intraoperative blood loss, number of pelvic lymph nodes excised, and postoperative recovery were analyzed retrospectively. **Results** Frozen section method during operation proved the diagnosis of ovarian malignant tumor and cytological examination proved a negative result of the peritoneal irrigation liquid. The operation time was 298 min ± 60 min for the laparoscopy group and 182 min ± 43 min for the laparotomy group ($P < 0.05$). The intraoperative blood loss was 280 ml ± 156 ml for the laparoscopy group and 346 ml ± 170 ml for the laparotomy group ($P < 0.05$). The number of pelvic lymph node resected was 25 ± 5 and 27 ± 7 for the laparoscopy group and laparotomy group respectively ($P > 0.05$). The postoperative illness rate was 20.0% and 72.7% for the laparoscopy group and laparotomy group respectively ($P < 0.01$). Seven patients and 1 case in the laparoscopy group and laparotomy group left their beds 48 hours after operation ($P < 0.05$). The right obtuator nerve was injured and was sutured on 1 patient in the laparoscopy group. **Conclusion** The whole procedure of total hysterectomy, bilateral adnexectomy, pelvic lymph node dissection, ovarian aortic and vein high ligation, omentectomy, and additional appendectomy may be performed under laparoscope in the treatment of early stage ovarian malignant tumor with lower risk. The laparoscopic operation has the advantage of less intraoperative bleeding, less morbidity and rapid recovery.

【**Key words**】 Laparoscopy; Ovarian neoplasms; Lower risk

应用腹腔镜手术治疗良性卵巢肿瘤已逐渐成为手术治疗的首选方法。目前，应用腹腔镜治疗良性肿瘤时一旦术中冰冻切片证实为卵巢恶性肿瘤，仍要中转剖腹完成手术的全过程。20 世纪 90 年代起，一些国外学者试图通过腹腔镜来完成上述手术操作，但国内尚无此方面报道。2002 年起我院运用腹腔镜手术技术对 10 例早期低危型卵巢恶性肿瘤患者施行全子宫切除＋双附件切除＋卵巢动静脉高位结扎＋盆腔淋巴结清扫＋大网膜切除＋阑尾切除的全过程，取得了令人满意的临床效果。

作者单位：200011 上海，复旦大学附属妇产科医院妇科

对象与方法

一、对象

腹腔镜组为 2002 年 9 月至 2004 年 5 月，我院经妇科检查、彩色超声检查、肿瘤标志物检查诊断为卵巢肿瘤的 10 例患者，其中拟诊恶性卵巢肿瘤 6 例，拟诊卵巢囊肿 4 例，10 例术中冰冻切片检查均诊断为恶性卵巢肿瘤。对照组（开腹组）为 11 例行开腹手术治疗的卵巢恶性肿瘤患者。两组患者临床及病理情况差异均无统计学意义（表 1）。

二、方法

腹腔镜组患者在腹腔镜下行全子宫切除 + 双附件切除 + 卵巢动静脉高位结扎 + 盆腔淋巴结清扫 + 大网膜切除 + 阑尾切除，剖腹组患者施行经腹与腹腔镜同类手术。

