Cochrane Library

**Cochrane** Database of Systematic Reviews

# Impact of residual disease as a prognostic factor for survival in women with advanced epithelial ovarian cancer after primary surgery (Review)

Bryant A, Hiu S, Kunonga PT, Gajjar K, Craig D, Vale L, Winter-Roach BA, Elattar A, Naik R

Bryant A, Hiu S, Kunonga PT, Gajjar K, Craig D, Vale L, Winter-Roach BA, Elattar A, Naik R.
Impact of residual disease as a prognostic factor for survival in women with advanced epithelial ovarian cancer after primary surgery.
*Cochrane Database of Systematic Reviews* 2022, Issue 9. Art. No.: CD015048.
DOI: 10.1002/14651858.CD015048.pub2.

**www.cochranelibrary.com**




WILEY

# TABLE OF CONTENTS

[Prognosis Review]

# Impact of residual disease as a prognostic factor for survival in women with advanced epithelial ovarian cancer after primary surgery

Andrew Bryant[1], Shaun Hiu[1], Patience T Kunonga[1], Ketankumar Gajjar[2], Dawn Craig[1], Luke Vale[1], Brett A Winter-Roach[3], Ahmed Elattar[4], Raj Naik[5]

[1]Population Health Sciences Institute, Newcastle University, Newcastle upon Tyne, UK. [2]Department of Gynaecological Oncology, 1st Floor Maternity Unit, City Hospital Campus, Nottingham, UK. [3]The Department of Surgery, Christie Hospital NHS Foundation Trust, Manchester, UK. [4]City Hospital & Birmingham Treatment Centre, Birmingham, UK. [5]Gynaecological Oncology, Northern Gynaecological Oncology Centre, Gateshead, UK

**Contact:** Andrew Bryant, andy.bryant@ncl.ac.uk.

**Editorial group:** Cochrane Gynaecological, Neuro-oncology and Orphan Cancer Group.
**Publication status and date:** New, published in Issue 9, 2022.

**Citation:** Bryant A, Hiu S, Kunonga PT, Gajjar K, Craig D, Vale L, Winter-Roach BA, Elattar A, Naik R. Impact of residual disease as a prognostic factor for survival in women with advanced epithelial ovarian cancer after primary surgery. *Cochrane Database of Systematic Reviews* 2022, Issue 9. Art. No.: CD015048. DOI: 10.1002/14651858.CD015048.pub2.



## ABSTRACT

### Background

Ovarian cancer is the seventh most common cancer among women and a leading cause of death from gynaecological malignancies. Epithelial ovarian cancer is the most common type, accounting for around 90% of all ovarian cancers. This specific type of ovarian cancer starts in the surface layer covering the ovary or lining of the fallopian tube. Surgery is performed either before chemotherapy (upfront or primary debulking surgery (PDS)) or in the middle of a course of treatment with chemotherapy (neoadjuvant chemotherapy (NACT) and interval debulking surgery (IDS)), with the aim of removing all visible tumour and achieving no macroscopic residual disease (NMRD). The aim of this review is to investigate the prognostic impact of size of residual disease nodules (RD) in women who received upfront or interval cytoreductive surgery for advanced (stage III and IV) epithelial ovarian cancer (EOC).

### Objectives

To assess the prognostic impact of residual disease after primary surgery on survival outcomes for advanced (stage III and IV) epithelial ovarian cancer. In separate analyses, primary surgery included both upfront primary debulking surgery (PDS) followed by adjuvant chemotherapy and neoadjuvant chemotherapy followed by interval debulking surgery (IDS). Each residual disease threshold is considered as a separate prognostic factor.

### Search methods

We searched CENTRAL (2021, Issue 8), MEDLINE via Ovid (to 30 August 2021) and Embase via Ovid (to 30 August 2021).

### Selection criteria

We included survival data from studies of at least 100 women with advanced EOC after primary surgery. Residual disease was assessed as a prognostic factor in multivariate prognostic models. We excluded studies that reported fewer than 100 women, women with concurrent malignancies or studies that only reported unadjusted results. Women were included into two distinct groups: those who received PDS followed by platinum-based chemotherapy and those who received IDS, analysed separately. We included studies that reported all RD thresholds after surgery, but the main thresholds of interest were microscopic RD (labelled NMRD), RD 0.1 cm to 1 cm (small-volume residual disease (SVRD)) and RD > 1 cm (large-volume residual disease (LVRD)).

**Data collection and analysis**

Two review authors independently abstracted data and assessed risk of bias. Where possible, we synthesised the data in meta-analysis. To assess the adequacy of adjustment factors used in multivariate Cox models, we used the 'adjustment for other prognostic factors' and 'statistical analysis and reporting' domains of the quality in prognosis studies (QUIPS) tool. We also made judgements about the certainty of the evidence for each outcome in the main comparisons, using GRADE.

We examined differences between FIGO stages III and IV for different thresholds of RD after primary surgery. We considered factors such as age, grade, length of follow-up, type and experience of surgeon, and type of surgery in the interpretation of any heterogeneity.

We also performed sensitivity analyses that distinguished between studies that included NMRD in RD categories of < 1 cm and those that did not. This was applicable to comparisons involving RD < 1 cm with the exception of RD < 1 cm versus NMRD. We evaluated women undergoing PDS and IDS in separate analyses.

**Main results**

We found 46 studies reporting multivariate prognostic analyses, including RD as a prognostic factor, which met our inclusion criteria: 22,376 women who underwent PDS and 3697 who underwent IDS, all with varying levels of RD.

While we identified a range of different RD thresholds, we mainly report on comparisons that are the focus of a key area of clinical uncertainty (involving NMRD, SVRD and LVRD). The comparison involving any visible disease (RD > 0 cm) and NMRD was also important.

*SVRD versus NMRD in a PDS setting*

In PDS studies, most showed an increased risk of death in all RD groups when those with macroscopic RD (MRD) were compared to NMRD. Women who had SVRD after PDS had more than twice the risk of death compared to women with NMRD (hazard ratio (HR) 2.03, 95% confidence interval (CI) 1.80 to 2.29; $I^2$ = 50%; 17 studies; 9404 participants; moderate-certainty). The analysis of progression-free survival found that women who had SVRD after PDS had nearly twice the risk of death compared to women with NMRD (HR 1.88, 95% CI 1.63 to 2.16; $I^2$ = 63%; 10 studies; 6596 participants; moderate-certainty).

*LVRD versus SVRD in a PDS setting*

When we compared LVRD versus SVRD following surgery, the estimates were attenuated compared to NMRD comparisons. All analyses showed an overall survival benefit in women who had RD < 1 cm after surgery (HR 1.22, 95% CI 1.13 to 1.32; $I^2$ = 0%; 5 studies; 6000 participants; moderate-certainty). The results were robust to analyses of progression-free survival.

*SVRD and LVRD versus NMRD in an IDS setting*

The one study that defined the categories as NMRD, SVRD and LVRD showed that women who had SVRD and LVRD after IDS had more than twice the risk of death compared to women who had NMRD (HR 2.09, 95% CI 1.20 to 3.66; 310 participants; $I^2$ = 56%, and HR 2.23, 95% CI 1.49 to 3.34; 343 participants; $I^2$ = 35%; very low-certainty, for SVRD versus NMRD and LVRD versus NMRD, respectively).

*LVRD versus SVRD + NMRD in an IDS setting*

Meta-analysis found that women who had LVRD had a greater risk of death and disease progression compared to women who had either SVRD or NMRD (HR 1.60, 95% CI 1.21 to 2.11; 6 studies; 1572 participants; $I^2$ = 58% for overall survival and HR 1.76, 95% CI 1.23 to 2.52; 1145 participants; $I^2$ = 60% for progression-free survival; very low-certainty). However, this result is biased as in all but one study it was not possible to distinguish NMRD within the < 1 cm thresholds. Only one study separated NMRD from SVRD; all others included NMRD in the SVRD group, which may create bias when comparing with LVRD, making interpretation challenging.

*MRD versus NMRD in an IDS setting*

Women who had any amount of MRD after IDS had more than twice the risk of death compared to women with NMRD (HR 2.11, 95% CI 1.35 to 3.29, $I^2$ = 81%; 906 participants; very low-certainty).

**Authors' conclusions**

In a PDS setting, there is moderate-certainty evidence that the amount of RD after primary surgery is a prognostic factor for overall and progression-free survival in women with advanced ovarian cancer. We separated our analysis into three distinct categories for the survival outcome including NMRD, SVRD and LVRD.

After IDS, there may be only two categories required, although this is based on very low-certainty evidence, as all but one study included NMRD in the SVRD category. The one study that separated NMRD from SVRD showed no improved survival outcome in the SVRD category, compared to LVRD. Further low-certainty evidence also supported restricting to two categories, where women who had any amount of MRD after IDS had a significantly greater risk of death compared to women with NMRD.

Therefore, the evidence presented in this review cannot conclude that using three categories applies in an IDS setting (very low-certainty evidence), as was supported for PDS (which has convincing moderate-certainty evidence).

## PLAIN LANGUAGE SUMMARY

### The impact of remaining (residual) disease after surgery on the survival prognosis for women with advanced epithelial ovarian cancer

#### Review question

We aimed to assess the effect on survival (the 'prognostic impact') of the amount of disease remaining after surgery (residual disease) during the initial treatment stage for women with advanced ovarian cancer. We looked at both surgery before chemotherapy ('primary debulking surgery') followed by adjuvant (additional) chemotherapy and chemotherapy first ('neoadjuvant chemotherapy') followed by surgery ('interval debulking surgery'). This review should help to determine the prognostic impact of residual disease after surgery on survival and work out acceptable definitions of residual disease thresholds.

#### Background

Ovarian cancer is the seventh most common cancer among women and a leading cause of death in women with gynaecological cancers. Ovarian cancers can develop from different cell types within the ovary/fallopian tubes. Most ovarian cancers are 'epithelial', arising from either the surface layer of the ovary or the lining of the fallopian tube. Newly diagnosed ovarian cancer is treated with a combination of surgery and chemotherapy, with surgery performed either before (called upfront or primary debulking surgery) or around the mid-point of chemotherapy (called interval debulking surgery). Ovarian cancer has normally spread throughout the abdominal cavity by the time of diagnosis, so, unlike many other cancers, surgery is still performed, even though it may not remove the cancer in its entirety. The aim of surgery is to remove as much of the visible (macroscopic) cancer tissue as possible, which is called debulking or cytoreductive surgery. Studies have shown that the amount of the visible cancer that can be removed is likely to be an important prognostic factor for survival of women with advanced epithelial ovarian cancer. The aim of this review was to investigate how well the amount of remaining (residual) disease after surgery for newly diagnosed ovarian cancer predicts how long women will survive following a diagnosis of epithelial ovarian cancer (prognosis).

#### Review methods

We searched electronic databases up to the end of August 2021 and we also searched for unpublished studies. We included studies that reported residual disease as a prognostic factor, which also examined other prognostic factors at the same time.

#### Key results

We found 46 studies (including 22,376 women in 31 primary debulking surgery studies and 3697 women in 15 interval debulking surgery studies). Each study included more than 100 women, used statistical adjustment for important prognostic factors (multivariate analysis) and met our inclusion criteria. Our analyses showed the prognostic importance of surgery leaving no visible tumour deposits ('no macroscopic residual disease') both when women had upfront debulking surgery or interval debulking surgery. Both overall survival and progression-free survival (survival without disease worsening, which was reported for upfront debulking surgery) were prolonged if this was achieved.

*Primary debulking surgery for newly diagnosed ovarian cancer*

Complete surgical removal of all visible tumour after upfront or primary debulking surgery improved survival, and this was also the case for those with a small amount of residual disease (0.1 cm to 1 cm). There was evidence to suggest that three categories of residual disease should be used (no macroscopic residual disease, small-volume and large-volume residual disease (more than 1 cm).

*Interval debulking surgery for newly diagnosed ovarian cancer*

When chemotherapy was given before surgery (interval debulking surgery), there was an association with improved survival if the remaining tumour was reduced to 'no macroscopic residual disease' (removal of all visible tumour). Women with small-volume residual disease had no survival advantage compared to those with large-volume residual disease, with both groups having a poorer prognosis compared to those with no visible tumour deposits; however, this evidence was of very low certainty. Any visible residual disease after interval debulking surgery was associated with poorer survival compared to women with none.

Most interval debulking surgery studies included no visible tumour deposits in the small-volume residual disease category, which limits our interpretation of these findings.

#### Certainty of the evidence

We judged our certainty of the evidence as 'moderate' for overall survival and progression-free survival in the analyses involving primary debulking surgery studies. For the interval debulking surgery studies, the certainty of evidence was very low for overall survival in all

comparisons and those that involved progression-free survival. This was largely due to all but one study including 'no macroscopic residual disease' in the small-volume residual disease category.

**Main conclusions**

The evidence in the review suggests that following primary debulking surgery three categories for the amount of residual disease should be used: no macroscopic residual disease, small-volume and large-volume residual disease. The evidence is more limited for interval debulking surgery and further studies are needed, but there may not be a survival difference between those with small- and large-volume residual disease. Until there is evidence for a survival benefit for those with small-volume compared to large-volume residual disease, it may only be important to use two residual disease categories when classifying surgical outcomes: 'no macroscopic residual disease' and 'macroscopic residual disease' (remaining visible disease of more than 0 cm). However, this is based on very low-certainty evidence and more information may change this finding.

## SUMMARY OF FINDINGS

### Summary of findings 1. Small-volume residual disease (SVRD) < 1 cm versus NMRD in PDS studies

**Small-volume residual disease (SVRD) (< 1 cm) compared with NMRD after upfront primary debulking surgery (PDS) in women with advanced epithelial ovarian cancer (EOC)**

**Population:** women with advanced EOC after PDS

**Settings:** all settings in adult women aged 18 years or older worldwide

**Prognostic factor:** SVRD compared with NMRD

| Outcomes | Relative effect (95% CI) | No of participants (studies) | Quality of the evidence (GRADE) | Comments |
|---|---|---|---|---|
| **Overall survival:**<br>Median length of follow-up[1]:<br>Range: 28 to 77.7 months | **Adjusted HR 2.03** (1.80 to 2.29) | 9404 participants (17 studies) | ⊕⊕⊕⊖ **moderate**[2] | We could not present illustrative absolute effects because a representative control group risk could not be ascertained from the studies. The HR estimates were adjusted for in multivariable analyses and this cannot be done in absolute terms so we did not attempt it, as the numbers were likely to mislead with any bias potentially favouring the NMRD threshold. |
| **Progression-free survival:**<br>Median length of follow-up[1]:<br>Range: 28 to 77.7 months | **Adjusted HR 1.88** (1.63 to 2.16) | 6596 participants (10 studies) | ⊕⊕⊕⊖ **moderate**[2] | There were no concerns with inconsistency and imprecision across studies due to restrictive inclusion criteria in a generally representative cohort of women with advanced EOC. Data were considerable in size in PDS studies with > 9000 and > 6500 women in the analyses of OS and PFS, respectively.<br><br>The percentage of the variability in effect estimates that was due to heterogeneity rather than sampling error (chance) may appear to represent moderate heterogeneity (as measured by the $I^2$ statistic), but we had no major concerns as the direction of effect was consistent throughout.<br><br>There did not appear to be any evidence of small study biases, such as publication bias, or any irregularities with the data by visual inspection of funnel plots. While publication bias cannot be dismissed, it would take a lot of large statistically insignificant studies to overhaul the current results. Furthermore, studies showing harmful survival in women with NMRD compared to other thresholds of RD is implausible. |

**CI:** confidence interval; **HR:** hazard ratio; **EOC:** epithelial ovarian cancer; **NMRD:** no macroscopic residual disease; **OS:** overall survival; **PDS:** upfront primary debulking surgery; **PFS:** progression-free survival; **SVRD:** small-volume residual disease

GRADE Working Group grades of evidence
**High quality:** Further research is very unlikely to change our confidence in the estimate of effect.

**Moderate quality:** Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
**Low quality:** Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
**Very low quality:** We are very uncertain about the estimate.

[1]Range in Klar 2016 was 0 to 144 months.
[2]Downgraded by one level because was assessed the statistical analysis and reporting domain in the QUIPS tool as being at high or unclear risk of bias in all included studies. Either no conceptual framework was reported, where the variable selection criteria in multivariate model was unclear, or quite often the authors reported that significant variables from the univariate analysis were included in the multivariable model, but with no further details. This was the most serious bias from the QUIPS domains that could influence the effect estimates.

### Summary of findings 2. Large-volume residual disease (LVRD) > 1 cm versus no macroscopic residual disease (NMRD) in PDS studies

**LVRD (> 1 cm) compared with NMRD after upfront primary debulking surgery (PDS) in women with advanced epithelial ovarian cancer (EOC)**

**Population:** women with advanced ovarian cancer after PDS

**Settings:** all settings in adult women aged 18 years or older worldwide

**Prognostic factor:** LVRD > 1 cm compared with NMRD

| Outcomes | Relative effect (95% CI) | No of participants (studies) | Quality of the evidence (GRADE) | Comments |
|---|---|---|---|---|
| **Overall survival:**<br>Median length of follow-up:<br>Range: 28 to 77.7 months | **Adjusted HR**<br>**2.50** (2.13 to 2.94) | 7988 participants (14 studies) | ⊕⊕⊕⊖<br>**moderate**[1] | We could not present illustrative absolute effects because a representative control group risk could not be ascertained from the studies. The HR estimates were adjusted for in multivariable analyses and this cannot be done in absolute terms so we did not attempt it, as the numbers were likely to mislead with any bias potentially favouring the NMRD threshold. |
| **Progression-free survival:**<br>Median length of follow-up:<br>Range: 28 to 77.7 months | **Adjusted HR**<br>**2.10** (1.84 to 2.40) | 2629 participants (6 studies) | ⊕⊕⊕⊖<br>**moderate**[1] | There were no concerns with inconsistency and imprecision across studies due to restrictive inclusion criteria in a generally representative cohort of women with advanced EOC. Data were considerable in size in PDS studies with nearly n = 8000 in the analysis of OS and to lesser extent > 2500 for PFS.<br><br>The percentage of the variability in effect estimates that was due to heterogeneity rather than sampling error (chance) may appear to represent moderate heterogeneity (as measured by the $I^2$ statistic), but we had no major concerns as the direction of effect was consistent throughout.<br><br>There did not appear to be any evidence of small study biases, such as publication bias, or any ir- |

| | |
|---|---|
| | regularities with the data by visual inspection of funnel plots. While publication bias cannot be dismissed, it would take a lot of large statistically insignificant studies to overhaul the current results. Furthermore, studies showing harmful survival in women with NMRD compared to other thresholds of RD is implausible. |

**CI:** confidence interval; **HR:** hazard ratio; **EOC:** epithelial ovarian cancer; **LVRD**: large-volume residual disease; **NMRD:** no macroscopic residual disease; **OS:** overall survival; **PDS:** upfront primary debulking surgery; **PFS:** progression-free survival

GRADE Working Group grades of evidence
**High quality:** Further research is very unlikely to change our confidence in the estimate of effect.
**Moderate quality:** Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
**Low quality:** Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
**Very low quality:** We are very uncertain about the estimate.

[1]Downgraded by one level because we assessed the statistical analysis and reporting domain in the QUIPS tool as being at high or unclear risk of bias in all included studies. Either no conceptual framework was reported, where the variable selection criteria in multivariate model was unclear, or quite often the authors reported that significant variables from the univariate analysis were included in the multivariable model, but with no further details. This was the most serious bias from the QUIPS domains that could influence the effect estimates.

### Summary of findings 3. Large-volume residual disease (LVRD) > 1 cm versus small-volume residual disease (SVRD) < 1 cm in PDS studies

**LVRD (> 1 cm) compared with SVRD (< 1 cm) after upfront primary debulking surgery (PDS) in women with advanced epithelial ovarian cancer (EOC)**

**Population:** women with advanced EOC after PDS

**Settings:** all settings in adult women aged 18 years or older worldwide

**Prognostic factor:** LVRD > 1 cm compared with SVRD < 1 cm

| Outcomes | Relative effect (95% CI) | No of participants (studies) | Quality of the evidence (GRADE) | Comments |
|---|---|---|---|---|
| **Overall survival:**<br>Median length of follow-up[1]:<br>Range: 28 to 34.1 months | **Adjusted HR 1.22** (1.13 to 1.32) | 6000 participants (5 studies) | ⊕⊕⊕⊖ **moderate**[2] | We could not present illustrative absolute effects because a representative control group risk could not be ascertained from the studies. The HR estimates were adjusted for in multivariable analyses and this cannot be done in absolute terms, so we did not attempt it as the numbers were likely to mislead with any bias potentially favouring the NMRD threshold. |
| **Progression-free survival:**<br>Median length of follow-up[1]: 28 months | **Adjusted HR**<br>1.30 (1.08 to 1.56) | 3402 participants (2 studies) | ⊕⊕⊕⊖ **moderate**[2] | There were no concerns with inconsistency and imprecision across studies (the smallest study comparison (n = 100) was imprecise but there were only n = 23 women with sub-optimal RD) due to restrictive inclusion criteria in a generally representative cohort of women with advanced EOC. Data were considerable in size in PDS studies with n > 6000 in the analysis of OS and to lesser extent > 3000 for PFS. |

| | |
|---|---|
| | The percentage of the variability in effect estimates that was due to heterogeneity rather than sampling error (chance) may not be important (as measured by the $I^2$ statistic) in meta-analyses including PDS studies. |

**CI:** confidence interval; **HR:** hazard ratio; **EOC:** epithelial ovarian cancer; **LVRD**: large-volume residual disease; **NMRD:** no macroscopic residual disease; **OS:** overall survival; **PDS:** upfront primary debulking surgery; **PFS:** progression-free survival; **SVRD:** small-volume residual disease

GRADE Working Group grades of evidence
**High quality:** Further research is very unlikely to change our confidence in the estimate of effect.
**Moderate quality:** Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
**Low quality:** Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
**Very low quality:** We are very uncertain about the estimate.

[1]Range in Klar 2016 was 0 to 144 months.
[2]Downgraded by one level because we assessed the statistical analysis and reporting domain in the QUIPS tool as being at high or unclear risk of bias in all included studies. Either no conceptual framework was reported, where the variable selection criteria in multivariate model was unclear, or quite often the authors reported that significant variables from the univariate analysis were included in the multivariable model, but with no further details. This was the most serious bias from the QUIPS domains that could influence the effect estimates.

### Summary of findings 4. Small-volume residual disease (SVRD) (< 1 cm) versus NMRD in IDS studies

**SVRD (< 1 cm) compared with NMRD after primary interval debulking surgery (IDS) in women with advanced epithelial ovarian cancer (EOC)**

**Population:** women with advanced EOC after primary IDS

**Settings:** all settings in adult women aged 18 years or older worldwide

**Prognostic factor:** SVRD < 1 cm compared with NMRD

| Outcomes | Relative effect (95% CI) | No of participants (studies) | Quality of the evidence (GRADE) | Comments |
|---|---|---|---|---|
| **Overall survival:**<br><br>Median length of follow-up<br><br>Not reported | **Adjusted HR 2.09** (1.20 to 3.60) | 310 participants (1 study reporting on 2 groups) | ⊕⊖⊖⊖<br>**very low[123]** | We could not present illustrative absolute effects because a representative control group risk could not be ascertained from the studies. The HR estimates were adjusted for in multivariable analyses and this cannot be done in absolute terms so we did not attempt it, as the numbers were likely to mislead with any bias potentially favouring the NMRD threshold. |
| **Progression-free survival:**<br><br>Median length of follow-up: 47 months | P = 0.001 | 322 participants (1 study) | ⊕⊖⊖⊖<br>**very low[123]** | The authors of Petrillo 2014 found that the risk of disease progression for women with RD < 1 cm after IDS was significantly higher than those with complete cytoreduction, but the magnitude of effect was not reported. |

| | | | | |
|---|---|---|---|---|
| | Range: 3 to 181 months | | | |

**CI:** confidence interval; **HR:** hazard ratio; **EOC:** epithelial ovarian cancer; **IDS:** interval debulking surgery; **NMRD:** no macroscopic residual disease; **OS:** overall survival; **PDS:** upfront primary debulking surgery; **PFS:** progression-free survival; **SVRD:** small-volume residual disease

GRADE Working Group grades of evidence
**High quality:** Further research is very unlikely to change our confidence in the estimate of effect.
**Moderate quality:** Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
**Low quality:** Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
**Very low quality:** We are very uncertain about the estimate.

[1]Downgraded by one level because we assessed the statistical analysis and reporting domain in the QUIPS tool as being at high or unclear risk of bias in all included studies. Either no conceptual framework was reported, where the variable selection criteria in multivariate model was unclear, or quite often the authors reported that significant variables from the univariate analysis were included in the multivariable model, but with no further details. This was the most serious bias from the QUIPS domains that could influence the effect estimates.
[2]Downgraded by one level for sparse data.
[3]Downgraded by one level for lack of generalisability and validity of results as reported in single analysis or very few included studies.

### Summary of findings 5. Large-volume residual disease (LVRD) > 1 cm versus no macroscopic residual disease (NMRD) in IDS studies

**Large-volume residual disease (LVRD) (> 1 cm) compared with NMRD after primary interval debulking surgery (IDS) in women with advanced epithelial ovarian cancer (EOC)**

**Population:** women with advanced EOC after primary IDS

**Settings:** all settings in adult women aged 18 years or older worldwide

**Prognostic factor:** LVRD > 1 cm compared with NMRD

| Outcomes | Relative effect (95% CI) | No of participants (studies) | Quality of the evidence (GRADE) | Comments |
|---|---|---|---|---|
| **Overall survival:**<br><br>Median length of follow-up:<br><br>Not reported | **Adjusted HR 2.23** (1.49 to 3.34) | 343 participants (1 study reporting on 2 groups) | ⊕⊖⊖⊖<br>**very low**[123] | We could not present illustrative absolute effects because a representative control group risk could not be ascertained from the studies. The HR estimates were adjusted for in multivariable analyses and this cannot be done in absolute terms, so we did not attempt it as the numbers were likely to mislead with any bias potentially favouring the NMRD threshold. |
| **Progression-free survival** | Not reported | | | |

**CI:** confidence interval; **HR:** hazard ratio; **EOC:** epithelial ovarian cancer; **IDS:** interval debulking surgery; **LVRD:** large-volume residual disease; **NMRD:** no macroscopic residual disease; **OS:** overall survival; **PDS:** upfront primary debulking surgery; **PFS:** progression-free survival

GRADE Working Group grades of evidence

**High quality:** Further research is very unlikely to change our confidence in the estimate of effect.
**Moderate quality:** Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
**Low quality:** Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
**Very low quality:** We are very uncertain about the estimate.

[1]Downgraded by one level because we assessed the statistical analysis and reporting domain in the QUIPS tool as being at high or unclear risk of bias in all included studies. Either no conceptual framework was reported, where the variable selection criteria in multivariate model was unclear, or quite often the authors reported that significant variables from the univariate analysis were included in the multivariable model, but with no further details. This was the most serious bias from the QUIPS domains that could influence the effect estimates.
[2]Downgraded by one level for sparse data.
[3]Downgraded by one level for lack of generalisability and validity of results as reported in single analysis or very few included studies.

### Summary of findings 6. Large-volume residual disease (LVRD) > 1 cm versus small-volume residual disease (SVRD) < 1 cm in IDS studies

**Large-volume residual disease (LVRD) > 1 cm compared with small-volume residual disease (SVRD) < 1 cm after primary interval debulking surgery (IDS) in women with advanced epithelial ovarian cancer (EOC)**

**Population:** women with advanced EOC after primary IDS

**Settings:** all settings in adult women aged 18 years or older worldwide

**Prognostic factor:** LVRD > 1 cm compared with SVRD < 1 cm

| Outcomes | Relative effect (95% CI) | No of participants (studies) | Quality of the evidence (GRADE) | Comments |
|---|---|---|---|---|
| **Overall survival:**<br>Median length of follow-up:<br>Range: 34.3 to 43.5 months | **Adjusted HR 1.60** (1.21 to 2.11) | 1572 participants (6 studies) | ⊕⊕⊕⊖ **verylow**[123] | We could not present illustrative absolute effects because a representative control group risk could not be ascertained from the studies. The HR estimates were adjusted for in multivariable analyses and this cannot be done in absolute terms, so we did not attempt it as the numbers were likely to mislead with any bias potentially favouring the NMRD threshold.<br>The percentage of the variability in effect estimates that was due to heterogeneity rather than sampling error (chance) may represent substantial heterogeneity (as measured by the $I^2$ statistic) in meta-analyses. |
| **Progression-free survival:**<br>Median length of follow-up<br>Range: 38 to 43.5 months | **Adjusted HR 1.76** (1.23 to 2.52) | 1145 participants (4 studies) | ⊕⊕⊕⊖ **verylow**[123] | |

**CI:** confidence interval; **HR:** hazard ratio; **EOC:** epithelial ovarian cancer; **IDS:** interval debulking surgery; **LVRD:** large-volume residual disease; **NMRD:** no macroscopic residual disease; **OS:** overall survival; **PDS:** upfront primary debulking surgery; **PFS:** progression-free survival; **SVRD:** small-volume residual disease

GRADE Working Group grades of evidence
**High quality:** Further research is very unlikely to change our confidence in the estimate of effect.
**Moderate quality:** Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
**Low quality:** Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

**Very low quality:** We are very uncertain about the estimate.

[1]Downgraded by one level because we assessed the statistical analysis and reporting domain in the QUIPS tool as being at high or unclear risk of bias in all included studies. Either no conceptual framework was reported, where the variable selection criteria in multivariate model was unclear, or quite often the authors reported that significant variables from the univariate analysis were included in the multivariable model, but with no further details. This was the most serious bias from the QUIPS domains that could influence the effect estimates.
[2]Downgraded by one level for heterogeneity across studies.
[3]Only one study reported a comparison of SVRD < 1 cm versus LVRD > 1 cm in the strict sense that SVRD < 1 cm was mutually exclusive of NMRD (Phillips 2018).

### Summary of findings 7. Residual disease (RD) > 0 cm versus NMRD in IDS studies

**Any remaining residual disease (RD) (> 0 cm) compared with NMRD after primary interval debulking surgery (IDS) in women with advanced epithelial ovarian cancer (EOC)**

**Population:** women with advanced EOC after primary IDS

**Settings:** all settings in adult women aged 18 years or older worldwide

**Prognostic factor:** RD > 0 cm compared with NMRD

| Outcomes | Relative effect (95% CI) | No of participants (studies) | Quality of the evidence (GRADE) | Comments |
|---|---|---|---|---|
| **Overall survival:**<br><br>Median length of follow-up: range: 37 to 39 (reported in 2 studies) | **Adjusted HR 2.11** (1.35 to 3.29) | 906 participants (4 studies) | ⊕⊖⊖⊖<br>**very low[123]** | We could not present illustrative absolute effects because a representative control group risk could not be ascertained from the studies. The HR estimates were adjusted for in multivariable analyses and this cannot be done in absolute terms, so we did not attempt it as the numbers were likely to mislead with any bias potentially favouring the NMRD threshold.<br><br>The percentage of the variability in effect estimates that was due to heterogeneity rather than chance may represent considerable heterogeneity ($I^2$ = 81%).<br><br>The authors of Lecuru 2019 additionally found that the risk of death for women with any remaining RD (> 0 cm) after IDS was significantly higher than those with NMRD (n = 163, P < 0.01), but the magnitude of effect was not reported. |
| **Progression-free survival:**<br><br>Median length of follow-up: not reported | **Adjusted HR 1.36** (1.05 to 1.76) | 471 participants (1 study) | ⊕⊖⊖⊖<br>**very low[123]** | The authors of Lecuru 2019 additionally found that the risk of disease progression for women with RD > 0 cm after IDS was significantly higher than those with NMRD (n = 163, P < 0.01), but the magnitude of effect was not reported. |

**CI:** confidence interval; **HR:** hazard ratio; **EOC:** epithelial ovarian cancer; **IDS:** interval debulking surgery; **NMRD:** no macroscopic residual disease; **OS:** overall survival; **PDS:** upfront primary debulking surgery; **PFS:** progression-free survival

GRADE Working Group grades of evidence
**High quality:** Further research is very unlikely to change our confidence in the estimate of effect.

**Moderate quality:** Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
**Low quality:** Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
**Very low quality:** We are very uncertain about the estimate.

[1]Downgraded by one level because we assessed the statistical analysis and reporting domain in the QUIPS tool as being at high or unclear risk of bias in all included studies. Either no conceptual framework was reported, where the variable selection criteria in multivariate model was unclear, or quite often the authors reported that significant variables from the univariate analysis were included in the multivariable model, but with no further details. This was the most serious bias from the QUIPS domains that could influence the effect estimates.
[2]Downgraded by one level for heterogeneity across studies.
[3]Downgraded by one level for lack of generalisability and validity of results as reported in single analysis or very few included studies.

# BACKGROUND

## Description of the health condition and context

Ovarian cancer is the seventh most common cancer among women and a leading cause of death in women with gynaecological malignancies (GLOBOCAN 2018). Globally, there are approaching 300,000 new cases per year, with approximately 6.6 new cases per 100,000 women per year. A woman's cumulative risk of developing ovarian cancer by the age of 75 years is 0.72%: 0.52% in low-income countries and 0.92% in high-income countries (GLOBOCAN 2018). Ovarian cancer is rare in women under 40 years of age and most cancers in this age group are germ cell tumours. Above age 40, more than 90% are epithelial tumours and the risk increases with age (Kurman 2014; Webb 2017). Epithelial ovarian cancer is the most common type, accounting for around 90% of all ovarian cancers. This specific type of ovarian cancer starts in the surface layer covering the ovary or lining of the fallopian tube.

Ovarian cancer is best regarded as a peritoneal malignancy. The current understanding on the pathogenesis of epithelial ovarian cancer (EOC) recognises two pathways and two clinical groupings, classified as Type 1 and Type 2. Type 1 tumours comprise low-grade serous, low-grade endometrioid, clear-cell and mucinous carcinomas, and Brenner tumours. Type 2 tumours comprise the high-grade serous and endometrioid carcinomas, mixed mullerian tumours and undifferentiated carcinomas. Type 2 tumours are more common and are thought to have their origin within the fallopian tube (Perets 2016). They are associated with the BRCA (breast cancer gene) germline and somatic mutations, and histopathologically identified with aberrant p53 expression and other characteristic immunohistochemical features (Kurman 2010; Kurman 2011).

The extent of dissemination of the disease is described using the International Federation of Gynecology and Obstetrics (FIGO) staging system; stage I disease is confined to the ovaries; stage II disease is confined to the true pelvis, stage III disease is an abdominal disease where there is spread to the lining (peritoneum) of the abdominal cavity outside the pelvis or regional lymph node spread; whilst stage IV disease is outside the abdomen or parenchymatous metastases, e.g. disease with spread to distant organs such as the chest or liver (Berek 2018). Thirty per cent of women with ovarian cancer present with early-stage disease, whilst 70% have advanced stage at presentation (Torre 2018). In Europe, just over a third of women with ovarian cancer are alive five years after diagnosis (EUROCARE 2015), largely because most women with ovarian cancer are diagnosed when the cancer is already at an advanced stage (Jemal 2017). This is, in part, due to the biology of the disease and immediate acces to the abdominal cavity and non-specific symptoms, which include progressive feelings of: abdominal distension, bloating, indigestion, urinary frequency, urgency, early satiety, weight loss, reduced appetite, abdominal and pelvic pain and, less commonly, vaginal bleeding (Shafi 2018).

## Description of the surgical interventions and residual disease as a prognostic factor

Surgery and chemotherapy are the mainstay of treatment for the 70% of women who present with advanced disease (FIGO stage III/IV) when surgery alone cannot be curative (Fader 2007; Torre 2018).

Appropriate initial investigations usually include ultrasonography, tumour markers and a CT scan, if malignancy is suggested by tumour markers and ultrasound. If required, an ultrasound-guided biopsy of metastatic spread is carried out to obtain histological diagnosis (Shafi 2018).

Traditionally, upfront debulking surgery (PDS) is performed to remove as much visible disease as possible, as the amount of residual tumour is one of the most important prognostic factors for survival of epithelial ovarian cancer (Bristow 2002; Chang 2013; du Bois 2009; Griffiths 1975; Hoskins 1994; Wimberger 2010). Platinum-based chemotherapy is the standard of care, in combination with debulking surgery (Colombo 2019; National Comprehensive Cancer Network 2020).

Chemotherapy followed by interval debulking surgery (IDS) is an alternative primary treatment option for women diagnosed with advanced ovarian cancer. A Cochrane Review, which comprised five randomised controlled trials (RCTs), comprehensively reviewed the evidence in this area (Coleridge 2021). The review assessed survival, quality of life and morbidity outcomes in trials that compared upfront primary and interval debulking surgery. The five trials included two large, well-documented RCTs (CHORUS (Kehoe 2015) and EORTC 55971 (Vergote 2010)), which reported no significant difference in survival between IDS compared with PDS. It was suggested that IDS may have better overall survival in stage IV disease. One included study suggested that women with FIGO stage IIIC disease with extrapelvic metastases smaller than 5 cm may have better progression-free survival after upfront debulking (Vergote 2018). The selection of women with advanced ovarian cancer for PDS or IDS remains controversial (Vergote 2013). An investigation of maximum effort cytoreductive surgery during the initial treatment of epithelial ovarian cancer comparing PDS versus IDS is being investigated in the TRUST trial (Trial of Radical Upfront Surgical Therapy in advanced ovarian cancer (ENGOT ov33/AGO-OVAR OP7)), and results are expected in 2024 (Reuss 2019).

The terms cytoreductive and debulking surgery are often used interchangeably to indicate surgical efforts aimed at removing the bulk of the tumour. No macroscopic residual disease (NMRD) (also known as 'complete' macroscopic resection or R0) is achieved when there is no visible tumour left at the end of surgery. Previously, the term 'optimal cytoreduction' had been variably defined as referring to a maximal diameter of residual tumour left behind after surgery measuring 0 to 2 cm, and in 1994 the Gynaecologic Oncology Group (GOG) defined optimal cytoreduction as having residual disease < 1 cm (Hoskins 1994). However, in 2010 the Gynaecological Cancer Inter-Group defined 'optimal' as having no visible residual tumour nodules, i.e. NMRD ('complete' is a misnomer as microscopic disease remains in the majority of patients) (Stuart 2011), which has been shown to result in better survival than small-volume residual disease (SVRD) to < 1 cm (also referred to as near-optimal) and large-volume residual disease (LVRD) which is > 1 cm (also referred to as suboptimal) and to be a better predictor of survival (Bookman 2009; Chang 2013; du Bois 2009; Sørensen 2019; Wimberger 2010). While there is now less controversy about the prognostic importance of maximum cytoreduction, there remains divided opinion about the effects of any remaining residual disease after PDS or IDS, and about what attempts should be made for maximal efforts at debulking. All women would potentially do better if there was NMRD after surgery, and obviously no surgeon sets out for suboptimal cytoreduction from the onset. However,

different philosophies are evident within the surgical community and there are also other important considerations, such as surgical skills and training, surgical and critical care resources, the woman's fitness for more radical treatment, morbidity, mortality and quality of life. The questions about PDS in ovarian cancer that appear to have become more important and relevant over the last 10 years of practice as other evidence has emerged relate to the timing of maximal surgical effort (still within initial treatment phase), and to consideration of whether there are some histological subtypes that may have better outcomes with PDS. In this review we only consider the epithelial subtype of ovarian cancer, since it comprises 90% of histological subtypes.

Surgery to achieve NMRD appears to be associated with the best chance of prolonged survival (Bookman 2009). An attempt to achieve NMRD is the recommended standard for cytoreductive surgery for advanced ovarian cancer, as advised by the British Gynaecological Cancer Society (BGCS) (BGCS 2017), European Society of Medical Oncology (ESMO) and European Society of Gynaecological Oncology (ESGO) (Colombo 2019), and the National Comprehensive Cancer Network (NCCN) (National Comprehensive Cancer Network 2020).

A Cochrane Review assessed the role of a further attempt at cytoreductive effort after LVRD remained after primary surgery (Tangjitgamol 2016). The results from three studies in the review found that a further attempt at cytoreductive surgery after chemotherapy in first-line treatment was only of benefit to those who had not had their initial surgery performed by a gynaecological oncologist (Redman 1994; Rose 2004; Van der Burg 1995).

Over the last few decades, efforts have been made to increase NMRD resection rates. It has been shown that surgery performed by gynaecologists with training in gynaecological oncology, by high-volume surgeons and high-volume centres, is associated with increased likelihood of NMRD (Bristow 2009; Greggi 2016; Woo 2012).

There is a widespread belief that tumour biology has a significant role to play in ovarian cancer outcomes. The relationship between surgical outcome and tumour biology is complex and remains unclear. The biological rationale behind the benefit of surgical cytoreduction is that removal of certain ovarian cancer tumour cells will create a supportive microenvironment to enhance chemotherapy effect (Covens 2000; Napoletano 2010). Whether it is the intrinsic biological behaviour of the tumour or the surgeon's ability to cytoreduce that determines optimal cytoreduction is not well studied. However, among the relevant prognostic factors, the extent of surgery and consequent residual disease are the most important prognostic factors. The extent of surgical effort (standard versus extensive surgery) to achieve NMRD and its impact on survival is not fully understood, as determined by a previous Cochrane Review (Hui 2022).