1. 手术器械：腹腔镜常规手术器械采用美国 storz 公司产品，超声发生器、超声刀采用美国强生公司产品，高频电刀采用 Commed 公司产品。

2. 手术方法：（1）术前准备：采用气管插管全身麻醉，取膀胱截石位，头低臀高，常规消毒后铺巾，插导尿管留置集尿袋。用 10 mm 套管针自脐孔中央穿刺置入腹腔镜，左、右下腹部、耻骨联合上方3 cm、脐孔旁左 4 cm 处分别置入第 2、3、4、5 套（5、5、10、5 mm）管针。常规环视盆腹腔，检查子宫、双附件形态、大小、活动度，检查肝、胆、膈肌、胃、肠管、大网膜、直肠陷窝有无病灶，并抽取腹腔液或用生理盐水 100 ml 冲洗盆腹腔后留取盆腹腔冲洗液，离心沉淀后找癌细胞。（2）手术步骤：腹腔镜下行全子宫切除 + 双附件切除 + 卵巢动静脉高位结扎 + 盆腔淋巴结清扫 + 大网膜切除 + 阑尾切除：未切除子宫前，卵巢肿瘤标本先置于 7.5 寸消毒乳胶手套内或经阴道行囊肿穿刺缩减瘤体后与切除的子宫一起取出。术中冰冻切片诊断为卵巢恶性肿瘤，腹腔冲洗液细胞学检查均为阴性。① 盆腔淋巴结切除术：打开侧腹膜，超声刀剪开髂外动脉血管鞘，撕拉法切除髂外淋巴结，切除髂外静脉顶端腹股沟深淋巴结，在髂外静脉内侧暴露闭孔神经，清除其上方的闭孔淋巴结，再沿髂内动脉切除周围淋巴结，切除的淋巴结置入 7.5 寸消毒乳胶手套自制的取瘤袋内取出。② 卵巢动静脉和子宫动脉处理：仔细辨认输尿管，在其跨过髂血管处用双极电刀高位凝切卵巢动静脉，游离髂内动脉向内下暴露膀胱上动脉，在其内侧分离子宫动脉，用双极电刀凝切断。③子宫、阴道残端处理：超声刀凝切圆韧带，并打开膀胱反折腹膜，经阴道切开阴道穹隆，分离子宫颈与膀胱及直肠间隙，剪开后腹膜，分次钳夹，切断骶、主韧带后切除子宫。在阴道内放置充气消毒手套，以保证人工气腹的维持。之后自阴道取出切除的大网膜和阑尾，用可吸收线关闭阴道顶。④大网膜及阑尾切除：将腹腔镜置于耻骨联合上方套管中，沿横结肠下缘用超声刀凝切大网膜。双极电刀凝切阑尾周围血管，不吸收套扎线分 3 次套扎阑尾根部切除阑尾，以双极电凝凝固残端。开腹手术组全子宫切除 + 双附件切除 + 卵巢动静脉高位结扎 + 盆腔淋巴结清扫 + 大网膜切除 + 阑尾切除：手术步骤同常规手术[1]。

三、统计学分析

数据输入 SPSS 9.0 软件，结果以 $\bar{x} \pm s$ 表示，两数间均数比较采用 t 检验，计量资料采用 χ^2 检验。

结　果

1. 术中情况比较：腹腔镜组手术时间长于开腹组（$P < 0.05$），但术中出血量少于开腹组（$P < 0.05$），两组手术中切除淋巴结的数目差异无统计学意义（$P > 0.05$，表 2）。

表 2　两组卵巢肿瘤患者术中情况比较（$\bar{x} \pm s$）

组别	例数	手术时间（min）	术中出血量（ml）	切除淋巴结（个数）	损伤（例数）
腹腔镜组	10	298 ± 60	280 ± 156	25 ± 5	1
开腹组	11	182 ± 43	346 ± 170	27 ± 6	0
P 值		< 0.05	< 0.05	> 0.05	

2. 术后情况比较：腹腔镜组术后病率低于开腹组（$P < 0.01$），术后排气时间早于开腹组（$P < 0.01$）。开腹组术后并发症包括尿潴留、淋巴囊肿、伤口感染各 1 例。腹腔镜组术后用镇痛药物例数少

表 1　两组卵巢肿瘤患者的临床及病理情况

组别	例数	年龄（岁，$\bar{x} \pm s$）	卵巢直径（cm，$\bar{x} \pm s$）	Ca125（IU/L）	临床分期（例数）		情况			
							病理分型（例数）		肿瘤分化程度（例数）	
					Ⅰ期	Ⅱ～Ⅳ期	上皮性	性索间质	高	低
腹腔镜组	10	40 ± 8	6.6 ± 2.8	49.8	10	0	9	1	10	0
开腹组	11	42 ± 6	6.8 ± 5.2	52.6	11	0	9	2	11	0

于剖腹组($P<0.01$);术后48 h下床活动例数多于剖腹组($P<0.05$,表3)。

表3 两组卵巢肿瘤患者术后情况比较

组别	例数	术后病率(%)	术后排气(h,$\bar{x}\pm s$)	术后并发症			术后镇痛(例)	48 h下床活动(例)
				尿潴留	淋巴囊肿(例)	伤口感染		
腹腔镜组	10	20.0	13±10	0	0	0	1	7
剖腹组	11	72.7	42±17	1	1	1	10	1
P值		<0.01	<0.01				<0.01	<0.05

讨 论

一、腹腔镜手术治疗卵巢恶性肿瘤的选择

腹腔镜手术已广泛运用于妇科良性疾病的手术治疗,20世纪90年代以来,腹腔镜手术已开始尝试于妇科恶性肿瘤。研究者对早期子宫内膜癌和宫颈癌应用腹腔镜手术治疗争议较少[2],而对卵巢恶性肿瘤的腹腔镜手术治疗多持否定态度,认为CO_2气腹可能增加腹壁切口肿瘤种植或转移,故腹腔镜手术治疗卵巢恶性肿瘤一直被认为是不可行的。我们对术前妇科检查、肿瘤标志物、彩色超声及标本探查可疑为卵巢恶性肿瘤者,术中进行快速冰冻检查,同时腹腔液或腹腔冲洗液细胞检查阴性的低危型早期恶性卵巢肿瘤患者在腹腔镜下完成手术的全过程,本组10例手术均成功。我们认为腹腔镜下行全子宫切除+双附件切除+盆腔淋巴结清扫+大网膜切除+阑尾切除仅限于低危型早期恶性卵巢肿瘤患者。

二、腹腔镜手术治疗卵巢恶性肿瘤的优势和难点

1. 手术在放大的视野中进行,尤其在清扫盆腔较深部位闭孔淋巴结时使视觉明显变浅,手术视野更清晰。

2. 手术在相对闭合的腔内进行,保持机体内环境稳定,对腔内其他脏器干扰小,术后肠功能恢复快,术后病率显著降低。

3. 手术在正压环境中进行,出血量更少,对患者损伤更小,术后疼痛减轻,恢复快。

4. 盆腔淋巴清扫术数目并未减少,在有开腹的临床经验的基础上,可达到与开腹手术同样的效果[3]。

5. 可以更全面、细致、直观地检查盆、腹腔脏器,特别是横膈有无肿瘤转移,可避免开腹手术大切口造成盆腹腔粘连及术后患者疼痛。

三、腹腔镜手术治疗卵巢恶性肿瘤成功的关键

1. 操作技巧掌握:只要术者具备经腹手术的技术及娴熟腹腔镜的操作技巧,并熟悉盆腔解剖,按照开腹手术基本程序,结合腹腔镜技术操作技巧进行,完全能达到开腹手术的范围要求,并具有创伤少,恢复快的优点。同时应有配合默契的助手,否则非但不能发挥腹腔镜手术微创的优势,反而增加手术风险,也达不到恶性肿瘤的治疗效果。

2. 手术并发症的防治:腹腔镜手术治疗卵巢恶性肿瘤,手术难度大,技术要求高。盆腔淋巴清扫手术过程中,髂血管的静脉壁薄而脆,极易损伤,因此操作动作必须轻柔,血管小分支宜选用双极电凝或超声刀处理,一旦发生大静脉损伤,应沉着冷静,迅速判断损伤部位。由于患者取膀胱截石位,下肢静脉受阻,若为髂外静脉微小损伤,出血并不多,只要具备熟练的镜下缝合技术,应立即选用不吸收尼龙缝线连续缝合,一般可成功修补。若为髂内或髂总静脉损伤,往往出血汹涌,应立即用无损伤钳阻断血管,随即中转剖腹止血。若勉强镜下止血,可能因失血过多,导致无法挽回的后果。淋巴管需用电凝或超声刀阻断,以防术后淋巴囊肿形成。腹腔镜下盆腔视觉变浅,闭孔神经位置较浅,若未将闭孔神经与其上方的淋巴组织充分游离,即清扫神经上方的闭孔淋巴结,可能损伤闭孔神经。若术中及时发现可在镜下用3-0可吸收线先逢合1~2针,作为牵引,再用6-0不可吸收尼龙线间断逢合3针,以维持一定的神经张力。本研究腹腔镜组中发生1例右侧闭孔神经损伤,腹腔镜术中及时修补,术后随访患者无下肢运动及功能障碍。因此在清扫闭孔淋巴时必须完全暴露其下方的闭孔神经,同时避免超声刀操作孔在置镜孔的正下方所造成的视盲区,一般可选择位于左侧脐旁6~8 cm处放置超声刀。