Within the advanced ovarian cancer group, women with stage IV ovarian cancer represent a heterogeneous group with extraperitoneal metastases. While it has been shown in a previously published guideline that NMRD resection is associated with the best chance of prolonged survival (Vergote 2016), the data are not as convincing for stage IV ovarian cancer. The presence of microscopic disease in the extraperitoneal locations has not been assessed and can potentially be even more frequent. While some stage IV diseases could be amenable to resection to NMRD (isolated splenic parenchymal lesion or resectable liver metastasis), others could be difficult (extensive mediastinal, axillary, or supraclavicular nodes or multiple, unresectable hepatic metastases). Therefore, it is worth investigating the impact of residual disease in stage IV cancers, and in particular in relation to extra-peritoneal residual disease (thoracic, mediastinum, groin, axilla, neck). The EORTC55971 trial confirmed that neoadjuvant chemotherapy results in superior survival compared with primary debulking surgery in the management of women with stage IV disease (Vergote 2010). However, there is a need for further investigation into the impact of residual disease on survival between the PDS and IDS subgroups.

This review sets out to determine the prognostic impact of residual disease on survival rates in women with advanced epithelial ovarian cancer. There are no universally established patient selection criteria, but certain baseline characteristics are important when investigating the impact of residual disease on prognosis. These include age, nutritional status, FIGO stage, comorbidities, ASA score (American Society of Anaesthesiologists' (ASA) classification of Physical Health), ECOG (Eastern Cooperative Oncology Group) performance status (score of symptom and functional status with respect to ambulatory status and need for care), BRCA status, presence of ascites on preoperative imaging and histological grade (du Bois 2009). To date, there are no specific predictive models for surgical success that are clinically useful, and the majority of previous studies have limitations in design that make their interpretation difficult (Borley 2012).

If the surgical outcome and prognosis are to be determined by tumour biology alone, the residual disease after surgery may have little influence on overall survival. However, tumour biology and the extent of disease may influence the likelihood of achieving NMRD after surgery (Colombo 2019). The extent of residual disease and prognosis could be influenced by the extent of disease measured intraoperatively by the peritoneal cancer index (PCI) score, surgical complexity score (SCS) (Elzarkaa 2018), type and extent of surgery (Aletti 2007), characteristics of the surgical team (gynaecological oncologist in a specialist centre with a high volume of cases) (Bristow 2009) and presence of ascites during surgery (du Bois 2009).

## Why it is important to do this review

A greater understanding of the biology of ovarian cancer variants, especially with respect to BRCA gene mutations, has led to more sophisticated treatment regimens. These include the emergence of tailored adjuvant and maintenance chemotherapeutic options for women with BRCA somatic and germline mutations, and greater options for the chemotherapeutic approach to recurrent disease (Colombo 2019).

While the place of surgery in the context of treatment of ovarian cancer is well established, the distinctive biological phenotypes (e.g. type and grade of disease, extent of disease) should be anticipated to lead to some heterogeneity in the level of benefit derived from maximal surgical effort. There may be a greater willingness to rely on PDS for women with known subtypes of disease, such as low-grade serous cancer, that are known to be less chemo-responsive (Grabowski 2016). PDS for highly chemo-responsive disease has also been questioned by a growing acceptance of the non-inferiority of interval debulking surgery

(Coleridge 2021). The current position in many settings, in the UK and elsewhere, is to reserve PDS in advanced disease for those women who have a good performance status, and in whom it is anticipated that NMRD or SVRD can be achieved. Performance status is relevant in consideration of PDS. Though true advocates of PDS remain, many clinicians recognise that women presenting with poor performance status are likely to be too frail to undergo a PDS without significant comorbidity. In such a situation, clinical optimisation and initiation of treatment with chemotherapy is preferable with a possible benefit of reduced morbidity by reduction in disease burden with chemotherapy (Kumar 2017).

There is consensus that the surgery performed during the initial treatment of ovarian cancer, whether PDS or IDS, should aim to leave NMRD, if possible. The need for clarity on the location (cancer centre or unit) and timing from diagnosis of first look surgery (intensive staging and cytoreductive surgery) for advanced ovarian cancer has never been more relevant. Women, clinicians and commissioners of specialist cancer services need to know what the overall benefit of cytoreductive surgery for ovarian cancer is, and to determine if there are subgroups of women for whom this intervention is of greater value. Given the diversity recognised within the overall group of women with advanced-stage ovarian cancer, it is anticipated that an ethos of individualised surgical planning, whilst recognising overarching principles, would be appropriate. One recent cohort study compared operative approaches/philosophies, where an ultra-radical approach to surgery was introduced at a population level (Falconer 2020). In this population-based cohort study, all women with suspected EOC in a region of Stockholm in two national cancer registries were selected in two three-year cohorts, based on year of diagnosis (before (cohort 1) or after (cohort 2) change in surgical treatment algorithm) and followed for at least three years. The study reported five-year overall survival in non-surgically and surgically treated women. A similar study into system reorganisation that uses either a controlled before-and-after component or interrupted time series design would be able to look at the impact of any centralisation of more radical surgery on survival.

Although the size of residual tumour mass after surgery has been shown to be an important prognostic factor for advanced ovarian cancer, there is limited evidence to support the conclusion that the surgical procedure is directly responsible for the superior outcome associated with less residual disease (Girling 1996; Hunter 1992).

Whether optimal cytoreduction is more feasible in women with biologically less aggressive tumours is a subject of continued debate. Tumour biology is not thought to be the only factor affecting prognosis (Sørensen 2019), and its impact seems to be partially overruled by the extent of residual disease, i.e. whether NMRD or SVRD was achieved (du Bois 2009). It has also been suggested that further evaluation of biological factors may help select women who are most likely to benefit from PDS (du Bois 2009; Markar 2016). It has been suggested that women whose cancer is cytoreduced to NMRD and SVRD at PDS may have super-imposable progression-free survival, meaning that women with high tumour load, completely resected at the time of surgery, may have micro/macroscopic unrecognised residual disease (Fagotti 2020). In this review, we will analyse PDS and IDS separately, as PDS achieving cytoreduction to < 1 cm may be equivalent to IDS achieving cytoreduction to NMRD.

The aim of this review is to investigate the effects of residual disease in women who received PDS or IDS for advanced epithelial ovarian cancer. This review should help to determine the prognostic impact of residual disease after surgery on survival.

## OBJECTIVES

To assess the prognostic impact of residual disease after primary surgery on survival outcomes or advanced (stage III and IV) epithelial ovarian cancer. In separate analyses, primary surgery included both upfront primary debulking surgery (PDS) followed by adjuvant chemotherapy and neoadjuvant chemotherapy followed by interval debulking surgery (IDS). Each residual disease threshold is considered as a separate prognostic factor.

### Investigation of sources of heterogeneity

We examined differences between FIGO stages III and IV in different thresholds of residual disease after primary surgery. We considered factors such as age, grade, length of follow-up, type and experience of surgeon, and type of surgery in the interpretation of any heterogeneity.

We also performed sensitivity analyses that distinguished between studies that included NMRD in residual disease (RD) categories of < 1 cm and those that did not. This was applicable to comparisons involving RD < 1 cm with the exception of RD < 1 cm versus NMRD.

We evaluated women undergoing PDS and IDS in separate analyses.

## METHODS

### Criteria for considering studies for this review

#### Types of studies

We included data from RCTs, prospective and retrospective cohort studies, and unselected case series of 100 or more women that included a concurrent comparison of different RD thresholds after primary surgical intervention. Any data collected from RCTs were retrospective and taken from trials that randomised groups of women to various chemotherapy protocols after primary or interval debulking surgery. We categorised the surgical outcome as macroscopic, optimal and suboptimal debulking, based on the maximum size of postoperative residual disease.

In order to minimise bias, we only included studies of multivariate Cox regression models that used sensible adjustment factors associated with survival in women with advanced EOC (e.g. age, stage, grade, extent of disease at diagnosis). We excluded studies that only reported unadjusted results. To assess the adequacy of adjustment factors used in multivariate Cox models, we used the 'adjustment for other prognostic factors' and 'statistical analysis and reporting' domains of the quality in prognosis studies (QUIPS) tool (Riley 2019). Therefore, in theory, only one other factor needed to be adjusted for the study to meet the criteria for inclusion in the review, but we judged such studies as being at high risk of bias in these domains.

We excluded case-control studies, studies that did not have concurrent comparison groups and case series of fewer than 100 women. This was to attempt to optimise the quality of the review, as poor study designs would have introduced additional forms of bias. The inclusion of adequately sized studies, although pragmatic, may

also provide more reliable estimates due to restricting results to those reporting multiple adjustments in statistical models.

### Types of participants

We included adult women (over 18 years of age) with surgically staged advanced epithelial ovarian cancer (FIGO stages III and IV) who had confirmed histological diagnoses. We excluded women with other concurrent malignancies.

Women were included into two distinct groups: those who received primary debulking surgery (PDS) followed by platinum-based chemotherapy and those who received interval debulking surgery (IDS), which involves receiving the surgery sandwiched between a schedule of chemotherapy. We analysed these distinct groups separately.

### Details of prognostic factor

The surgical intervention for which we assessed the resulting prognostic factor was primary debulking surgery (upfront and interval debulking).

We included studies that reported all RD thresholds after surgery but we defined optimal RD as surgery leading to residual tumours with a maximum diameter of any threshold up to 1 cm. The main RD thresholds of interest were microscopic RD (labelled as no macroscopic residual disease (NMRD)); RD < 1 cm and exclusive of 0 cm, categorised as small-volume residual disease (SVRD); and RD > 1 cm, categorised as large-volume residual disease (LVRD). However, we included studies reporting any size of RD but restricted to the most pertinent comparisons in key summary sections. We noted details of any women who had primary surgery that resulted in RD that did not meet the criteria specified in the study as 'optimal', namely not categorised as NMRD or SVRD cytoreduction.

We applied the above RD thresholds to both PDS (primary debulking surgery followed by platinum-based chemotherapy) and IDS (platinum-based chemotherapy followed by interval debulking surgery) settings.

- No macroscopic residual disease (NMRD) after PDS (RD = 0 cm).
- Small-volume residual disease (SVRD) after primary cytoreduction (RD 0.1 cm to 1 cm).
- Large-volume residual disease (LVRD) after cytoreduction (RD > 1 cm).

### Types of outcome measures

- Overall survival: survival until death from any cause. We assessed survival from the time at which women were enrolled in the study.
- Progression-free survival.

We extracted survival estimates as time-to-event data from an adjusted multivariate Cox model (as outlined above in 'Types of studies'). This is the most appropriate way to analyse these outcomes as it accounts for any loss to follow-up and will correctly allow for censoring.

## Search methods for identification of studies

We sought papers in all languages and translated them when necessary.

We searched the following electronic databases on 30 August 2021:

- the Cochrane Central Register of Controlled Trials (CENTRAL; 2021, Issue 8), in the Cochrane Library;
- MEDLINE via Ovid (1950 to 30 August 2021);
- Embase via Ovid (1950 to 2021 week 34).

The MEDLINE, EMBASE and CENTRAL search strategies were based on terms related to the review topic and are presented in Appendix 1, Appendix 2 and Appendix 3, respectively. We searched the databases from 1950 up to end of August 2021.

We identified all relevant articles found on PubMed and used the 'related articles' feature to carry out a further search for newly published articles.

### Searching other resources

#### *Unpublished and grey literature*

We searched metaRegister, Physicians Data Query, www.controlled-trials.com/rct, www.clinicaltrials.gov and www.cancer.gov/clinicaltrials for ongoing trials.

#### *Handsearching*

We checked the citation lists of relevant publications, abstracts of scientific meetings and included studies through handsearching, and we contacted experts in the field to identify further reports of studies. We handsearched reports of conferences from the following sources.

- *Gynecologic Oncology* (Annual Meeting of the American Society of Gynecologic Oncologists).
- *International Journal of Gynecological Cancer* (Annual Meeting of the International Gynecologic Cancer Society).
- *British Journal of Cancer*.
- British Cancer Research Meeting.
- Annual Meeting of European Society of Medical Oncology (ESMO).
- Annual Meeting of the American Society of Clinical Oncology (ASCO).

#### *Correspondence*

We contacted authors of relevant trials to ask if they knew of further data, which may or may not have been published.

## Data collection and analysis

### Selection of studies

We downloaded all titles and abstracts retrieved by electronic searching to the reference management database Endnote. After removing duplicates, three review authors (AB, PK, SH) examined the remaining references independently. We excluded those studies that clearly did not meet the inclusion criteria and obtained copies of the full text of potentially relevant references. Three review authors (AB, PK, SH) assessed the eligibility of retrieved papers independently. We resolved disagreements by discussion between the three review authors or, when necessary, by appeal to a fourth review author (RN, KG). We documented the reasons for exclusion.

### Data extraction and management

For included studies, we extracted items relevant to prognostic factor studies, derived from the checklist for critical appraisal and data extraction for systematic reviews of prediction modelling studies (CHARMS) (Moons 2014). This included data on the following:

- Author, year of publication and journal citation (including language).
- Country.
- Setting.
- Inclusion and exclusion criteria.
- Study design, methodology.
- Study population:
  - total number enrolled in each group;
  - participant characteristics;
  - age;
  - comorbidities.
- Ovarian cancer details at diagnosis:
  - FIGO stage (III or IV);
  - histological cell type;
  - preoperative tumour volume;
  - ascites (large or small volume);
  - tumour grade;
  - extent of disease.
- Surgical intervention details:
  - details of primary optimal cytoreductive surgery;
  - upfront and interval debulking settings.
- Details of platinum-based chemotherapy:
  - dose;
  - number of chemotherapy cycles before and after surgery;
  - type of surgeon (gynaecological oncologist, gynaecologist, general surgeon);
  - experience of surgeon;
  - type of surgery (ultra-radical or standard).
- Details of prognostic factor:
  - details of residual disease;
  - definition of residual disease thresholds in study;
  - covariates included in multivariate Cox models for survival that include residual disease.
- Risk of bias in study (see 'Assessment of risk of bias in included studies').
- Duration of follow-up.
- Outcomes (see 'Types of outcome measures').

For time-to-event data (survival and progression-free survival), we extracted the log of the hazard ratio (log(HR)) and its standard error from study reports. If the study did not report these, we did not attempt to estimate the log(HR) and its standard error using the methods of Parmar 1998, as we only included adjusted analyses.

We noted the time points at which outcomes were collected and reported.

Three review authors (AB, PK, SH) independently extracted data using a data collection form specially designed for the review. We resolved differences between review authors by discussion or by appeal to a fourth review author (KG), when necessary.

### Assessment of risk of bias in included studies

Three review authors independently extracted data and assessed risk of bias. We extracted the data using the CHARMS-PF (checklist for critical appraisal and data extraction for systematic reviews - prognostic factor studies; Riley 2019). We assessed the risk of bias for each outcome (overall survival and progression-free survival) in each study. We assessed risk of bias (and appraised quality) in the prognostic assessment of residual disease in the included studies using the quality in prognosis studies (QUIPS) tool (Appendix 4). QUIPS is a tool designed to assess risk of bias in prognostic factor studies (Riley 2019). It assesses bias across the following six domains using intermediate signalling questions to aid the decision-making process.

1. Participant selection
2. Study attrition
3. Prognostic factor measurement
4. Outcome measurement
5. Adjustment for other prognostic factors
6. Statistical analysis and reporting

In addition, we considered the applicability of the study for four of the domains, as reported in other tools (Whiting 2011; Wolff 2019). We judged risk of bias and concerns regarding applicability using the tools shown in Appendix 4. The questions regarding applicability included the following.

- Domain 1: participant selection. Are there concerns that the included women do not match the review question?
- Domain 3: prognostic factor measurement. Are there concerns that residual disease, the way that it is measured, or the way that it is interpreted, differ from the review question?
- Domain 4: outcome measurement. Are there concerns that the outcome does not match the review question or that follow-up was not of sufficient duration?
- Domain 5: adjustment for other prognostic factors. Did the prognostic factors adjusted for match the review question?

Three review authors (AB, PK, SH) applied the risk of bias tool independently and resolved differences by discussion or by appeal to a fourth review author (KG). We presented the results in a risk of bias summary table. We interpreted the results of meta-analyses in light of the findings with respect to risk of bias.

### Measures of effect

For time-to-event data (overall and progression-free survival), we used the adjusted hazard ratio (HR). We did not use unadjusted results, as outlined above in 'Types of studies'.

### Dealing with missing data

We did not impute missing outcome data for any of the outcomes.

### Assessment of heterogeneity

We assessed heterogeneity between studies by visual inspection of forest plots, by estimation of the percentage of heterogeneity between trials that cannot be ascribed to sampling variation (Higgins 2003), by a formal statistical test of the significance

of the heterogeneity (Deeks 2001), and, where possible, by subgroup analyses (see 'Subgroup analysis and investigation of heterogeneity'). If there was evidence of substantial heterogeneity, we investigated and reported the possible reasons for this.

### Assessment of reporting biases

We examined the symmetry of funnel plots corresponding to meta-analyses of overall survival to assess the potential for small study effects in analyses containing 10 or more studies. We tested for asymmetry where evidence of asymmetry may have been an indicator of publication bias (Debray 2018; Sterne 2011).

### Data synthesis

If sufficient clinically similar studies were available, we pooled their adjusted results in meta-analyses. We reported results by FIGO stage (see 'Subgroup analysis and investigation of heterogeneity').

- For time-to-event data, we pooled hazard ratios (HRs) using the generic inverse variance facility of Review Manager 2020.
- We used random-effects models with inverse variance weighting for all meta-analyses (DerSimonian 1986).
- We reported analyses separately for women who received upfront and interval debulking surgery.

### Subgroup analysis and investigation of heterogeneity

We considered factors such as age, grade, length of follow-up, type and experience of surgeon, and type of surgery in the interpretation of any heterogeneity.

We performed subgroup analyses grouping studies by women with FIGO stage III versus stage IV disease.

We analysed women undergoing PDS and IDS in separate analyses (see above).

### Sensitivity analysis

We had planned to perform sensitivity analysis that restricted the analyses to studies we judged to be at an overall low risk of bias. However, the overall profiles of the included studies were largely very similar.

We performed a sensitivity analysis that distinguished between studies that included NMRD in residual disease categories of < 1 cm and those that did not. This was applicable to some comparisons involving RD < 1 cm, with the exception of SVRD versus NMRD. In this area, RD <1 cm should be exclusive of NMRD and is often described as RD = 0.1 cm to 1 cm in the literature, for clarity.

We also conducted a number of post hoc sensitivity analyses. This included excluding one study (Klar 2016), which included a proportion of women with early and unknown stage disease.