3. 先进手术器械的合理应用:超声刀震荡所产生的热量远小于高频电刀,周围组织产生的热损伤小,可安全地在重要脏器或髂血管旁进行切割、分离[4]。超声刀具有一次完成分离、凝固、切割的功能,减少器械更换,使手术更便捷。超声刀在操作过程中产生烟雾小、组织焦痂少,手术视野清晰,安全系数更高[5,6]。此外由于超声刀叶振动产生低压带引起的局部低压使细胞内的水分在37℃状态下汽化,在含蛋白质低的脂肪组织中起主要作用[7],因此本研究用超声刀施行盆腔淋巴清扫及大网膜切除,既避免了在清扫盆腔淋巴结及大网膜过程中可能造成对后腹膜大血管及横结肠的损伤,同时又可达到对脂肪淋巴组织中淋巴管及大网膜血管的超声

凝固和切割作用，但在处理 3 mm 以上血管时，宜先用双极电凝处理后再断离，以确保手术的安全[8]。

通过本组资料分析，腹腔镜手术治疗早期卵巢恶性肿瘤具有潜在的临床价值。但由于腹腔镜手术治疗妇科恶性肿瘤时间短，大样本资料尚欠缺，需要进行一系列前瞻性对照研究，如生存率、复发率等疗效随访指标的分析，才能取得更客观的临床效果。

参 考 文 献

1 Cao ZY, ed. Gynecologic oncology. Beijing: Beijing Publishing House, 1998. 1005-1011.
曹泽毅，主编. 妇科肿瘤学. 北京：北京出版社，1998. 1005-1011.

2 Li GY, Wang G, Chen LS. A comparision study between laparoscopic and open operation for the treatment of early-stage cervix cancer. Clinical analysis of 37 cases. Chin J Min Inv Surg, 2003, 3: 315-316.
李光仪，王刚，陈霖诗. 腹腔镜与剖腹手术治疗早期子宫颈癌 37 例临床分析. 中国微创外科手术杂志，2003，3：315-316.

3 Spirtos NM, Eisenkop SM, Schlaerth JB, et al. Laparoscopic radical hysterectomy (type Ⅲ) with aortic and pelvic lymphadenectomy in patients with stage I cervical cancer: surgical morbidity and intermediate follow up. Am J Obatet Gynecol, 2002, 187: 340-348.

4 Sietses C, Eijsbouts QAJ, von Blomberg BM, et al. Ultrasonic energy vs monopolar electrosurgery in laparoscopic cholecystectomy: influence on the postoperative systemic immune response. Surg Endosc, 2000, 83: 915-920.

5 Stringer NH, Strassner HT, Lawson L, et al. Laparoscopic ultrasonic energy and laparoscopic suturing of the endometrial cavity. J Am Assoc Gynecol Laparosc, 2001, 8: 129-136.

6 Liang ZQ, Xu HC, Chang Q, et al. Role of laparoscopic ultrasonic scalpel in lymphadenectomy of gynecologic malignancies. Acta Academiae Medicinae Militaris Tertiae, 2001, 23: 1466-1468.
梁志清，徐惠成，常青，等. 腹腔镜超声刀在妇科恶性肿瘤淋巴结切除术中的应用. 第三军医大学学报，2001，23：1466-1468.

7 Lin JF, Feng ZC, Ding AH, eds. Practical gynecologic endoscopy. Shanghai: Fudan University Pubishing House, 2001. 55-56.
林金芳，冯璋冲，丁爱华，主编. 实用妇科内镜学. 上海：复旦大学出版社，2001. 55-56.

8 Gyr T, Ghezzi F, Arslanagic S, et al. Minimal invasive laparoscopic hysterectomy with ultrasonic scalpel. Am J Surg, 2001, 181: 516-519.

（收稿日期：2004-11-03）
（本文编辑：朱瑶）

·读者·作者·编者·

本刊“循证医学”栏目征稿

循证医学是临床研究证据与临床经验、临床决策以及患者价值观的结合，就是在临床医疗实践中，对患者的诊治决策建立在最新的科学依据的基础上。本刊于 2000 年曾连续刊登了有关循证医学的讲座，目的是进一步在疾病防治中实践循证医学，通过循证医学的方法，得到有关诊断、诊疗、预防的证据，来解决我们每日碰到的临床问题。

本刊将组织稿源开辟“循证医学”栏目，为您提供最佳临床诊疗方案，实践循证医学的园地。同时也邀请您把自己临床工作中成功运用循证医学的方法交给我刊，为广大临床医师借鉴。让我们共同促进 21 世纪的医学从经验医学向循证医学转化。

具体形式为临床循证，临床证据，循证病例报告（如：手术的方法？手术时机？用药是否有效？能多大程度地预防并发症？药物的副作用有多大？）等等。

书写要求按实践循证医学的 5 个步骤书写，第一步，提出问题；第二步，查询证据；第三步，评价证据（根据文献）；第四步，应用证据；第五步，后效评价。

如要索取具体写作方式可电话联系：010-65273362 中华医学杂志编辑部 李群。

本刊“临床医学影像”栏目征稿

医学影像学检查是临床常用的诊断手段。影像学改变是病理改变的反映，但不同的病理改变往往有相似的影像学表现，这给诊断带来很大困难。为了促进临床影像诊断经验的交流和诊断、鉴别诊断水平的提高，中华医学杂志自 2001 年第 1 期开辟“临床医学影像”栏目，为特殊的、少见的、但具有临床启发意义的影像学表现提供一个展示园地，使局部的、个人的经验尽快地为广大临床医师借鉴，为临床医学影像诊断积累宝贵的第一手资料。本栏目是一个以图片展示为主的栏目，要求提供高质量的影像图片，图片必须清晰、对比度好、病变特征显示明确。每篇文章可提供 2～4 幅不同影像技术的图片，如 X 线、CT、磁共振成像、超声、核素显像或病理图片等。文字部分则宜简练，描述患者的简要病史，主要影像学表现，经病理或临床科学手段确定的最后诊断结果，不进行讨论，不引用参考文献，字数在 400 字以内。欢迎踊跃投稿。

Avaliação de Cirurgia Laparoscópica para Tumor Ovariano Maligno de Baixo Risco e em Estágio Inicial

SUJEITOS E MÉTODO DA PESQUISA

1. Sujeitos da pesquisa

O Grupo de Laparoscopia foi composto de 10 pacientes (sujeitos) diagnosticados com casos de tumor ovariano por meio de exames ginecológicos, ultra-sonografia Doppler colorida e os biomarcadores tumorais realizados por este hospital durante o período de setembro de 2002 a maio de 2004. Dos dez casos, seis foram diagnosticados como casos de tumor ovariano maligno, e quatro como casos de cisto ovariano. Não obstante, todos esses dez casos foram diagnosticados como casos de tumor ovariano maligno por meio de criosecção (também denominado "procedimento de secção congelada"). O grupo controle (isto é, o Grupo de Laparotomia) consistiu de onze casos de tumor ovariano maligno com laparotomia a ser realizada. Nenhuma diferença significativa do ponto de vista estatístico foi observada nas condições patológicas e clínicas pré-cirúrgicas entre esses dois grupos (Ver Tabela 1).

2. Método de pesquisa

Os pacientes do Grupo de Laparoscopia foram submetidos a histerectomia laparoscópica total, salpingooforectomia bilateral, alta ligação da artéria e da veia ovarianas, linfadenectomia pélvica (ressecção dos linfonodos pélvicos), omentectomia e apendicectomia, enquanto os pacientes do Grupo de Laparotomia foram submetidos à cirurgia transabdominal e ao mesmo tipo de cirurgia laparoscópica.

(1) Instrumentos cirúrgicos:

Instrumentos cirúrgicos fabricados pela empresa Storz Instruments dos EUA foram empregados como instrumentos laparoscópicos convencionais para este estudo, juntamente com o uso de geradores ultra-sônicos e escalpelos ultra-sônicos, fabricados pela empresa Johnson & Johnson e eletrocautério de alta freqüência de Commed.

(2) Métodos cirúrgicos:

(i) Preparações pré-cirúrgicas: os pacientes foram submetidos a anestesia sistêmica através de intubação endotraqueal e, a seguir, foram colocados em uma posição de litotomia abaixo da sua cabeça e acima dos seus quadris. Toalhas foram estendidas após esterilização de rotina. Após a inserção de um cateter urinário, uma bolsa coletora de urina foi acoplada. Um laparoscópio foi colocado puncionado-se um primeiro trocarte com 10 mm através do meio do umbigo,

seguido da colocação de um segundo, terceiro, quarto e quinto trocarte (de 5, 5, 10, e 5 mm respectivamente, no quadrante inferior esquerdo do abdômen, no quadrante inferior direito do abdômen, 3 cm acima da sínfise púbica, e 4 cm da esquerda ao lado do umbigo. Através de visualizações panorâmicas de rotina da cavidade peritoneal pélvica, a morfologia, tamanho e nível de atividade do útero e o anexo bilateral foram examinados, juntamente com exames do fígado, da vesícula biliar, diafragma, estômago, duto intestinal, omento maior e a bolsa retal para verificar a existência de qualquer foco de infecção. O fluido peritoneal foi coletado ou a solução de pós-enxágüe foi retida após o enxágüe da cavidade peritoneal pélvica com 100 ml de solução salina normal. Por último, a existência de células cancerígenas foi examinada usando-se precipitação centrífuga.