### Summary of findings and assessment of the certainty of the evidence

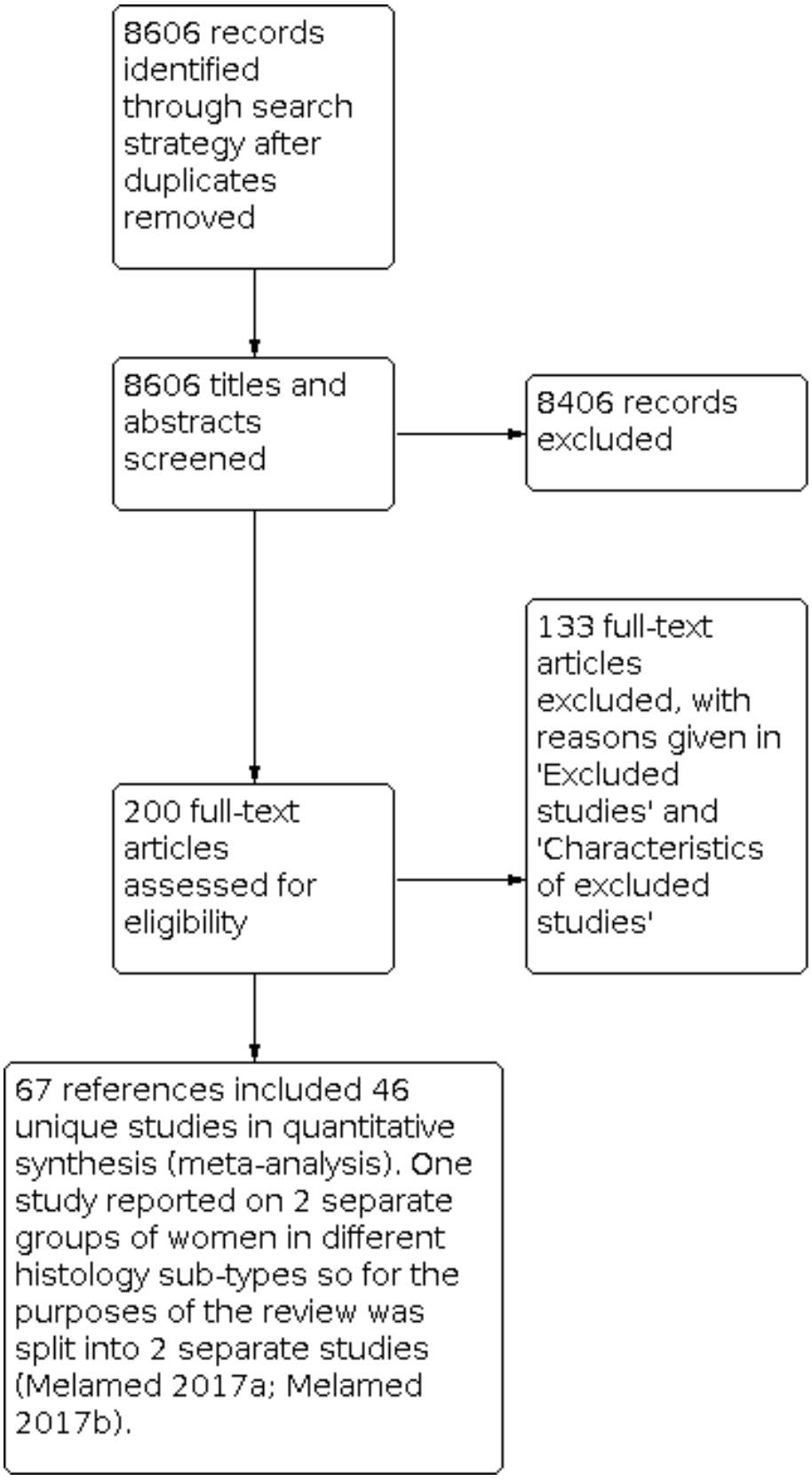
Guidance on the use of GRADE for prognostic factor studies has not yet been published (Foroutan 2020; GRADE Working Group), but we attempted to appraise the quality and certainty of the evidence where possible. We constructed summary of findings tables to present the results of outcomes in the review for the main comparisons involving prognostic factor thresholds of NMRD, SVRD (0.1 cm to 1cm) and LVRD. We used the GRADE system to rank the certainty of the evidence (Foroutan 2020; GRADE Working Group). Two review authors (AB, SH) independently graded the evidence and resolved differences by discussion or by involving a third review author (PK). We based our judgements on the strength of the body of evidence based on the domains presented in Appendix 5. Where the evidence was based on single studies, or where there was no evidence on a specific outcome for comparisons, we included the outcome in the summary of findings table and graded or explained in a narrative account accordingly. We gave the rationale for each judgement in the table footnotes. We interpreted the results of the review in light of this graded evidence. Summary of findings tables are given for PDS studies in Summary of findings 1, Summary of findings 2 and Summary of findings 3 and in IDS studies in Summary of findings 4, Summary of findings 5 and Summary of findings 6. The comparison involving any remaining macroscopic disease (RD > 0 cm) and NMRD in an IDS setting was also an important comparison so we additionally gave this a certainty of evidence judgement (Summary of findings 7).

## RESULTS

### Results of the search

The search strategy identified 8606 unique references (Figure 1). The title and abstract screening of these references identified 200 studies as potentially eligible for the review. The full-text screening of the 200 references identified 13 references, reporting on two RCTs (Kehoe 2015; Vergote 2010), but these trials did not meet the inclusion criteria as they did not report results across residual disease thresholds; instead they gave comparisons of residual disease by type of surgery. These trials were reported in a recent Cochrane Review (Coleridge 2021), which assessed chemotherapy versus surgery for initial treatment in advanced ovarian epithelial cancer along with another three trials (Chekman 2015; Fagotti 2020; Onda 2020), which did not report any of their outcomes for extent of disease by type of initial surgery.

**Figure 1. Study flow diagram.**



We excluded 133 references reporting on 115 studies that investigated the effects of residual disease after primary surgery for the reasons described in the table Characteristics of excluded studies. The remaining 67 references, reporting on 46 unique studies, met our inclusion criteria and are described in the table Characteristics of included studies. Fifty-two of these, reporting on 30 unique studies, reported on residual disease for PDS. One included publication, Klar 2016, reported results based on four individual RCTs but each one alone did not meet the inclusion criteria due to different scope so we included the combined analysis reported in Klar 2016. One study reported on two separate groups of women in different histology sub-types so for the purposes of the review we split it into two separate studies (Melamed 2017a; Melamed 2017b), therefore we refer to 31 included studies throughout. The other 15 studies reported on residual disease for IDS.

Searches of the grey literature did not identify any additional relevant trials.

There were three RCTs evaluating the effectiveness of surgery in advanced-stage epithelial ovarian cancer (Redman 1986; Rose 2004; Van Der Burg 1996). However, we excluded all three of these trials as they were designed to evaluate the benefits of surgery after an induction period with chemotherapy treatment, where the surgery was performed as a secondary procedure after initial (primary) surgery and they have been evaluated in a separate Cochrane Review (Tangjitgamol 2016).

## Characteristics of included studies

See Characteristics of included studies table.

### Residual disease after upfront primary debulking surgery (PDS)

The 31 included studies assessed a total of 22,376 women (Akahira 2001; Aletti 2006; Ataseven 2016; Bristow 2011; Chan 2003; Chang 2012a; Chang 2012b; Chi 2001; Chi 2006; Cuylan 2018; Eisenkop 2003; Feng 2016; Hofstetter 2013; Kahl 2017; Klar 2016; Langstraat 2011; Luger 2020; McGuire 1995; Melamed 2017a; Melamed 2017b; Paik 2018; Peiretti 2010; Peiretti 2012; Polterauer 2012; Shim 2016; Tewari 2016; Tseng 2018; Van Geene 1996; Wimberger 2010; Winter 2007; Winter 2008). Three studies included a small proportion of women with early-stage (predominantly stage II) or unknown disease. Although not stringently part of our initial inclusion criteria, we included a study if the proportion with unknown or early-stage disease in the entire cohort was small. The proportion of women with early or unknown stage of disease in Feng 2016 (9.3%), Polterauer 2012 (6.6%) and Klar 2016 (12.5%) was not going to affect the applicability of the results. The analyses in Klar 2016 included 1182 women with stage IIB to IIIB disease and 3684 had stage IIIC to IV disease. The study contributed heavily to the analyses, but the results were robust to its exclusion in a sensitivity analysis. The four individual RCTs used in the analyses could not be included separately because residual disease (RD) was not reported.

Four studies reported exclusively on women with stage IV epithelial ovarian cancer (EOC) and included 225, 326, 573 and 360 stage IV women respectively (Akahira 2001; Ataseven 2016; Wimberger 2010; Winter 2008).

Five studies reported exclusively on women with stage IIIC EOC (Aletti 2006; Bristow 2011; Chang 2012b; Chi 2006; Eisenkop 2003); whereas Cuylan 2018 and Winter 2007 reported women with stage IIIA to C disease; whilst 16 studies reported on both stage III and IV EOC (Chan 2003; Chang 2012a; Chi 2001; Hofstetter 2013; Langstraat 2011; McGuire 1995; Melamed 2017a; Melamed 2017b; Paik 2018; Peiretti 2010; Peiretti 2012; Polterauer 2012; Shim 2016; Tewari 2016; Tseng 2018; Van Geene 1996).

The number of women included in all studies varied from 104 in the Chan 2003 study to 5055 women in the Klar 2016 analysis. The larger studies tended to combine results from primary studies but generally it was not possible to report the results of these separately due to the scope of the original publications that had a different focus.

For a summary of the total number of women included in each study, as well as stage and residual disease details see Table 1.

#### *Design*

All analyses examining residual disease thresholds following surgery were retrospective in nature.

Four studies were primarily prospective cohort studies (Eisenkop 2003; Hofstetter 2013; Polterauer 2012; Van Geene 1996).

The Winter 2007, Winter 2008 and Klar 2016 studies were retrospective analyses of six, four and four randomised controlled trials of various chemotherapy protocols, respectively. The Winter 2007 study reported on women with stage III EOC, Winter 2008 reported on women with stage IV EOC and Klar 2016 a mix of stages included a small proportion of early and unknown. Winter 2007 included women from GOG protocols 111, 114, 132, 152, 158 and 172 (Armstrong 2006; Markman 2001; McGuire 1996; Muggia 2000; Ozols 2003; Rose 2004), Winter 2008 included women from GOG protocols 111, 132, 152 and 162 (McGuire 1996; Muggia 2000; Rose 2004; Spriggs 2007) and Klar 2016 reported a combined analysis of four individual RCTs (OVAR 3, 5, 7 and 9). Likewise, the McGuire 1995 study was a retrospective analysis of a randomised controlled trial of two different chemotherapy protocols.

All remaining studies were analyses of retrospective data from hospital databases, medical records and cancer registries.

#### *Participant characteristics*

Fourteen studies were conducted in the USA (Aletti 2006; Bristow 2011; Chan 2003; Chi 2001; Chi 2006; Eisenkop 2003; Langstraat 2011; Melamed 2017a; Melamed 2017b; McGuire 1995; Tewari 2016; Tseng 2018; Winter 2007; Winter 2008), whilst four were set in South Korea (Chang 2012a; Chang 2012b; Paik 2018; Shim 2016), nine set predominantly in Europe including Germany, Belgium, France, Spain, Italy, Austria and the UK (Ataseven 2016; Hofstetter 2013; Kahl 2017; Klar 2016; Luger 2020; Peiretti 2010; Polterauer 2012; Van Geene 1996; Wimberger 2010); the study Cuylan 2018 was set in Turkey, Feng 2016 in China and the Akahira 2001 study was conducted in 24 centres in Japan. One of the studies included populations from multiple locations: Peiretti 2012 (Italy and the USA).

The mean or median age reported for women with advanced EOC varied between 50.9 years (Tewari 2016) to 73.5 (Langstraat 2011) years with the range between 16 to 91 years.

***Details of PDS reported in studies***

RD thresholds ranged from NMRD up to > 5 cm across the included studies. The most common comparisons were of RD thresholds NMRD, SVRD (described in most studies as being < 1 cm, but exclusive of NMRD) and LVRD. We did identify studies where optimal RD was defined up to < 2 cm, but more recent studies and guidelines (BGCS 2017; du Bois 2009) state that surgery should not be considered optimal beyond 1 cm (however, we assessed RD as a prognostic factor and we included studies that included all RD thresholds, but only reported the most pertinent comparisons in the key sections of the review).

Women in all the studies described above underwent PDS followed by platinum-based adjuvant chemotherapy. All women were confirmed histologically to have invasive epithelial ovarian cancer.

The speciality of the surgeon who performed PDS (for example, general surgeon, gynaecologic surgeon or specialist gynaecologic oncology surgeon) was not reported in 20 of the included studies (Akahira 2001; Aletti 2006; Chang 2012a; Feng 2016; Hofstetter 2013; Klar 2016; Langstraat 2011; McGuire 1995; Melamed 2017a; Melamed 2017b; Paik 2018; Peiretti 2010; Polterauer 2012; Shim 2016; Tewari 2016; Tseng 2018; Van Geene 1996; Wimberger 2010; Winter 2007; Winter 2008); whereas specialist gynaecologic oncology surgeons undertook PDS in 11 studies (Ataseven 2016; Bristow 2011; Chan 2003; Chang 2012b; Chi 2001; Chi 2006; Cuylan 2018; Eisenkop 2003; Kahl 2017; Luger 2020; Peiretti 2012).

The mean duration of PDS was reported to be 210 minutes (range: 40 to 480 minutes) in Aletti 2006. Similarly the median duration of PDS was reported to be 194 minutes (range: 60 to 750 minutes) and 180 minutes (range: 55 to 480 minutes) in the Chi 2006 and Eisenkop 2003 studies respectively. All three studies reported on women with stage IIIC disease. On the other hand, the Akahira 2001 study reported on women with stage IV disease and the median duration of PDS was found to be 240 minutes (range 40 to 780 minutes). Two studies reported on the mean duration of PDS on women with stage III and IV disease: 270 minutes (range: 70 to 480 minutes) in Peiretti 2010 and 280 minutes (range: 36 to 893 minutes) in Tseng 2018.

The duration of PDS was not reported in the remaining 25 studies (Ataseven 2016; Bristow 2011; Chan 2003; Chang 2012a; Chang 2012b; Chi 2001; Cuylan 2018; Feng 2016; Hofstetter 2013; Kahl 2017; Klar 2016; Langstraat 2011; Luger 2020; McGuire 1995; Melamed 2017a; Melamed 2017b; Paik 2018; Peiretti 2012; Polterauer 2012; Shim 2016; Tewari 2016; Van Geene 1996; Wimberger 2010; Winter 2007; Winter 2008).

The median estimated operative blood loss was 500 mL (range 20 mL to 7500 mL); 850 mL (range 30 mL to 5000 mL) and 1085 mL (range 40 mL to 11,000 mL) in the Chi 2006, Eisenkop 2003 and Akahira 2001 studies, respectively. In the latter study, blood transfusion was given to 112 women (50%) intra- and postoperatively. Peiretti 2010 and Peiretti 2012 reported the estimated blood loss using different measures as 700 mL (range 50 mL to 6000 mL) and 1000 mL (range 200 mL to 8500 mL), respectively. Intraoperative blood transfusion was given to 112 (43.2%) and 152 (64%) women in Peiretti 2010 and Peiretti 2012 respectively, while postoperative blood transfusion was given to 140 (50.1%) women in Peiretti 2010 and 150 (63%) women in Peiretti 2012. The Hofstetter 2013 study did not report on the estimated blood loss, however they reported that nine of 185 women (4.86%) required blood transfusion.

Only five studies reported on the length of hospital stay (LHS). In the studies by Chi 2006, Eisenkop 2003 and Peiretti 2012 the median LHS was 10 days, with a range of 0 to 59, 0 to 93 and 4 to 24 days, respectively. The median LHS was 9 days and 8 days (range: 1 to 22 days) in Peiretti 2010 and Tseng 2018, respectively.

Postoperative mortality within 30 days of PDS ranged from 0.4% to 4.3% in eight studies reporting this outcome (Ataseven 2016; Aletti 2006; Bristow 2011; Chi 2001; Chi 2006; Eisenkop 2003; Langstraat 2011; Tseng 2018). One study reported a postoperative mortality rate of 45% but this was during a median follow-up period of 49.6 months (interquartile range (IQR) 32.9 to 66.3) (Luger 2020).

Postoperative mortality and morbidity were not reported in 19 studies (Akahira 2001; Chan 2003; Chang 2012a; Chang 2012b; Feng 2016; Hofstetter 2013; Klar 2016; Melamed 2017a; Melamed 2017b; McGuire 1995; Paik 2018; Peiretti 2010; Peiretti 2012; Polterauer 2012; Shim 2016; Van Geene 1996; Wimberger 2010; Winter 2007; Winter 2008).

Two studies used a postoperative residual disease cutoff of < 2 cm to define an optimal level of remaining RD after surgery (Akahira 2001; Van Geene 1996). Eighteen studies considered that an optimal outcome was achieved only if NMRD was left behind at the conclusion of PDS (Ataseven 2016; Chang 2012a; Chang 2012b; Cuylan 2018; Eisenkop 2003; Feng 2016; Hofstetter 2013; Kahl 2017; Langstraat 2011; Luger 2020; Melamed 2017a; Melamed 2017b; Paik 2018; Peiretti 2010; Peiretti 2012; Tewari 2016; Tseng 2018; Wimberger 2010). Four studies used a postoperative RD cutoff of < 1 cm to define the optimal level of remaining RD (Aletti 2006; Bristow 2011; Chan 2003; Klar 2016). The remaining seven studies did not define what is considered optimal in the study methodology but analysed the outcome by a range of postoperative RD (Chi 2001; Chi 2006; McGuire 1995; Polterauer 2012; Shim 2016; Winter 2007; Winter 2008).

Four studies did not make direct comparisons against NMRD. These studies included NMRD in the RD < 1 cm (Chi 2001; Chan 2003) and RD < 2 cm categories (Akahira 2001; McGuire 1995). None of the studies reported the proportion of participants with NMRD. While Winter 2008 did give a breakdown of various RD categories, the authors additionally reported a comparison involving RD > 1 cm versus < 1 cm with the latter including NMRD (n = 29/107).

The rate of NMRD after surgery was reported in 20 studies (Aletti 2006; Ataseven 2016; Chang 2012b; Chi 2006; Cuylan 2018; Eisenkop 2003; Kahl 2017; Langstraat 2011; Luger 2020; Melamed 2017a; Melamed 2017b; Paik 2018; Peiretti 2010; Peiretti 2012; Polterauer 2012; Tewari 2016; Tseng 2018; Winter 2007; Winter 2008). It was achieved in 4906 out of 15,246 women (32.2%) with the lowest macroscopic disease rate reported by Tewari 2016 (4.9%) and the highest (86%) reported by Eisenkop 2003.

Postoperative RD < 1 cm (SVRD) was achieved in 8201 out of 19,185 women (42.75%) as calculated from 19 studies (Aletti 2006; Ataseven 2016; Bristow 2011; Chan 2003; Chang 2012a; Chi 2001; Chi 2006; Cuylan 2018; Eisenkop 2003; Klar 2016; Langstraat 2011;

Melamed 2017a; Melamed 2017b; Paik 2018; Polterauer 2012; Tewari 2016; Wimberger 2010; Winter 2007; Winter 2008). The lowest rate for RD < 1 cm was 25.3% (71/281) in the Chi 2001 study and the highest was 96% (392/408) in the Eisenkop 2003 study.

In 26 studies all women received postoperative platinum-based chemotherapy (Aletti 2006; Ataseven 2016; Bristow 2011; Chan 2003; Chang 2012a; Chang 2012b; Cuylan 2018; Eisenkop 2003; Feng 2016; Hofstetter 2013; Kahl 2017; Klar 2016; Langstraat 2011; Luger 2020; McGuire 1995; Melamed 2017a; Melamed 2017b; Paik 2018; Peiretti 2010; Peiretti 2012; Polterauer 2012; Tewari 2016; Van Geene 1996; Wimberger 2010; Winter 2007; Winter 2008). In four studies the majority of women (95.1%, 96%, 97%, 98.4%, 99% respectively) received postoperative platinum-based chemotherapy (Akahira 2001; Chi 2001; Chi 2006; Tseng 2018). The main reason for not receiving postoperative chemotherapy was postoperative death within 30 days of surgery and absent records (Chi 2001). Other reasons for not receiving postoperative chemotherapy or receiving non-platinum-based chemotherapy were poorly reported. The study by Shim 2016 did not report the number of women who received postoperative chemotherapy.

Fourteen studies reported the survival outcome for NMRD (Aletti 2006; Ataseven 2016; Bristow 2011; Chi 2006; Cuylan 2018; Eisenkop 2003; Feng 2016; Hofstetter 2013; Kahl 2017; Langstraat 2011; Paik 2018; Tewari 2016; Winter 2007; Winter 2008).

#### *Outcomes*

The median duration of follow-up varied from 28 months (Winter 2008) to 77.7 months (Tseng 2018), with a range between 1 and 199 months (Chi 2006). The duration of follow-up was not reported in seven studies (Chang 2012b; McGuire 1995; Peiretti 2012; Shim 2016; Tewari 2016; Van Geene 1996; Wimberger 2010).

Only two studies did not report overall survival (Peiretti 2010; Shim 2016), while 16 studies reported progression-free survival and used appropriate statistical techniques (hazard ratios to correctly allow for censoring) (Chang 2012a; Chang 2012b; Cuylan 2018; Feng 2016; Klar 2016; Luger 2020; McGuire 1995; Paik 2018; Peiretti 2010; Polterauer 2012; Shim 2016; Tewari 2016; Tseng 2018; Wimberger 2010; Winter 2007; Winter 2008). Prognostic factors were adjusted for in the analysis of survival outcomes in each study using Cox regression. Between them, the 30 studies (31 with Melamed split (Melamed 2017a; Melamed 2017b)) included 29 different prognostic factors in the analysis. The number of prognostic factors included in the analysis ranged from two in Eisenkop 2003 to 10 in Tewari 2016. The prognostic factors most frequently included in the analyses are (in order of frequency) residual disease (26 studies), age (23 studies), stage (21 studies), performance status (nine studies), histology (nine studies) and tumour grade (six studies). A list of the different prognostic factors is shown in Appendix 6.

For the distribution of these factors at baseline for each study and by residual disease, see the table Characteristics of included studies.