(ii) Procedimentos cirúrgicos: histerectomia laparoscópica total, salpingooforectomia bilateral, alta ligação da artéria e veia ovarianas, linfadenectomia pélvica (ressecção dos linfonodos pélvicos), omentectomia e apendicectomia foram realizados. Um espécime de tumor ovariano foi colocado dentro de uma luva de látex esterilizada de 19 cm antes da ressecção do útero, ou o útero foi resseccionado e a seguir removido juntamente com o tumor ovariano contraído obtido através da punção vaginal dos cistos. Os espécimes de ressecção congelados obtidos durante a operação foram diagnosticados como tumores ovarianos malignos, enquanto que a solução peritoneal pós-enxágüe apresentou resultado negativo no teste feito por meio de exame citológico. Os procedimentos cirúrgicos consistiram dos seguintes procedimentos:

i. Linfadenectomia pélvica: após a abertura do peritônio lateral, e a abertura por meio de corte da bainha do vaso da artéria ilíaca externa usando um escalpelo ultra-sônico, os linfonodos ilíacos externos foram divididos e resseccionados, seguido da ressecção dos linfonodos inguinais profundos localizados no vértice da veia ilíaca externa. Após a exposição do nervo obturador localizado no interior da veia ilíaca externa, os linfonodos obturadores localizados acima do nervo obturador foram removidos, e os linfonodos ao longo da periferia da artéria ilíaca interna foram então resseccionados. Os linfonodos assim resseccionados foram colocados em uma bolsa coletora de tumor feita de uma luva de látex de 19 cm esterilizada.

ii. Procedimentos cirúrgicos para a artéria ovariana, a veia ovariana e a artéria uterina: após a cuidadosa identificação do ureter, a coagulação laparoscópica e a ressecção da artéria e da veia ovariana localizadas através dos vasos ilíacos foram realizadas usando-se eletrocautério bipolar colocado em uma alta posição, desassociando assim a artéria ilíaca interna e expondo a artéria vesical superior em direção às regiões internas inferiores. A artéria uterina foi então separada da artéria vesical superior , e coagulada e seccionada com eletrocautério bipolar.

iii. Procedimentos cirúrgicos para o coto (ou remanescente) uterino e vaginal: O ligamento redondo (ou ligamentum teres) do útero foi coagulado e resseccionado com um escalpelo ultra-sônico. Após a abertura da reflexão peritoneal da bexiga urinária e corte do fórnice vaginal aberto via vagina, o espaço cervical foi separado da bexiga urinária e do reto. A seguir, o retroperitônio foi aberto por meio de corte. Através de pinçamento separado, os ligamentos uterossacro e cardinal foram cortados antes da ressecção do útero. Uma luva inflada e

esterilizada foi colocada dentro da vagina para garantir a manutenção do pneumoperitônio artificial. O grande omento e o apêndice, que foram resseccionados, foram removidos da vagina. O vértice da vagina foi finalmente fechado usando-se suturas absorvíveis.

iv. Omentectomia e apendicectomia; após a colocação do laparoscópio dentro do trocarte acima da sínfise púbica, o grande omento foi coagulado com um escalpelo ultra-sônico ao longo da margem inferior do cólon transversum (ou cólon transverso). Os vasos sangüíneos localizados na periferia do apêndice foram coagulados com eletrocautério bipolar. A ligação foi realizada na raiz do apêndice usando suturas de ligação não absorvíveis em três ocasiões separadas antes da ressecção do apêndice. A seguir, coagulação bipolar foi realizada no remanescente.

Por outro lado, o Grupo de Laparotomia foi submetido à histerectomia laparoscópica total, salpingooferectomia bilateral, alta ligação da artéria e veia ovarianas, linfadenectomia pélvica (resseccção dos linfonodos pélvicos), omentectomia e apendicectomia realizadas da mesma maneira como os procedimentos cirúrgicos de rotina[1].