### Residual disease after interval debulking surgery (IDS)

The 15 included studies assessed a total of 3697 women (Bixel 2020; Cioffi 2018; Davidson 2019; Iwase 2015; Kaban 2017; Lecointre 2020; Lecuru 2019; Liu 2020; Lorusso 2016; Petrillo 2014; Phillips 2018; Shibutani 2020; Stoeckle 2014; Zhang 2018; Zhu 2016). One study, whilst it reported descriptive statistics for 102 women, only had 85 women who underwent interval debulking surgery (IDS) (Cioffi 2018). Although this was not strictly part of our inclusion criteria (i.e. n ≥ 100), we noted this study as a caveat. Additionally, adjusted hazard ratios (HRs) and 95% confidence intervals (CIs) were not reported in Petrillo 2014 and Lecuru 2019 in their multivariate Cox models; however, P values were reported in both. Two of the included studies were abstracts only (Lecuru 2019; Lorusso 2016).

All studies included women with advanced EOC who underwent IDS (neoadjuvant chemotherapy (NACT) given prior to surgery). Twelve of the studies provided descriptive statistics of FIGO stage - all of which included samples of women with FIGO stages III and IV (Bixel 2020; Cioffi 2018; Davidson 2019; Iwase 2015; Lecointre 2020; Liu 2020; Petrillo 2014; Phillips 2018; Shibutani 2020; Stoeckle 2014; Zhang 2018; Zhu 2016). For the three remaining studies, only Kaban 2017 and Lecuru 2019 reported in their methods that women with stage IIIC and IV ovarian cancer were included; we could not determine FIGO staging for Lorusso 2016.

Study sample size varied from 102 (Cioffi 2018) to 672 (Zhu 2016).

For a summary of the total number of women included in each study, as well as stage and residual disease details see Table 2.

#### *Design*

All analyses examining RD thresholds were retrospective in nature with data collected from past medical records and databases. The exceptions were Lecuru 2019, which was a secondary analysis of the CHIVA double-blind randomised phase II GINECO study that sought to examine the effects of nintedanib in combination with NACT (Ferron 2019); Davidson 2019, whose sample comprised data collected retrospectively from medical records as well as prospective participants(the purpose of the prospective data collection being to explore the role of minimally invasive surgery following NACT); and Lecointre 2020, whose sample was from a multicentre cohort study of women with histologically confirmed advanced epithelial ovarian cancer who all consented to participation.

#### *Participant characteristics*

Three of the studies were conducted in Italy (Cioffi 2018; Lorusso 2016; Petrillo 2014), three in France (Lecointre 2020; Lecuru 2019; Stoeckle 2014), three in China (Liu 2020; Zhang 2018; Zhu 2016), two in the USA (Bixel 2020; Davidson 2019), two in Japan (Iwase 2015; Shibutani 2020), and one study was conducted in Turkey (Kaban 2017), and the UK (Phillips 2018) each. Five of the studies were conducted across multiple centres: Lecointre 2020 collected data from nine French referral centres, Davidson 2019 from three US institutions, Bixel 2020 from two US institutions, Lorusso 2016 from five Italian centres, and Zhu 2016 from two Chinese institutions.

The median age reported for women with advanced EOC varied between 55 years (Zhu 2016) and 64 years (Stoeckle 2014) with the range between 28 and 88 years.

#### *Details of interval debulking surgery reported in studies*

RD thresholds ranged from NMRD up to > 2 cm across the included studies. The most common comparisons were of RD thresholds NMRD, ≤ 1 cm (although the majority included NMRD in this threshold, rather than 0.1 cm to 1 cm, which we defined as SVRD), and > 1 cm (LVRD). Optimal RD was commonly defined as less than 1 cm (RD < 1) or less than or equal to 1 cm (≤ 1 cm), consistent

with recent studies and guidelines (BGCS 2017; du Bois 2009), which state that surgery should not be considered optimal beyond 1 cm. Four studies did not provide an explicit definition of optimal RD (Lecointre 2020; Lecuru 2019; Lorusso 2016; Petrillo 2014). This was due to the nature of the information for the middle two cases (i.e. abstracts). For Petrillo 2014, although no definition of optimal RD was given, thresholds of NMRD, RD ≤ 1 cm and RD > 1 cm were provided. For Lecointre 2020, thresholds of NMRD, RD ≤ 0.25 cm, and RD 0.25 cm to 2.5cm were used. Davidson 2019 utilised two definitions of optimal RD (NMRD and RD ≤ 1 cm) in their study although only the latter was used in their multivariate Cox model.

Six studies compared SVRD versus LVRD (Cioffi 2018; Davidson 2019; Kaban 2017; Phillips 2018; Zhang 2018; Zhu 2016). Six of the studies did not make direct comparisons against NMRD and included NMRD in their SVRD category (Cioffi 2018; Davidson 2019; Kaban 2017; Shibutani 2020; Zhang 2018; Zhu 2016). Consequently, comparisons of SVRD (0.1 cm to 1 cm) and LVRD (> 1 cm) suffered from serious bias as a result of the inclusion of NMRD in the near-optimal category. Of these six studies, only three reported the number of participants with NMRD within the SVRD category: Cioffi 2018 (n = 37/57 participants with SVRD), Davidson 2019 (n = 165/228) and Zhang 2018 (n = 59/156). Only one study appropriately treated NMRD as a distinct category from SVRD (Phillips 2018).

Women in all the studies were treated by platinum-based neoadjuvant chemotherapy followed by IDS. One possible exception may be Lorusso 2016, but it was assumed that the NACT was platinum-based. All women were confirmed histologically to have invasive EOC.

The median number of NACT cycles varied from three (Zhang 2018) to six (Iwase 2015), with a range of 1 to 13. The large range is partially contributed by Stoeckle 2014, which was conducted in women receiving delayed IDS (after six or more cycles). Two studies did not provide descriptive statistics of NACT cycles (Lecuru 2019; Zhu 2016), but reported in their methodology that women received three or between three to four cycles. Information on the NACT regimen was provided in all but one study (Lorusso 2016). Carboplatin plus paclitaxel was most commonly reported and varied between 37.2 % (Zhu 2016), 96.6% (Stoeckle 2014), and 100%; although no details were reported for Kaban 2017, Lecuru 2019 (with reference to Ferron 2019), and Zhang 2018, they reported all women received carboplatin plus paclitaxel in their methods. Route of administration was reported in Bixel 2020 in which NACT was administrated intraperitoneally in 28% and intravenously in 72%, and Zhang 2018 in which NACT was administrated intraperitoneally in 45% and intravenously in 55%. Response to NACT according to RECIST criteria was reported in three studies in which complete/partial response was observed in 66.6% (Cioffi 2018), 66.1% (Zhu 2016), and in all participants in Lecointre 2020 (however, this was based on n = 380/501 with data on NACT response).

Information on the specialty of the surgeon performing the IDS was only reported in Stoeckle 2014 where all 118 surgeries were conducted by two surgeons with experience in ovarian cancer surgery and Shibutani 2020 where gynaecologic oncologists were involved in all surgeries. Duration of IDS was only reported in two studies and varied from a median of 194 minutes (Davidson 2019) to 419 minutes (Iwase 2015), with a range of 45 to 611 minutes. Length of hospital stay (LHS) was only reported in Stoeckle 2014 with a median of 10 days (range: 2 to 44) and Lecointre 2020 (median of 10 days (range: 6 to 13) in the group with ≤ 4 NACT cycles and median of 11 days (range: 7 to 14) in the group with > 4 NACT cycles). Postoperative morbidity/complications and mortality (defined as death within 30 days of IDS) was only reported in two studies (Davidson 2019; Stoeckle 2014). Postoperative mortality varied from 0% to 1.7%, whilst postoperative morbidity/complications varied from 18% to 22% in these studies. Complications after discharge and within 30 days of surgery were reported only in Davidson 2019. Approximately 11% experienced post-discharge complications of whom 6.4% were re-admitted. Operative blood loss was reported in Iwase 2015, with a median blood loss of 1291 mL (range: 220 mL to 5640 mL) and Lecointre 2020, where 57% of patients required blood transfusion (based on n = 77/501 with available data). Lecointre 2020 reported intraoperative complications in 15% of patients (based on n = 387/501 with available data). Lecointre 2020 also reported postoperative complications in 22% of participants (based on n = 421/501 patients with available data) but this was across an undefined time frame.

Information on postoperative chemotherapy following IDS was reported in 11 studies, albeit with varying levels of detail (Bixel 2020; Cioffi 2018; Iwase 2015; Kaban 2017; Lecuru 2019; Liu 2020; Petrillo 2014; Phillips 2018; Shibutani 2020; Stoeckle 2014; Zhang 2018). Clear reporting of platinum-based adjuvant chemotherapy was observed in five studies (Bixel 2020; Iwase 2015; Lecuru 2019; Petrillo 2014 (with reference to Ferron 2019); Zhang 2018), whilst it was implied (Kaban 2017; Liu 2020; Phillips 2018; Shibutani 2020; Stoeckle 2014) or unstated (Cioffi 2018) in the remaining six studies. Six of the studies did not provide descriptive statistics for adjuvant chemotherapy cycles or regimen and only reported in their methods that participants received chemotherapy following IDS (Cioffi 2018; Kaban 2017; Lecuru 2019 (with reference to Ferron 2019); Petrillo 2014; Shibutani 2020; Stoeckle 2014). However, with the exception of Cioffi 2018, they did report in their methods that their participants received two (Petrillo 2014; Liu 2020; Stoeckle 2014), two to three (Lecuru 2019 (with reference to Ferron 2019)), or two to six (Kaban 2017) cycles of adjuvant chemotherapy. Shibutani 2020 did not report the number of adjuvant cycles but did report the total (NACT + adjuvant chemotherapy) cycles. Six studies reported descriptive statistics (Bixel 2020; Iwase 2015; Liu 2020; Phillips 2018; Shibutani 2020; Zhang 2018). The median number of cycles ranged from three (Iwase 2015; Phillips 2018) to five (Zhang 2018), and ranged from one to eight in these three studies.

Optimal RD was most commonly defined as RD < 1 cm (Cioffi 2018; Iwase 2015; Phillips 2018; Shibutani 2020; Stoeckle 2014; Zhang 2018) or RD ≤ 1 cm (Bixel 2020; Davidson 2019; Kaban 2017; Liu 2020; Zhu 2016). Four studies did not provide a definition of optimal RD in their methodology but included RD thresholds in their multivariable Cox models (Lecointre 2020; Lecuru 2019; Lorusso 2016; Petrillo 2014). Davidson 2019 utilised two definitions of optimal RD (NMRD and SVRD) in their study, although only the latter was used in their multivariate Cox model. NMRD was reported in 12 studies (Bixel 2020; Cioffi 2018; Davidson 2019; Iwase 2015; Lecointre 2020; Lecuru 2019; Liu 2020; Lorusso 2016; Petrillo 2014; Phillips 2018; Stoeckle 2014; Zhang 2018), however descriptive statistics for the rate of NMRD were only reported in 10 studies (Bixel 2020; Cioffi 2018; Davidson 2019; Iwase 2015; Lecointre 2020; Liu 2020; Petrillo 2014; Phillips 2018; Stoeckle 2014; Zhang 2018). Lecointre 2020 reported missing data for RD in n = 30/501 women and did not report any imputation method. Rate of

NMRD varied from the lowest of 29.5% (Zhang 2018) to the highest of 79% (Iwase 2015). Across the 10 studies that reported descriptive statistics, NMRD was achieved in 1451 out of 2237 women (64.9%).

Across the six studies that provided descriptive statistics for RD < 1 cm (Cioffi 2018; Iwase 2015; Phillips 2018; Shibutani 2020; Stoeckle 2014; Zhang 2018), RD < 1 cm was achieved in 897 out of 1096 women (81.8%). Rates per study varied from 71% (Cioffi 2018) to 94% (Stoeckle 2014).

Across the four studies that provided descriptive statistics for RD ≤ 1 cm (Davidson 2019; Kaban 2017; Petrillo 2014; Zhu 2016), RD ≤ 1 cm was achieved in 1151 out of 1466 women (78.5%). Rates per study varied from 72% (Zhu 2016) to 84% (Davidson 2019; Petrillo 2014).

Nine studies reported the survival outcome in models comparing RD threshold(s) against NMRD (Bixel 2020; Iwase 2015; Lecointre 2020; Lecuru 2019; Liu 2020; Lorusso 2016; Petrillo 2014; Phillips 2018; Stoeckle 2014).

***Outcomes***

The median duration of follow-up was reported in nine studies (Bixel 2020; Iwase 2015; Kaban 2017; Lecuru 2019; Petrillo 2014; Shibutani 2020; Stoeckle 2014; Zhang 2018; Zhu 2016), and varied from a median of 29.5 months (Bixel 2020) to 47 months (Petrillo 2014), with a range between 1 and 181 months. The duration of follow-up was not reported in four studies (Cioffi 2018; Davidson 2019; Lecointre 2020; Liu 2020; Lorusso 2016; Phillips 2018).

Only one study did not report overall survival (Davidson 2019). Three studies did not provide adjusted HRs and 95% confidence intervals from their multivariate survival models predicting overall survival (Bixel 2020; Lecuru 2019; Petrillo 2014). One study only brought RD forward into the "multivariate" model for overall survival after univariate analysis, however the criteria for selection was not mentioned in the methods (Liu 2020). Eight studies reported progression-free survival and used appropriate statistical techniques (hazard ratios to correctly allow for censoring) (Cioffi 2018; Lecointre 2020; Lecuru 2019; Liu 2020; Petrillo 2014; Zhang 2018; Zhu 2016). One study reported using multivariate logistic regression to predict progression-free survival in their methods but reported hazard ratios in their results, so it may be inferred that multivariate Cox regression had actually been used (Bixel 2020). Disease-specific overall survival (DSS) was reported in Davidson 2019. Disease-free survival (DFS) was reported in Liu 2020. Prognostic factors were adjusted for in the analysis of survival outcomes in each study using Cox regression. Between them, the 15 studies included 29 different prognostic factors in the analysis. The precise prognostic factors used in Lorusso 2016 could not be determined beyond the complete cytoreduction, ECOG performance status and number of NACT cycles. The number of prognostic factors included in the analysis ranged from one in Petrillo 2014 to nine in Cioffi 2018. The prognostic factors most frequently included in the analyses are (in order of frequency): residual disease (15 studies), number of NACT cycles (eight studies), age (seven studies), FIGO stage (seven studies), performance status (six studies), ascites (four studies), response to NACT (four studies), NACT regimen (three studies), CA-125 (two studies) and lymphadenectomy (two studies). A list of the different prognostic factors is shown in Appendix 7.

One study, which included 501 women, had missing RD data for 30 (6%) (Lecointre 2020). Furthermore, other variables in the multivariate Cox model for overall survival had larger rates of missing data such as the Charlson Index (missing data for n = 203, 41%) and response to NACT (missing for n = 121, 24%). It is likely that the multivariate Cox model was based on a complete case analysis and therefore the estimates reported are based on ≤ 298 women, but the exact number cannot be known. For the multivariate model for progression-free survival, the estimates are based on ≤ 380 women as response to NACT was included as a covariate.

For the distribution of these factors at baseline for each study and by RD threshold see the table Characteristics of included studies.

## Excluded studies

We excluded 133 references reporting on 115 studies after obtaining the full text, for the following primary reasons.

- We excluded 42 references reporting on 40 studies because they did not include at least 100 women with advanced epithelial ovarian cancer (Alphs 2006; Andersen Soegaard 2005; Benedetti-Panici 1996; Bristow 1999; Cai 2007; Ceresoli 2018; Colozza 1997; Del Campo 1994; Gao 2001; Gershenson 1989; Gershenson 1995; Grem 1991; Hainsworth 1990; Hakes 1992; Hamid 2002; Hardy 1991; Hoskins 1996; Kaern 2005; Kirmani 1994; Kristensen 1995; Loizzi 2016; Lorusso 1998; Malik 1998; Marchetti 1993; Ngan 1989; Palmer 1992; Risum 2012; Redman 1986; Rutten 2014; Shapiro 1998; Son 2017; Strauss 1996; Sutton 1989; Tay 1996; Taylor 1994; Vallejos 1997; Willemse 1992; Wils 1990; Zang 1999; Zhang 2015).
- Twenty-two studies either did not report multivariate analyses or did not include or adequately report residual disease as a variable to enable an analysis (Alberts 1996; Altman 2012; Bertelsen 1990; Bian 2016; Brinkhuis 1996a; Clamp 2018; Greggi 2016; Heitz 2016; Kessous 2017; Keyver-Paik 2016; Lee 2018; McGuire 1996; Piver 1991; Raspagliesi 2018; Rodriguez 2013; Sessa 1991; Sioulas 2017; Stewart 2016; Suidan 2015; Vidal 2016; Wallace 2017; Wimberger 2007).
- Fourteen studies did not report survival by residual disease (Alberts 1993; Bertelsen 1993; Brinkhuis 1996b; Conte 1991; Conte 1996; Creasman 1990; Gershenson 1992; Hoskins 1992; Hoskins 1997; Itamochi 2002; Solmaz 2015; Uyar 2005; Wadler 1996; Warwick 1995).
- Non-platinum based chemotherapy was given to all women in one study (Van Driel 2017), a proportion of women in four studies (Barda 2004; Bonnefoi 1999; de Oliviera 1990; Tingulstad 2003), and chemotherapy data were absent in the Bailey 2006 study. Women received preoperative chemotherapy in two studies (Shinozuka 1999; Sun 2000).
- Four studies included women who received neoadjuvant chemotherapy and interval debulking surgery but did not report an appropriate comparison by extent of disease (Dao 2016; Todo 2003; Van Der Burg 1996; van Vliet 2015).
- Seven studies included women with early-stage disease and it was not possible to distinguish between early- and advanced-stage participants (Crawford 2005; di Re 1996; Geisler 2004; Skarlos 1996; Smits 2015; Takano 2006; Takano 2007). The Le 1997 study did not report the survival data from the stage IIIC and IV subgroup and the authors no longer had access to these data.

- Two studies reported a HR for overall survival but did not include the corresponding 95% confidence interval, standard error (SE) (lnHR) or exact P value (Baker 1994; Omura 1989).
- The study Rose 2004 reported on outcomes after secondary debulking surgery. However, the trial statistician (Dr Mark Brady) of the included study Winter 2007 alerted us to the results of GOG 152, which reported by residual disease after primary cytoreductive surgery.
- Salani 2007 was excluded because it was a case-control study.
- The Yamamoto 2007 study included 67 selected women with rare histological subtypes and the Gasimli 2016 study included a selective group of women with cytoreduction of tumour to macroscopic optimal disease (0 cm).
- The Anuradha 2016 study focused only on the time interval between surgery and chemotherapy and the Michaan 2018 study focused on chemotherapy response score as an outcome, which is a histopathological scoring system based on morphological features of cancer tissue removed at IDS, but the same as optimal cytoreduction.
- Six references reporting on three RCTs comparing upfront versus delayed surgery did not report outcomes for extent of residual disease by type of initial primary surgery (Chekman 2015; Fagotti 2020; Onda 2020).
- Sixteen references reporting on three studies compared the threshold of residual disease based on type of intervention delivered (Kehoe 2015; Vergote 2010; Vergote 2018).
- Four studies were excluded because there was inadequate reporting and/or the full text was not available (Cummins 2019; Elgamal 2019; Stewart 2015; Trhlík 2013).
- One study did not distinguish between upfront and interval debulking primary surgery (Ruscito 2016).

For further details of all the excluded studies see the Characteristics of excluded studies table.

## Risk of bias and quality appraisal in included studies

We assessed the risk of bias at outcome level for overall survival and progression-free survival for each study using the QUIPS tool (Riley 2019). Most studies reported overall survival (only two of all PDS studies (Peiretti 2010; Shim 2016), and just one study of all IDS studies (Davidson 2019) did not report overall survival). The detailed assessments are depicted in the 'Risk of bias (QUIPS)' section in the Characteristics of included studies.

We judged most studies included in the review as being at an overall 'moderate' risk of bias as they satisfied some but not all of the domains using the QUIPS tool. (See Table 3; Table 4; Table 5; Table 6 for risk of bias assessment using the QUIPS tool for overall survival and progression-free survival in the PDS and IDS studies).

### Study participation

Most studies provided adequate details of study participation, which included details of eligible women, descriptions of the population and of the baseline study sample and recruitment, period and place of recruitment, and a description of inclusion and exclusion criteria. We assessed four studies as 'unclear' for this domain (two PDS studies (Hofstetter 2013; Van Geene 1996) and two IDS studies (Iwase 2015; Kaban 2017)), mostly due to a lack of detailed reporting of inclusion criteria. We assessed three studies (one PDS study (Shim 2016) and two IDS studies (Lecuru 2019; Lorusso 2016)) as being at a high risk of bias because they were in abstract form only, providing insufficient information on study participation.

### Applicability: Are there concerns that the included women do not match the review question?

All studies matched the review question and there were no applicability concerns. Many studies reported one particular stage of advanced disease, but we were not concerned about this as we performed subgroup analyses by stage.

Ten PDS studies appeared to include a strictly representative sample of women with advanced epithelial ovarian cancer, by including stages III and IV combined (Chan 2003; Chang 2012a; Chi 2001; Hofstetter 2013; McGuire 1995; Peiretti 2010; Peiretti 2012; Shim 2016; Tewari 2016; Van Geene 1996). The Polterauer 2012 study included a small proportion of women with stage II disease (6.6%) and Feng 2016 included 9.3% early stage (I to II) disease, however both were otherwise representative of advanced disease. Klar 2016 included a small proportion of women with early-stage (IA to IIA) disease (3.6%) and an unknown proportion with stage IIB but the main scope was advanced disease so this was likely to be relatively few. The results of the meta-analyses were robust to the exclusion of this study in sensitivity analyses, so we did not deem the decision to include Klar 2016 in the review as being associated with any bias or issues with representativeness of women.

Of the 15 IDS included studies, four included a strictly representative sample of participants with advanced ovarian cancer (Iwase 2015; Petrillo 2014; Phillips 2018; Zhu 2016).

### Study attrition

It was unclear if women with incomplete follow-up were excluded before arriving at the stated sample size in each study. There was insufficient information to permit judgement in all cases as many studies did not examine RD as a prognostic factor as their primary objective.

### Prognostic factor measurement

Most studies reported a valid and reliable measurement of RD and we assessed these as being at a low risk of bias for the prognostic factor measurement domain. Even though multicentre studies are advantageous in terms of recruitment options and generalisability of participants as well as other positive features, we cautiously assessed the prognostic factor measurement to be unclear in 12 studies (eight PDS studies (Akahira 2001; Chan 2003; Cuylan 2018; Kahl 2017; Klar 2016; Peiretti 2012; Polterauer 2012; Van Geene 1996) and four IDS studies (Bixel 2020; Davidson 2019; Lecointre 2020; Zhu 2016)) that had this design, but these may well have been at a low risk too.

### Applicability: Are there concerns that residual disease, the way that it is measured, or the way that it is interpreted, differ from the review question?

RD is measured by the surgeons estimate in all centres and there are no guidelines on how RD should be objectively measured. Therefore, there will be some natural variability in measurement across different centres, but we did not have any concerns about applicability.

**Outcome measurement**

The majority of the studies reported a valid and reliable measurement of outcome for both overall survival and progression-free survival and we assessed these as being at low risk of bias for the outcome measurement domain.

***Overall survival***

Two studies reported an inappropriate definition of overall survival (one PDS study (Aletti 2006) and a IDS study (Davidson 2019)) by reporting disease-specific survival, rather than all-cause overall survival. Consequently, we assessed these two studies to be at a high risk of bias. Outcome measurement of overall survival was unclear in one PDS study (Van Geene 1996) (Table 3; Table 4).

***Progression-free survival***

All studies that reported progression-free survival will have done so based on imaging and tumour markers. However, this is a somewhat subjective outcome and in unblinded studies could be deemed as being at a greater risk of bias. Therefore we judged the outcome measurement domain to be at unclear risk of bias as the measurement of this outcome may or may not have been reliable in certain RD thresholds (Table 5; Table 6).

**Applicability: Are there concerns that outcome does not match the review question or that follow-up was not of sufficient duration?**

We had no applicability concerns for outcome measurement for overall survival and progression-free survival.

**Adjustment for other prognostic factors**

For this domain, we assessed the appropriateness of confounders and whether important ones that a study should have at least been adjusted for such as age were included in their prognostic models. In cases where other prognostic factors in models were inadequate, we rated the studies as having a high risk of bias.

***Overall survival***

The studies at high risk of bias included seven PDS studies (Akahira 2001; Bristow 2011; Eisenkop 2003; Melamed 2017a; Melamed 2017b; Peiretti 2012; Shim 2016) and nine IDS studies (Bixel 2020; Davidson 2019; Lecointre 2020; Lecuru 2019; Liu 2020; Lorusso 2016; Petrillo 2014; Phillips 2018; Zhu 2016). These studies did not adequately adjust for a sufficient number of other prognostic factors in multivariate models or ones included were not pertinent. Adequate adjustment for other prognostic factors was unclear in 12 PDS studies (Chang 2012b; Feng 2016; Hofstetter 2013; Kahl 2017; Klar 2016; Langstraat 2011; McGuire 1995; Paik 2018; Van Geene 1996; Wimberger 2010; Winter 2007; Winter 2008) and in three IDS studies (Kaban 2017; Stoeckle 2014; Zhang 2018) (Table 3; Table 4).

***Progression-free survival***

The studies at high risk of bias included two PDS studies (Peiretti 2010; Shim 2016) and six IDS studies (Bixel 2020; Lecointre 2020; Lecuru 2019; Liu 2020; Petrillo 2014; Zhu 2016). These studies did not adequately adjust for a sufficient number of other prognostic factors in multivariate models or ones included were not pertinent. Adequate adjustment for other prognostic factors was unclear in eight PDS studies (Chang 2012b; Feng 2016; Klar 2016; McGuire 1995; Paik 2018; Wimberger 2010; Winter 2007; Winter 2008) and in one IDS study (Zhang 2018) (Table 5; Table 6).

**Applicability: Did the prognostic factors adjusted for match the review question?**

There was no reason to doubt the applicability of prognostic factors that were adjusted for in the multivariable models. Some studies may have used a wider range and more pertinent prognostic factors in their models for both overall survival and progression-free survival, but all studies satisfied our inclusion criteria for appropriateness of prognostic factors in their prognostic models and we had no applicability concerns.

Adjusted hazard ratios for survival using multivariable Cox models were used in each study. Any imbalances at baseline between RD thresholds should therefore be accounted for and all adjustments in the included studies met the inclusion criteria for the review.

We had applicability concerns in one IDS study (Petrillo 2014), as the multivariable analyses for overall survival and progression-free survival only adjusted for pathological response to NACT, so there may still be differences between RD thresholds that have not been controlled for.

**Statistical analysis and reporting**

We assessed the statistical analysis and reporting domain as being at high or unclear risk of bias in all included studies for both overall survival and progression-free survival outcomes. Either no conceptual framework was reported, where the variable selection criteria in the multivariate model was unclear or quite often the authors reported that significant variables from the univariate analysis were included in the multivariable model, but with no further details. It is also questionable whether this is adequate.

Mainly applicable to an IDS setting, it was not possible to distinguish NMRD within the SVRD thresholds in all but one study reporting a comparison of NMRD and SVRD. Only one study separated NMRD from SVRD (RD = 0.1 cm to 1 cm) and all other studies included NMRD in the SVRD group, resulting in serious risk of bias. Inclusion of NMRD in the SVRD category creates a high risk of bias when comparing suboptimal RD.

## Findings

Meta-analyses of survival are based on hazard ratios (HRs) that were adjusted for prognostic variables (see Appendix 6 (PDS) and Appendix 7 (IDS) for details).

The percentage of the variability in effect estimates that was due to heterogeneity rather than sampling error (chance) may appear to represent substantial or considerable heterogeneity (as measured by the $I^2$ statistic) in some of the analyses below, but we had no major concerns as the direction of effect was consistent throughout.

We have reported the most pertinent comparisons involving SVRD (0.1 cm to 1 cm) versus NMRD, LVRD (> 1 cm) versus NMRD, and LVRD versus SVRD for overall survival and progression-free survival; these all provided moderate-certainty evidence. These are the most pertinent comparisons as they are included in clinical guidelines (NICE 2013), and are the focus of a key area of clinical uncertainty. Other RD comparisons were prespecified and have been provided.

The certainty of the evidence assessed using the GRADE approach (GRADE Working Group) was moderate for all comparisons involving overall survival and progression-free survival in a PDS setting and very low in an IDS setting. We restricted to comparisons of the three main reported RD thresholds (NMRD, SVRD and LVRD), since there is no firm guidance for grading the evidence in reviews of prognostic factor analyses (Riley 2019). Therefore, we did not grade beyond these key RD thresholds (Summary of findings 1; Summary of findings 2; Summary of findings 3; Summary of findings 4; Summary of findings 5; Summary of findings 6). The comparison involving any remaining macroscopic disease (RD > 0 cm) and NMRD in an IDS setting was also an important comparison, so this was included in the summary of findings and GRADE assessment (Summary of findings 7).

**Residual disease after upfront primary debulking (cytoreductive) surgery (PDS)**

Where possible the meta-analyses subgrouped studies by FIGO stage (stage III, IIIC, IV and all advanced stages, if studies included all advanced cases together). We conducted subgroup analyses to explore the underlying clinical heterogeneity between the studies. There was no evidence of subgroup differences in any of the subgroup analyses. The results of these subgroup analyses were robust to the findings of the overall pooled estimate for all comparisons, so the results of each subgroup are not discussed in this section (see Analysis 1.1 to Analysis 11.2).

The SVRD threshold included NMRD in some studies in comparison with LVRD, but only in a small number of studies. In PDS studies, RD < 1 cm means RD 0.1 cm to 1 cm (SVRD), unless otherwise stated. Due to only being an issue in a small number of studies, it was deemed to have a negligible impact on the results and did not affect the risk of bias profiles, the certainty of the evidence or distort the results. We performed sensitivity analyses when necessary.

We performed sensitivity analyses in comparisons that included meta-analysis of more than 10 studies. The use of a fixed-effect model aided the construction of the pseudo 95% confidence interval lines on the funnel plot (e.g. expected distribution of studies in the absence of heterogeneity and biases (such as publication bias, data irregularities)), as well as allowing us to see how robust the random-effects model results were in comparison. To further test the robustness of the findings, we additionally conducted a sensitivity analysis excluding studies with the largest weight in the meta-analyses comparing main RD thresholds, where appropriate.

We were cautious about any over-interpretation of funnel plots as they are typically underpowered. Given the nature of model selection procedures, we did not dismiss the possibility of publication bias. However, it is unclear as to the direction of any bias as, for instance, many highly significant studies only reporting unadjusted analyses found strong evidence that NMRD was associated with prolonged survival compared to other thresholds including SVRD (RD < 1 cm exclusive of 0 cm).

**Overall survival (risk of death from all causes)**

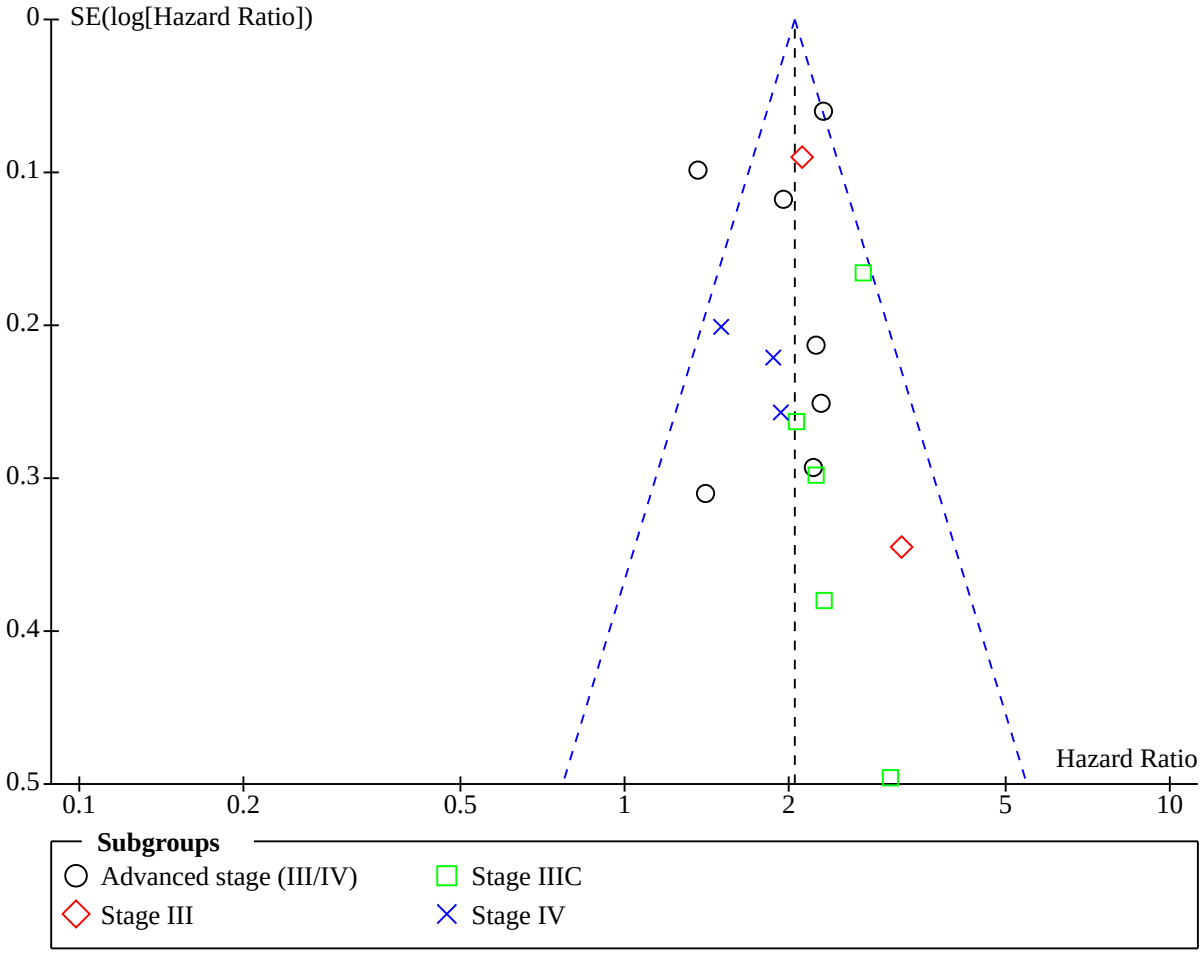
***Small-volume residual disease (SVRD) versus no macroscopic residual disease (NMRD)***

Meta-analysis of 17 studies, assessing 9404 participants, found that women with SVRD after PDS had more than twice the risk of death compared to women with NMRD (hazard ratio (HR) 2.03, 95% confidence interval (CI) 1.80 to 2.29). The percentage of the variability in effect estimates that was due to heterogeneity rather than chance may represent moderate heterogeneity ($I^2$ = 50%) (Analysis 1.1) (Summary of findings 1) (Aletti 2006; Ataseven 2016; Bristow 2011; Chang 2012a; Chang 2012b; Chi 2006; Cuylan 2018; Eisenkop 2003; Kahl 2017; Klar 2016; Langstraat 2011; Paik 2018; Tewari 2016; Tseng 2018; Wimberger 2010; Winter 2007; Winter 2008).

The results were robust to a sensitivity analysis that used a fixed-effect model and one that excluded the Klar 2016 study, which included a slight proportion of women with early or unknown stage (12.5%) disease. It also contributed the largest weight in the meta-analysis (see Analysis 1.2; Analysis 1.3).

There did not appear to be any evidence of small study biases, such as publication bias, or any irregularities with the data by visual inspection of a funnel plot (Figure 2).

**Figure 2. Funnel plot of comparison: 1 SVRD (< 1 cm) versus NMRD, outcome: 1.2 Overall survival**



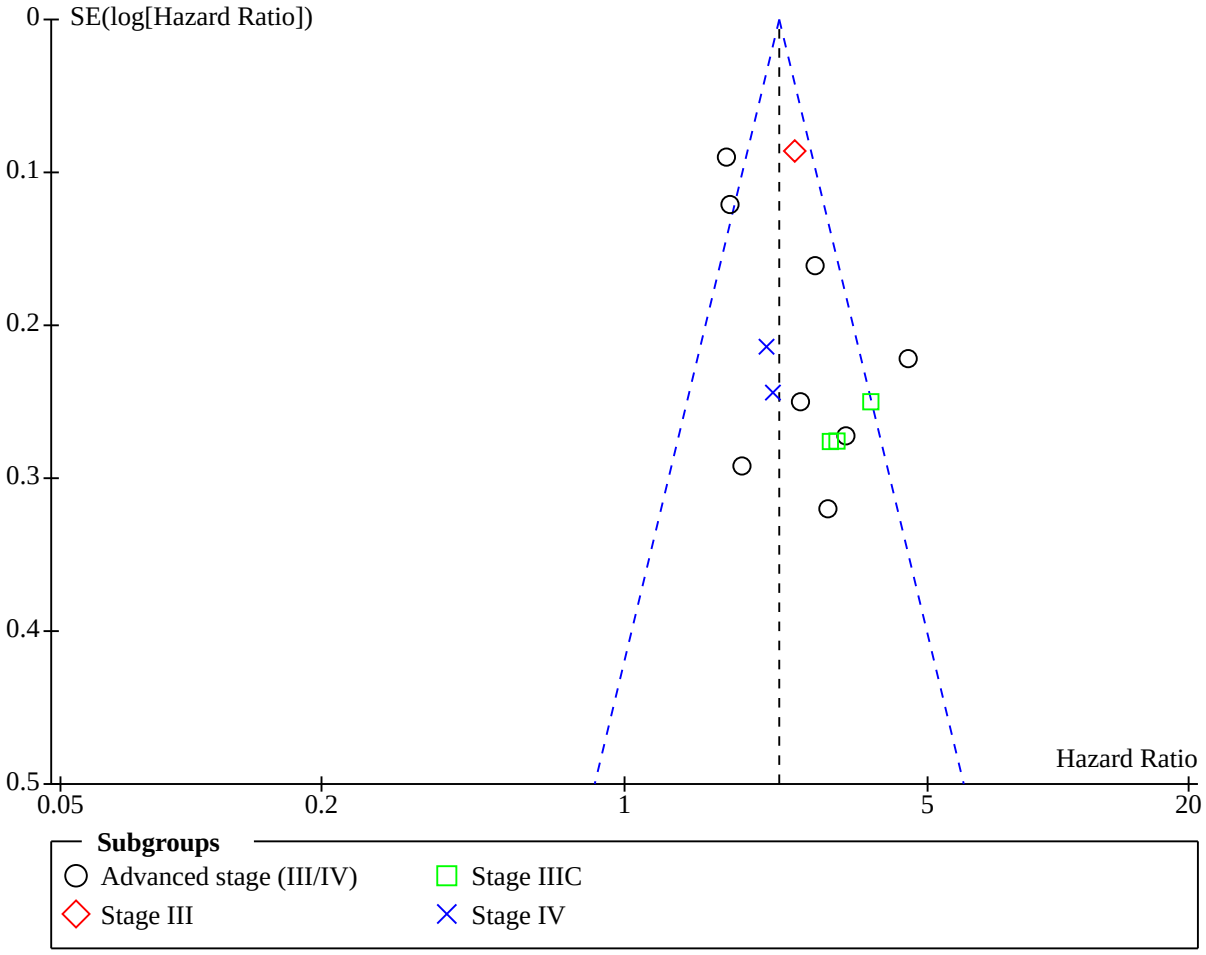

***Large-volume residual disease (LVRD) (> 1 cm) versus NMRD***

Meta-analysis of 14 studies, assessing 7988 participants, found that women with LVRD after PDS were associated with two and a half times the risk of death compared to women with NMRD (HR 2.50, 95% CI 2.13 to 2.94). The percentage of the variability in effect estimates that was due to heterogeneity rather than chance may represent substantial heterogeneity ($I^2$ = 63%) (Analysis 2.1) (Summary of findings 2) (Ataseven 2016; Chang 2012a; Chang 2012b; Chi 2006; Eisenkop 2003; Kahl 2017; Langstraat 2011; Melamed 2017a; Melamed 2017b; Paik 2018; Tewari 2016; Tseng 2018; Wimberger 2010; Winter 2007).

The results were robust to a sensitivity analysis that used a fixed-effect model and one that excluded the two studies with the largest weights in the meta-analysis (Melamed 2017b; Winter 2007) (see Analysis 2.2; Analysis 2.3).

There did not appear to be any evidence of small study biases, such as publication bias, or any irregularities with the data by visual inspection of a funnel plot (Figure 3).

**Figure 3. Funnel plot of comparison: 4 LVRD (> 1 cm) versus NMRD, outcome: 2.2 Overall survival**




***LVRD versus SVRD***

Meta-analysis of five studies, assessing 6000 participants, found that women with LVRD after PDS was associated with a greater risk of death compared to women with SVRD < 1 cm (HR 1.22, 95% CI 1.13 to 1.32; 6000 participants). The percentage of the variability in effect estimates that was due to heterogeneity rather than chance is not important ($I^2$ = 0%) (Analysis 3.1) (Summary of findings 3) (Chan 2003; Klar 2016; Melamed 2017a; Melamed 2017b; Winter 2008). The results were robust to a sensitivity analysis that excluded the Klar 2016 study with the largest weight in the meta-analysis (and a relatively small proportion of women with early or unknown stage (12.5%) disease) (see Analysis 3.2).