3. Análise estatística:

Os dados foram inseridos e analisados usando-se SPSS 9.0 (Pacote Estatístico para Ciências Sociais), com os resultados apresentados em termos de médias e desvios padrão. O teste T de Estudante foi realizado para determinar se as médias entre duas variáveis eram significativas do ponto de vista estatístico, enquanto os dados quantitativos foram analisados usando o teste "X-quadrado" (χ^2).

RESULTADOS

1. Comparação entre as condições intra-operatórias dos dois grupos: o Grupo de Laparoscopia teve um tempo operatório mais longo (ou tempo cirúrgico) do que o Grupo de Laparotomia (p<0.05). Adicionalmente, o Grupo de Laparoscopia teve uma perda sangüínea intra-operatória menor do que o Grupo de Laparotomia (p<0.05). Nenhuma diferença significativa do ponto de vista estatístico no número de linfonodos pélvicos resseccionados foi observada entre os dois grupos (p>0.05; Ver Tabela 2).
2. Comparação das condições pós-operatórias dos dois grupos: o Grupo de Laparoscopia teve uma doença pós-operatória menor do que a do Grupo de Laparotomia (p<0.01). Adicionalmente, o Grupo de Laparoscopia manifestou flatulência pós-cirurgica mais cedo do que o Grupo de Laparotomia (p<0.01). Em relação às complicações pós-operatórias do Grupo de Laparotomia, foi relatado um caso para cada uma das seguintes complicações: urosquese (ou retenção de urina), cistos linfáticos, e inflamação da ferida (s). Adicionalmente, um número menor de casos foi reportado no Grupo de Laparoscopia no uso pós-operatório de analgésicos do que no Grupo de Laparotomia (p<0.01), e mais casos foram relatados para o Grupo de Laparoscopia na mobilização pós-operatória (isto é, deixando seus leitos) em 48 horas do que no Grupo de Laparotomia (p<0.05; Ver Tabela 3).

[1] Cao ZY. Gynecological Oncology (1998). Beijing: Beijing Publishing House, 1005~1011.

Tabela 1: Comparação das Condições Clínicas e Patológicas Pré-Operatórias dos Dois Grupos de Pacientes com Tumor Ovariano

Grupo	Número de casos	Idade (Anos; Média ± Desvio Padrão)	Diâmetro Ovariano (cm; Média ± Desvio Padrão)	Cal 25 (IU/L)	Fase Clínica (No. de Casos)		Condições Pré-Operatórias			
							Tipo Patológico (No. de Casos)		Grau de Diferenciação do Tumor (N. de Casos)	
					Fase I	Fase II ~ Fase IV	Tumor Epitelial	Tumor Gonadal Estromal)	Alto	Baixo
Grupo de Laparoscopia	10	40 ± 8	6.6 ± 2.8	49.8	10	0	9	1	10	0
Grupo de Laparotomia	11	42 ± 6	6.8 ± 5.2	52.6	11	0	9	2	11	0

Tabela 2: Comparação das Condições Intra-Operatórias entre os Dois Grupos de Casos de Tumor Ovariano Maligno (média ± desvio padrão)

Grupo	Número de casos	Tempo de operação (min)	Perda Sangüínea Intra-Operatória (ml)	Número de Linfonodos Pélvicos Resseccionados (N. de Unidades)	Injuries (N. de Casos)
Grupo de Laparoscopia	10	298 ± 60	280 ± 156	25 ± 5	1
Grupo de Laparotomia	11	182 ± 43	346 ± 170	27 ± 6	0
Valor P		$p<0.05$	$p<0.05$	$p>0.05$	

Tabela 3: Comparação das Condições Pós-Operatórias entre os Dois Grupos de Casos de Tumor Ovariano Maligno

Grupo	Número de Casos	Morbidade Pós-Operatória (%)	Tempo para Passagem de Ar Pós-Operatória (Eliminação de gás) (h; Média ± Desvio Padrão)	Complicações Pós-Operatórias			Uso Pós-Operatória de Analgésicos (N. de Casos)	Mobilização Pós-Operatória em 48 horas (N. de Casos)
				Urosquese	Cistos Linfáticos (N. de Casos)	Infecção de ferida		
Grupo de Laparoscopia	10	20.0	13 ± 10	0	0	0	1	7
Grupo de Laparotomia	11	72.7	42 ± 17	1	1	1	10	1
Valor P		< 0.01	< 0.01				< 0.01	< 0.05