The results were also robust when only including the three studies that contributed majority of the weight in the meta-analysis and did not include NMRD in the SVRD category (HR 1.20, 95% CI 1.10 to 1.30; 5594 participants; $I^2$ = 0%) (Klar 2016; Melamed 2017a; Melamed 2017b)(see Analysis 3.3).

Similarly, meta-analysis of two studies that included NMRD in the SVRD category arrived at the same conclusion (HR 1.37, 95% CI 1.09 to 1.72; 435 participants; $I^2$ = 0%) (Chan 2003; Winter 2008).

Only Winter 2008 reported the proportion of women with NMRD (n = 29/107 of participants in the SVRD category)(see Analysis 3.4).

***Residual disease (RD) > 0 cm versus NMRD***

Meta-analysis of four studies, assessing 1220 participants, found that women who had RD greater than 0 cm after PDS were associated with a two-fold increase in the risk of death compared to women with NMRD (HR 1.96, 95% CI 1.44 to 2.67). The percentage of the variability in effect estimates that was due to heterogeneity rather than chance may represent moderate heterogeneity ($I^2$ = 49%) (Analysis 4.1) (Feng 2016; Hofstetter 2013; Luger 2020; Polterauer 2012). The authors of Peiretti 2012 additionally found that the risk of death for women with any remaining RD after PDS was higher than for those with NMRD (238 participants; P = 0.003), but the magnitude of effect was not reported.

***RD 1 cm to 2 cm versus NMRD***

The Aletti 2006 study, which included only women with stage IIIC disease, found that women who had RD between 1 cm and 2 cm after PDS were associated with a nearly four-fold increase in the risk of death compared to women with NMRD (HR 3.95, 95% CI 1.33 to 11.78; 68 participants) (Analysis 5.1).

***RD > 2 cm versus NMRD***

The Aletti 2006 study, which included only women with stage IIIC disease, found that women with LVRD > 2 cm after PDS were associated with more than eight times the risk of death compared to women with NMRD (HR 8.24, 95% CI 2.68 to 25.33; 87 participants) (Analysis 6.1).

***RD 1 cm to 5 cm versus NMRD***

The Winter 2008 study, which included only women with stage IV disease, found that women who had LVRD between 1 cm and 5 cm after PDS were associated with a greater risk of death compared to women with NMRD (HR 1.83, 95% CI 1.14 to 2.94; 193 participants) (Analysis 7.1).

***RD > 5 cm versus NMRD***

The Winter 2008 study, which included only women with stage IV disease, found that women who had LVRD > 5 cm after PDS were associated with more than two and a half times the risk of death compared to women with NMRD (HR 2.72, 95% CI 1.65 to 4.47; 118 participants) (Analysis 8.1).

***RD 1 cm to 2 cm versus SVRD***

The Chi 2001 study found that women who had LVRD between 1 cm and 2 cm after PDS were associated with a greater risk of death compared to women with SVRD (HR 1.70, 95% CI 1.10 to 2.60; 144 participants) (Analysis 9.1). The SVRD category in the Chi 2001 study included NMRD.

***RD > 2 cm versus SVRD***

The Chi 2001 study found that women with LVRD > 2 cm after PDS were associated with twice the risk of death compared to women with SVRD (HR 2.00, 95% CI 1.34 to 2.99; 208 participants) (Analysis 10.1). The SVRD category in the Chi 2001 study included NMRD.

***RD > 2 cm versus RD < 2 cm***

Meta-analysis of two studies, which included only women with stage IV disease and assessed 478 participants, found no statistically significant difference in the risk of death between women with LVRD > 2 cm after PDS and those with RD < 2 cm (HR 1.63, 95% CI 0.83 to 3.23). The percentage of the variability in effect estimates that was due to heterogeneity rather than chance alone may represent considerable heterogeneity ($I^2$ = 89%) (Akahira 2001; Winter 2008). The two studies were inconsistent: the Akahira 2001 study reported a large survival difference in favour of RD < 2 cm, whereas Winter 2008 found no difference in overall survival (Analysis 11.1). The < 2 cm category included NMRD in both studies, so this category had a mix of NMRD and SVRD < 1 cm as well as LVRD between 1 cm to 2 cm.

The authors of Van Geene 1996 reported the same comparison, but found evidence that more RD is associated with increased risk of death (HR 1.83, 95% CI not reported; 219 participants; P < 0.0001). Similarly, in two publications by McGuire 1995 in the same cohort of women, survival was significantly worse in women with LVRD > 2 cm compared to less remaining RD (1 cm to 2 cm as no women had SVRD) after PDS (n = 294 women with stage III disease, P < 0.01). The authors note that there was little notable difference in the risk of death between any volume of RD in comparisons of LVRD > 2 cm up to > 10 cm. In a further analysis including all advanced stages of disease (n = 458), women with stage III disease and LVRD > 2 cm had a lower risk of death than either those with stage III disease and LVRD > 2 cm, or those with stage IV disease (P = 0.012).

**Progression-free survival (risk of disease progression)**

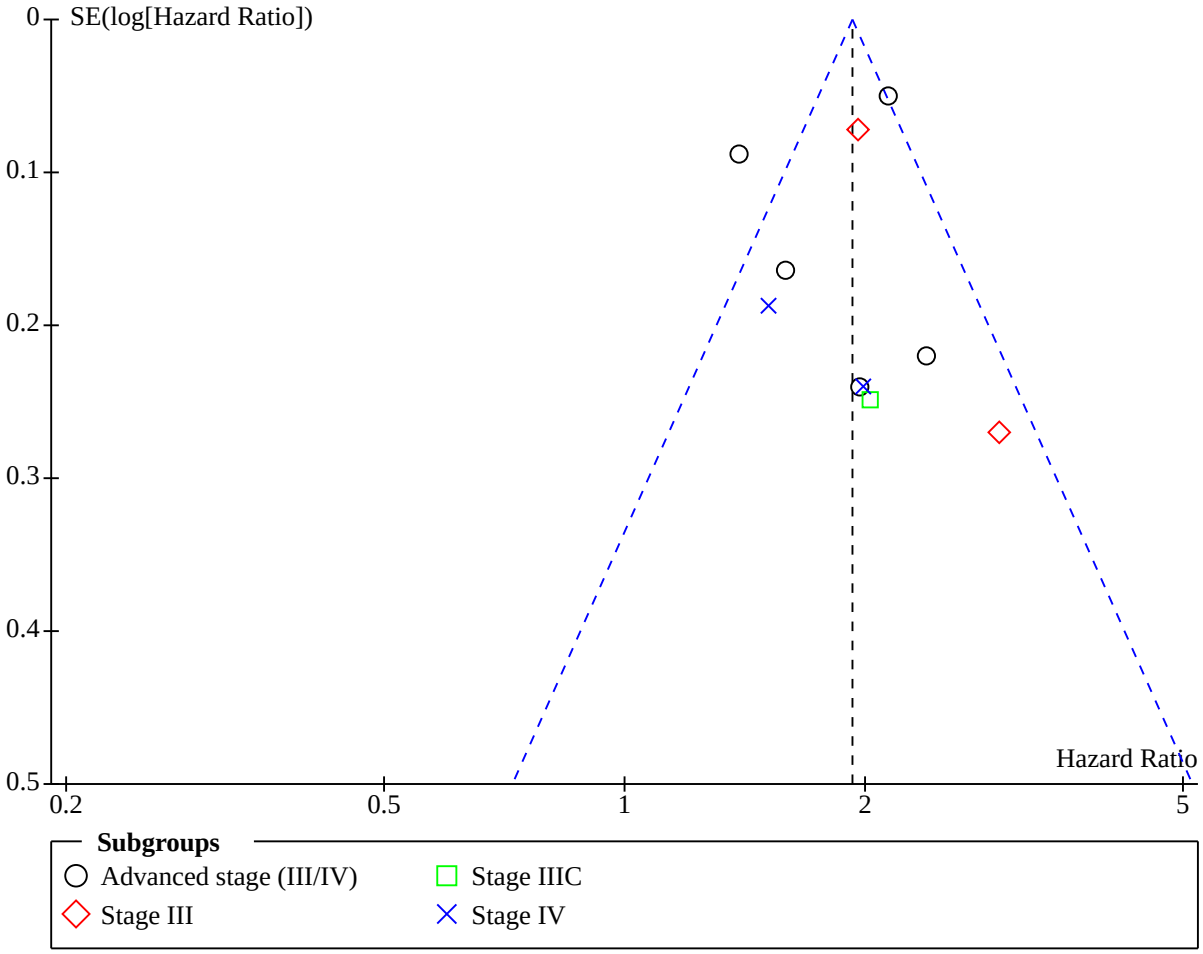
***SVRD versus NMRD***

Meta-analysis of 10 studies, assessing 6596 participants, found that women with SVRD after PDS were associated with nearly twice the risk of disease progression compared to women with NMRD (HR 1.88, 95% CI 1.63 to 2.16). The percentage of the variability in effect estimates that was due to heterogeneity rather than chance alone may represent substantial heterogeneity ($I^2$ = 63%) (Analysis 1.4) (Summary of findings 1) (Chang 2012a; Chang 2012b; Cuylan 2018; Klar 2016; Paik 2018; Shim 2016; Tseng 2018; Wimberger 2010; Winter 2007; Winter 2008).

The results were robust to a sensitivity analysis that used a fixed-effect model and one that excluded the Klar 2016 study with the largest weight in the meta-analysis (and a slight proportion of women with early or unknown stage (12.5%) disease) (see Analysis 1.5; Analysis 1.6).

There did not appear to be any evidence of small study biases, such as publication bias, or any irregularities with the data by visual inspection of a funnel plot (Figure 4).

**Figure 4. Funnel plot of comparison: 1 SVRD (< 1 cm) versus NMRD, outcome: 1.5 Progression-free survival**




***LVRD (> 1 cm) versus NMRD***

Meta-analysis of six studies, assessing 2629 participants, found that women with LVRD after PDS had more than twice the risk of disease progression compared to women with NMRD (HR 2.10, 95% CI 1.84 to 2.40). The percentage of the variability in effect estimates that was due to heterogeneity rather than chance may not be important ($I^2$ = 24%) (Analysis 2.4) (Summary of findings 2) (Chang 2012a; Chang 2012b; Paik 2018; Tseng 2018; Wimberger 2010; Winter 2007).

***LVRD versus SVRD***

Meta-analysis of two studies, assessing 3402 participants, found that women with LVRD > 1 cm after PDS had a greater risk of disease progression compared to women with SVRD (HR 1.30, 95% CI 1.08 to 1.56). The percentage of the variability in effect estimates that was due to heterogeneity rather than chance alone may represent moderate heterogeneity ($I^2$ = 53%) (Analysis 3.5) (Summary of findings 3) (Klar 2016; Winter 2008). Winter 2008 included NMRD in the SVRD category, but this only represented a small proportion in the analysis (n = 29/107 of participants in the SVRD category).

***RD > 0 cm versus NMRD***

Meta-analysis of three studies, assessing 1029 participants, found that women who had RD greater than 0 cm after PDS had more than one and a half times the risk of death compared to women with NMRD (HR 1.60, 95% CI 1.36 to 1.89). The percentage of the variability in effect estimates that was due to heterogeneity rather than chance is not important ($I^2$ = 0%) (Analysis 4.2) (Feng 2016; Luger 2020; Polterauer 2012). The authors of Peiretti 2010 additionally found that the risk of disease progression for women with any remaining RD was higher than those with NMRD (n = 259, P = 0.032), but the magnitude of effect was not reported.

***LVRD 1 cm to 5 cm versus NMRD***

The Winter 2008 study, which included only women with stage IV disease, found that women who had LVRD between 1 cm and 5 cm after PDS had more than twice the risk of disease progression compared to women with NMRD (HR 2.15, 95% CI 1.38 to 3.34; 193 participants) (Analysis 7.2).

***LVRD > 5 cm versus NMRD***

The Winter 2008 study, which included only women with stage IV disease, found that women who had LVRD between 1 cm and 5 cm after PDS had nearly three times the risk of disease progression

compared to women with NMRD (HR 2.96, 95% CI 1.86 to 4.71; 118 participants) (Analysis 8.2).

***RD > 2 cm versus RD < 2 cm***

The Winter 2008 study, which included only women with stage IV disease, found that women with LVRD > 2 cm after PDS had a slightly greater risk of disease progression compared to those with RD < 2 cm (HR 1.27, 95% CI 1.01 to 1.61; 253 participants) (Analysis 11.2).

Winter 2008 included NMRD in the < 2 cm category, but this only represented a small proportion in the analysis (n = 29/157 of participants in the RD < 2 cm category).

**Residual disease after interval debulking surgery (IDS)**

All meta-analyses included studies with participants with stage III and IV disease, other than in three studies where a specific breakdown was not reported (Kaban 2017; Lecuru 2019; Lorusso 2016). Therefore, we could not conduct subgroup analyses by stage to explore any underlying clinical heterogeneity between the studies as planned. However, we did perform subgroup analyses including cycle duration where possible (see Analysis 12.1 to Analysis 14.4). There was no evidence of any subgroup differences and all analyses were robust to the findings of the overall pooled estimates for all comparisons, with the exception of overall survival in the comparison of any remaining macroscopic disease versus NMRD (test for subgroup differences P = 0.01) (Analysis 15.1). However, the general direction of effect estimates was consistent and the findings were robust.

Davidson 2019 reported disease-specific survival (DSS) rather than overall survival, but this study did not appear to introduce statistical heterogeneity from visual inspection of the forest plot and the conclusions were robust to its exclusion in Analysis 14.2.

All comparisons involving SVRD included NMRD when compared to LVRD > 1 cm unless otherwise stated. The Phillips 2018 study was the only exception to this and reported an adequate comparison of LVRD > 1 cm versus SVRD using recognised RD threshold definitions, i.e. > 0 cm but < 1 cm residual disease as distinct from NMRD.

The comparison involving any remaining macroscopic disease (RD > 0 cm) and NMRD in an IDS setting was also an important comparison so we additionally gave this a certainty of the evidence judgement (Summary of findings 7).

**Overall survival (risk of death from all causes)**

***SVRD versus NMRD***

Meta-analysis of two groups of women from the same study undergoing different chemotherapy schedules found that women with SVRD after IDS had more than twice the risk of death compared to women with NMRD (HR 2.09, 95% CI 1.20 to 3.66; 310 participants) (Phillips 2018). The percentage of the variability in effect estimates that was due to heterogeneity rather than chance may represent moderate heterogeneity ($I^2$ = 56%) (Analysis 12.1). The magnitude of this effect was greater in this study in women who received > 4 cycles of neoadjuvant chemotherapy prior to their IDS (HR 2.78, 95% CI 1.66 to 4.65), but there was no significant difference or certainly a suggestion that there may be less of a difference between women with NMRD and those with SVRD when receiving ≤ 4 cycles of chemotherapy prior to IDS (HR 1.57, 95% CI 0.93 to 2.66) (Summary of findings 4).

The authors of Petrillo 2014 additionally found that the risk of death for women with SVRD after neoadjuvant chemotherapy (the majority received three or four cycles) before IDS was significantly higher than those with NMRD (n = 322, P = 0.001), but the magnitude of effect was not reported.

***LVRD (> 1 cm) versus NMRD***

Meta-analysis of two groups of women with different chemotherapy schedules, as outlined above, assessing 343 participants, found that women with LVRD > 1 cm after IDS had more than twice the risk of death compared to women with NMRD (HR 2.23, 95% CI 1.49 to 3.34). The percentage of the variability in effect estimates that was due to heterogeneity rather than chance may represent moderate heterogeneity ($I^2$ = 35%) (Analysis 13.1) (Phillips 2018). The magnitude of this effect was more pronounced in this study in women who received > 4 cycles of neoadjuvant chemotherapy prior to IDS (HR 2.67, 95% CI 1.76 to 4.06) (Summary of findings 5).

***RD > 1 cm versus RD < 1 cm***

Only Phillips 2018 compared SVRD versus LVRD > 1 cm in the strict sense that SVRD is mutually exclusive of NMRD. This was an important comparison and meta-analysis of the two groups in the study (three to six chemotherapy cycles) showed little difference in the risk of death between the SVRD and LVRD thresholds (HR 1.02, 95% CI 0.68 to 1.55; 343 participants). The percentage of the variability in effect estimates that was due to heterogeneity rather than chance was not important ($I^2$ = 0%) (Analysis 14.1).

The other five studies included NMRD in the SVRD category (referring to it as 'optimal') in their multivariate analyses. Nearly half of the women (261/550 (47%)) in the SVRD thresholds included NMRD in three studies (Cioffi 2018; Davidson 2019; Zhang 2018). This was not reported in the other two studies (Kaban 2017; Zhu 2016).

A sensitivity analysis that meta-analysed all six studies, assessing 1572 participants, found that women with LVRD > 1 cm after IDS had a statistically significant greater risk of death compared to women with SVRD or NMRD (HR 1.60, 95% CI 1.21 to 2.11). The percentage of the variability in effect estimates that was due to heterogeneity rather than chance may represent substantial heterogeneity ($I^2$ = 58%) (Analysis 14.2) (Summary of findings 6) (Cioffi 2018; Davidson 2019; Kaban 2017; Phillips 2018; Zhang 2018; Zhu 2016).

Sensitivity analysis, excluding Phillips 2018, led to an increase in effect estimates in a meta-analysis involving the five remaining studies (HR 1.84, 95% CI 1.34 to 2.52; 1429 participants; $I^2$ = 61%) (Analysis 14.3).

***RD > 0 cm versus NMRD***

Meta-analysis of four studies, assessing 906 women, found that any macroscopic RD after IDS was associated with more than twice the risk of death compared with NMRD (HR 2.11, 95% CI 1.35 to 3.29). The percentage of the variability in effect estimates that was due to heterogeneity rather than chance may represent considerable heterogeneity ($I^2$ = 81%) (Analysis 15.1) (Iwase 2015; Lecointre 2020; Lorusso 2016; Stoeckle 2014). For subgroup analysis by duration of NACT, we found evidence of a subgroup difference (P = 0.01, median of six cycles in two studies: N = 242, median four cycles in one study: N = 193, all range of cycles in one study: N = 471). However, the

direction of effect was consistent in all studies, showing a survival benefit in the NMRD group (Analysis 15.1).

The authors of Lecuru 2019 additionally found that the risk of death for women with any remaining RD (> 0 cm) after IDS was significantly higher than those with NMRD (n = 163, P < 0.01), but the magnitude of effect was not reported (Summary of findings 7).

**Progression-free survival (risk of disease progression)**

***SVRD (< 1 cm) versus NMRD***

Meta-analysis of two studies, assessing 248 women, found no difference in disease progression in women with SVRD after IDS and those with NMRD (HR 3.03, 95% CI 0.81 to 11.38). The percentage of the variability in effect estimates that was due to heterogeneity rather than chance alone may represent considerable heterogeneity ($I^2$ = 94%) (Analysis 12.2) (Bixel 2020; Liu 2020).

The authors of Petrillo 2014 found that the risk of disease progression for women with SVRD after IDS was higher than those with NMRD (n = 322, P = 0.001), but the magnitude of effect was not reported.

***LVRD (> 1 cm) versus SVRD***

Meta-analysis of four studies found that achieving LVRD > 1 cm after IDS was associated with a greater risk of disease progression compared to women in whom SVRD was achieved after surgery (HR 1.76, 95% CI 1.23 to 2.52; 1145 participants). The percentage of the variability in effect estimates that was due to heterogeneity rather than chance alone may represent substantial heterogeneity ($I^2$ = 60%) (Analysis 14.4) (Cioffi 2018; Shibutani 2020; Zhang 2018; Zhu 2016). These four studies included NMRD in the SVRD category (referring to it as 'optimal') in their multivariate analyses.

***RD > 0 cm versus NMRD***

The Lecointre 2020 study, assessing 471 women, found that RD > 0 cm after IDS was associated with an increased risk of disease progression compared those in whom NMRD was achieved (HR 1.36, 95% CI 1.05 to 1.76) (Analysis 15.2).

The authors of Lecuru 2019 found that the risk of disease progression for women with RD > 0 cm after IDS was higher than those with NMRD (n = 163, P < 0.01), but the magnitude of effect was not reported (Summary of findings 7).

# DISCUSSION

## Summary of main results

We found 46 studies reporting multivariate prognostic analyses that included residual disease (RD) as a prognostic factor, which met our inclusion criteria. These studies assessed survival after upfront primary debulking surgery (PDS) followed by adjuvant chemotherapy and neoadjuvant chemotherapy with interval debulking surgery (IDS) in advanced epithelial ovarian cancer. The review included 22,376 women who underwent PDS and 3697 women who underwent IDS, all with varying levels of RD. The main results of our review are summarised in the summary of findings tables (Summary of findings 1; Summary of findings 2; Summary of findings 3; Summary of findings 4; Summary of findings 5; Summary of findings 6; Summary of findings 7).

In PDS studies, meta- and single-study analyses demonstrate the prognostic importance of achieving no macroscopic residual disease (NMRD) after PDS for both overall and progression-free survival. Most studies showed an association with an increased risk of death in all groups with visible disease after surgery when compared to NMRD. The most pertinent comparison found that women who were debulked to leave small-volume residual disease (SVRD) after PDS had more than twice the risk of death compared to women with NMRD (meta-analysis of 17 studies: hazard ratio (HR) 2.03, 95% confidence interval (CI) 1.80 to 2.29; $I^2$ = 50%; 9404 participants; moderate-certainty evidence). Progression-free survival was not reported in all of the studies, but was sufficiently documented to allow conclusions to be drawn. The main comparison found that women who were debulked to SVRD after PDS had nearly twice the risk of disease progression compared to women with NMRD (meta-analysis of 10 studies: HR 1.88, 95% CI 1.63 to 2.16; $I^2$ = 63%; 6596 participants; moderate-certainty evidence). The fact that all of the studies included at least 100 women and used multivariate adjustment for important prognostic factors increased the level of certainty in the estimates.

When we compared large-volume residual disease (LVRD) (> 1 cm) versus SVRD cytoreduction the estimates were attenuated compared to the macroscopic RD comparisons. All analyses showed a survival benefit in women who had been debulked to leave SVRD (HR 1.22, 95% CI 1.13 to 1.32, $I^2$ = 0%, 6000 participants; moderate-certainty evidence). The results were robust to analyses of progression-free survival.

For neoadjuvant chemotherapy with IDS, the main comparisons involved any visible RD versus NMRD and LVRD (> 1 cm) versus SVRD. Unfortunately, it was not possible to distinguish those with NMRD after surgery within the SVRD thresholds in all but one study. A study reporting two groups of women on different chemotherapy schedules found that women who were debulked to leave SVRD and LVRD (> 1 cm) after IDS had more than twice the risk of death compared to women who had NMRD (HR 2.09, 95% CI 1.20 to 3.66; 310 participants; $I^2$ = 56% and HR 2.23, 95% CI 1.49 to 3.34; 343 participants; $I^2$ = 35%; very low-certainty evidence, for SVRD versus NMRD and LVRD versus NMRD, respectively). Women who had any amount of macroscopic RD (> 0 cm) after IDS had more than twice the risk of death compared to women with NMRD (HR 2.11, 95% CI 1.35 to 3.29, $I^2$ = 81%; 906 participants; very low-certainty evidence). Another study also found prolonged survival when RD was cytoreduced to NMRD (P < 0.01).

Unfortunately, in IDS studies the SVRD threshold included those with NMRD in all but one study (nearly half of women in the SVRD threshold had NMRD in three studies where it was reported). Therefore the reported comparison of NMRD or SVRD versus LVRD > 1 cm was of much lesser importance in IDS studies. Meta-analysis found that women who were debulked leaving LVRD > 1 cm had a greater risk of death and disease progression compared to women who were debulked to leave SVRD or NMRD (HR 1.60, 95% CI 1.21 to 2.11; 1572 participants; $I^2$ = 58% for overall survival and HR 1.76, 95% CI 1.23 to 2.52; 1145 participants; $I^2$ = 60% for progression-free survival; moderate-certainty evidence). The SVRD category included NMRD in all but one study, which suggests that only two categories of RD after IDS are being recognised at present, where NMRD remains of paramount prognostic importance.

## Overall completeness and applicability of evidence

The evidence from this review indicates that cytoreduction to NMRD after primary surgical cytoreduction is associated with prolonged survival in advanced epithelial ovarian cancer in both PDS and IDS settings. There is more strength in the evidence from studies reporting PDS, but the results suggest that the same conclusions apply in terms of the prognostic importance of NMRD in an IDS setting. More studies, including a larger number of women, will be needed to give more certainty in the effect estimates in comparisons of other RD thresholds, but there has been an emergence of studies using IDS in the last decade, so we expect this to be the case when the review is updated in the future. Interestingly, the comparison of SVRD versus LVRD (> 1 cm) is heterogeneously reported in the PDS and IDS analyses, as in the latter (IDS studies) the SVRD threshold included NMRD in all but one of the studies in the meta-analysis. Most studies included in the PDS analyses presented mutually exclusive RD thresholds, so there was less of a problem with NMRD being included in the SVRD category. The existing evidence does not currently support three categories of RD after IDS, as was recommended for PDS.

Although this review does not enable us to determine whether prolonged survival is a direct effect of the surgical intervention whereby women with NMRD do better, it appears that every effort should be made to attempt this, where possible, in both PDS and IDS settings. It may be particularly important in the latter due to issues with chemotherapy reaching allocation and further treatment options potentially being more limited thereafter. Where NMRD is considered not achievable for PDS, attempts should be made to obtain SVRD, defined as RD greater than 0 cm and less than or equal to 1 cm. There is limited evidence in this review to suggest that this may not be the case for IDS. Further data are needed, as understanding whether there is a benefit to IDS, if NMRD cannot be achieved, would be an important clinical question. However, as this is a prognostic review, we cannot answer this question from these data. Additionally, the data are of very low certainty - we are therefore very uncertain of this finding and drawing any conclusions would be unwise. Answering this question about the benefit of IDS, if NMRD cannot be achieved, would require an intervention study randomised controlled trial (RCT), rather than retrospective analysis of prognostic factors.

We found statistical heterogeneity between the studies in some analyses, but the direction of effect was consistent throughout so we had no concerns. We also did not have too many concerns about clinical heterogeneity as we applied restrictive inclusion criteria in terms of patient population, standardised measurement of RD as a prognostic factor and standard definitions of survival. Evaluation of other prognostic factors and biomarkers can often use different criteria for the interpretation of the results and different cut-off values may introduce levels of heterogeneity. Therefore, RD as a prognostic factor is unlikely to impact on the results or introduce any bias. That is, false-positive classifications seem much more unlikely than in other prognostic areas.

One of the strengths regarding the prognostic factor studies in this review was the ease of reporting in their statistical analyses. Authors mainly reported appropriate methods for their statistical analyses, with only a few studies not reporting the magnitude of effect estimates. We used hazard ratios (HRs) as the effect measure for time-to-event data in this review. We were able to provide pooled data for the majority of the included studies in the review. Of the studies that did not report appropriate statistics to extract for inclusion in the meta-analysis, we could not estimate the HR using other available data (Parmar 1998), as we restricted studies to those using multivariate analyses. We had limited success when contacting chief investigators to provide us with additional information or data from adjusted analyses.

In order to minimise bias, we only included studies of multivariate Cox regression models that used sensible adjustment factors associated with survival in women with advanced epithelial ovarian cancer (e.g. age, stage, grade, extent of disease at diagnosis). We excluded studies that only reported unadjusted results. To assess the adequacy of adjustment factors used in multivariate Cox models, we used the 'adjustment for other prognostic factors' and 'statistical analysis and reporting' domains of the quality in prognosis studies (QUIPS) tool (Riley 2019). Therefore, we prespecified in our protocol that we would only pool adjusted associations of the index prognostic factor. We felt that it was important to suggest a set of pertinent and established covariates a priori that were important to the disease under review (Riley 2019). This meant that we could better judge which models were adequate. We took these issues around the reporting in the studies into account when we assessed risk of bias and GRADE. The reported results in univariate analyses would have potentially been at a great risk of overestimating survival of RD as a prognostic factor. It is widely accepted that adjusting the predictive effect of a specific prognostic factor for the contribution of other prognostic factors strengthens the robustness of the evidence on the clinically relevant prognostic ability of that factor (Aldin 2020; Riley 2019).

Treatment-related morbidity very often degrades the quality of the time that women live, which is especially important after the completion of treatment for advanced cancer where women have poor prognosis and will want to enjoy a comfortable standard of living during their final months. It is unlikely that studies on prognosis will measure or report adverse events, so our focus was on survival as an outcome. This needs to be considered in the context of the findings from this review in that NMRD after PDS is associated with better survival - median survival for NMRD was 85.8 months (95% CI 77.5 to 94.1 months) in the Klar 2016 study, which included the largest analysis in the review. This study did include a small proportion of women with stage I and II cancer, but not to the extent of diluting the results too much. The next largest included study reporting median overall survival (71.9 months) also suggested that the potential benefits of prolonging survival may outweigh the disadvantages of any short-term morbidities associated with the surgical procedure (Winter 2007). Similarly, median survival in the NMRD group in IDS studies ranged from 50 months (Stoeckle 2014) to 51.8 (95% CI 45 to 58.5) months (Phillips 2018), the second largest analysis of IDS in this review. In terms of the overall survival rate in the NMRD group in IDS studies, Iwase 2015 reported a two-year and five-year overall survival rate of 88.8% and 43.4% respectively.

## Certainty of the evidence

Our certainty of the evidence is presented in the summary of findings tables (Summary of findings 1; Summary of findings 2; Summary of findings 3; Summary of findings 4; Summary of findings 5; Summary of findings 6; Summary of findings 7).

The 46 studies that met our inclusion criteria had reasonable risk of bias profiles when assessed using QUIPS as a prognostic risk of

bias tool (Riley 2019). We included only sufficiently large studies that controlled for various co-prognostic factors using multivariate analysis in order to reduce the possibility of bias.

The studies reported adjusted hazard ratio estimates using Cox proportional hazards models. A hazard ratio is the best statistic to summarise the difference in risk between groups over the duration of a study when there is 'censoring', that is the time to death (or disease progression) is unknown for some women as they are still alive (or disease-free) at the end of the study. Most studies were at moderate risk of bias as they satisfied some but not all of the criteria used to assess risk of bias. There were no real applicability concerns in any of the domains. This was largely due to the stringent and restrictive eligibility criteria. We were also cautious when deciding whether studies were selectively reported or whether any additional source of bias may have been present and assessed these items as being unclear.

In a PDS setting, for overall survival, all studies in the meta-analyses used adjusted results from multivariable analyses including important and well-established prognostic factors in women with advanced ovarian cancer, and the analyses all indicated the independent prognostic ability of thresholds of RD to predict overall survival. For comparisons of the three main reported RD thresholds (NMRD, SVRD and LVRD), we judged the certainty of the evidence as 'moderate' for all these comparisons (Summary of findings 1; Summary of findings 2; Summary of findings 3). We downgraded by one level for risk of bias due to some risk of bias concerns. With no firm guidance for grading the evidence in reviews of prognostic factor analyses (Riley 2019), we did not grade beyond these key RD thresholds. Similarly, progression-free survival was reported using the same methodology but in fewer studies. There was still a sufficient number to show that RD thresholds have an independent prognostic ability to predict progression-free survival. We also judged this outcomes to provide moderate-certainty evidence and we downgraded by one level for some risk of bias concerns (Summary of findings 1; Summary of findings 2; Summary of findings 3). We made the same certainty of the evidence judgements in an IDS setting for overall survival and progression-free survival. Only one study reported a comparison involving NMRD as a unique group (Phillips 2018). Furthermore, this same study was the only one to report the comparison of SVRD (< 1 cm) versus LVRD (> 1 cm) in the strict sense that SVRD was mutually exclusive of NMRD. The other studies reporting this outcome included NMRD in their SVRD group. Therefore, we downgraded overall survival and progression-free survival outcomes by one level. We also downgraded for some risk of bias concerns and insufficient and sparse data in the meta-analyses. Therefore the certainty of the evidence for overall survival and progression-free survival in an IDS setting was very low (Summary of findings 4; Summary of findings 5; Summary of findings 6; Summary of findings 7). The comparison of SVRD versus LVRD (> 1 cm) included one more study than the corresponding analysis involving PDS, but there were significantly fewer women in the analysis (less than a third) and the lack of separation of NMRD from the SVRD threshold was misleading, so that was reason it was judged to provide very low-certainty evidence (Summary of findings 6). Only one study truly reported an adequate comparison of LVRD versus SVRD.

In some cases, more data would be needed to see the full impact of leaving behind more considerable disease, although the evidence suggests that if it cannot be minimised to NMRD or SVRD it may not make a significant difference in terms of prolonged survival. The results are consistent and appear to be reliable and precise in terms of the conclusions drawn. Some comparisons were sparse, with wide confidence intervals, but even the lower 95% confidence interval would have been highly significant as a point estimate in many cases.

Further research is very unlikely to change our confidence in the estimates of effect in the larger and most pertinent meta-analyses (exclusively reported in a PDS setting), but may change the estimates for some of the comparisons involving head-to-head LVRD thresholds and in analyses that included IDS. However, in the latter the evidence base is likely to be strengthened in future years as there has been an emergence of evidence in the last decade that is expanding, given four RCTs have now demonstrated similar survival outcomes of PDS versus IDS, as reviewed by Coleridge 2021. However, this evidence needs to assess whether SVRD is associated with a survival benefit over LVRD in an IDS setting as the evidence is currently very uncertain.

### Strengths and weaknesses of the review

We performed a comprehensive search, including a thorough search of the grey literature, and two review authors working independently sifted and data extracted all studies. To prevent bias in this review, at least two review authors, along with willing arbiters, also independently performed all other relevant processes, such as risk of bias and GRADE assessment, and verification of all analyses. Although the methods for grading the evidence from prognosis studies are still under development, we felt that omitting it would be less transparent and potentially create bias in the review. Therefore we followed standard methodology for grading the certainty of the evidence and used specific exemplars from the Cochrane prognostic group for guidance, as well as examining other relevant prognostic factor reviews (Aldin 2020). We were not restrictive in our inclusion criteria with regards to types of studies, but limited to prognostic models that used multivariate analyses. This was to ensure that we minimised bias in getting accurate and reflective effect estimates for the prognostic performance of RD. We restricted to studies including at least 100 women in their analyses due to limiting the analyses to multivariate ones and the potential issue with adjustment for multiple prognostic factors in sparse data (Ogundimu 2016). There was more chance of drawing satisfactory conclusions in the review as the number of women in each study was adequate. We also conducted analyses using appropriate statistical methods for survival outcomes, namely hazard ratios, which correctly allow for censoring (see above).

In the analyses comparing SVRD versus LVRD (> 1 cm) for both PDS and IDS, we included studies that either treated NMRD as a distinct category of SVRD (Phillips 2018), or included NMRD within the SVRD category (Akahira 2001; Chan 2003; Chi 2001; Cioffi 2018; Davidson 2019; Kaban 2017; McGuire 1995; Winter 2008; Zhang 2018; Zhu 2016) during analyses. In keeping with the view that there is a dose-response relationship between RD thresholds and survival, the inclusion of these latter studies will have introduced an overestimation of the survival benefit of SVRD compared to LVRD (> 1 cm) and introduced serious bias. We attempted to determine the extent of this bias by identifying the number of participants with NMRD included in these latter studies, however this information was only provided in four studies – NMRD ranged from 27% (Winter 2008) to 72% (Davidson 2019). Results of analyses in PDS studies

were robust to the exclusion of studies that included NMRD within their SVRD category. Similar sensitivity analyses were not practical in the IDS case as only Phillips 2018 adequately reported the comparison involving SVRD that did not include NMRD.

A significant threat to the validity of the review is likely to be publication bias; that is, studies that did not find a positive association with the degree of surgical debulking achieved may not have been published. Although we conducted a test for funnel plot asymmetry and there did not appear to be any evidence of small study bias, such as publication bias, this type of test is not necessarily recommended for survival data due to issues of censoring (Debray 2018). Therefore, we cannot exclude potential publication bias and the presence of small study effects in our review (Riley 2019). Further investigation is beyond the scope of this review. Most included studies included in this review were retrospective and were probably not pre-registered. Studies are also not always labelled or indexed as prognosis studies, and search filters for studies on prognosis are still under development. Therefore the search had much wider scope than was necessary, but we felt it was better to be overly inclusive to reduce the chance of missing eligible studies for inclusion in the review.

### Agreements and disagreements with other studies or reviews

In our review, we included studies that have assessed residual disease (RD) as a prognostic factor after primary surgery in women with advanced epithelial ovarian cancer. Overall, the findings from this review are in agreement with similar reviews and studies that have investigated the prognostic value of NMRD in both PDS and IDS settings. They also support the findings that, in general, small-volume RD is associated with better survival after surgery. The majority of these studies reported univariate analyses and that was one of the exclusion criteria in our review. These univariate analyses widely reported larger magnitudes of effect giving greater levels of statistical significance, but our analyses restricted to estimates that adjusted for sensible covariates that were likely to give less biased and more reliable estimates. Many of these studies are documented in the Characteristics of excluded studies table.

The association with improved survival outcomes associated with NMRD categorisation consolidates use of the term 'optimal cytoreduction' by the Gynaecological Cancer Inter-Group (GCIG) to mean 'NMRD', from its former definition of < 1 cm RD, which we categorised as SVRD. Although the results of our review show that cytoreduction to SVRD is still superior to LVRD (> 1 cm).

In a PDS setting, if the term macroscopic cytoreduction is to be used solely for the group where there is NMRD, the moderate-certainty evidence in this review that women who undergo PDS and achieve SVRD still do better than women who achieve LVRD should prompt the surgical community to retain this category as well as SVRD for RD < 1 cm (and consider the term 'near-optimal'), while reserving the term LVRD (and consider using 'suboptimal') to cases where the RD is > 1 cm (a three-category classification of NMRD, SVRD and LVRD, or alternatively consider the terms 'optimal', 'near-optimal' and 'suboptimal' RD). In contrast, we obtained very low-certainty evidence from a single IDS study that showed a survival benefit for NMRD compared to SVRD (Phillips 2018). All but one study included NMRD in their comparison of SVRD versus LVRD (> 1 cm) so strong inferences were not possible. Evidence from this one study that reported a valid comparison found little difference in survival outcomes in this comparison of RD thresholds (Phillips 2018). Further evidence from a meta-analysis including four studies showed that achieving NMRD was associated with superior survival outcomes to having any remaining RD (> 0 cm) (Iwase 2015; Lecointre 2020; Lorusso 2016; Stoeckle 2014). Therefore, given the available evidence, the strongest conclusion renderable is a two-category classification following IDS (NMRD versus any RD > 0 cm).

The debate regarding whether a three-category classification should hold in both PDS and IDS has also surfaced amongst the surgical community in recent publications. To our knowledge, two retrospective studies of women with advanced epithelial ovarian cancer provided evidence pertinent to this debate (Ghirardi 2020; Kobal 2018). One rationale behind these studies was to address whether women in whom PDS achieved SVRD would be conferred similar or better survival compared to those in whom NMRD was achieved following IDS. In the Kobal 2018 study, amongst women achieving NMRD, the IDS group had poorer overall survival (36.3 versus 54.7 months; P = 0.012) but similar progression-free survival (19.9 versus 20.7 months; P = 0.251) compared to the PDS group. On the other hand, achieving NMRD following IDS was associated with similar overall survival (36.3 versus 34.7 months; P = 0.073), but better progression-free survival (19.9 versus 11.2 months; P = 0.005) compared to achieving SVRD following PDS. In contrast, Ghirardi 2020 found that achieving NMRD following IDS was associated with poorer overall survival compared to achieving SVRD following PDS (41.4 versus 52.4 months; P = 0.022). Given the unadjusted estimates and retrospective nature of these studies, and that these compare prognostic factors and not treatment effects, conclusions about the relative merits of different treatments cannot be made. However, they do reflect an ongoing point of discussion, and contribute towards a burgeoning empirical basis for either a two-threshold 'all-or-nothing' classification system following IDS (NMRD versus RD > 0 cm) or the retention of the three-threshold classification. The results of our review appear to lend support for the two-threshold classification following IDS based on the conduct of the included studies, although this is more on the grounds that there is a lack of evidence of significant differences in survival between SVRD and LVRD (> 1 cm) thresholds due to lack of reporting of this comparison.

A Cochrane Review by Coleridge 2021 compared intervention RCTs directly comparing PDS versus IDS (Chekman 2015; Fagotti 2020; Kehoe 2015; Onda 2020; Vergote 2010). The included studies did not meet our inclusion criteria, as they did not report results across RD thresholds. Within this review, Kehoe 2015 and Vergote 2010 randomised 1270 participants (of which 1220 were assessed), compared PDS versus IDS and provided a breakdown of extent of disease by type of surgery (but did not give breakdown of differences within RD thresholds for each type of surgery, so did not meet our inclusion criteria). Both trials recruited participants with stage IIIC and IV epithelial ovarian cancer. Both trials reported RD thresholds that included NMRD (optimal), SVRD (RD < 1 cm) and LVRD (RD > 1 cm). The two trials found no significant difference in overall survival for the comparison of extent of RD threshold (NMRD, SVRD and LVRD) by primary surgery (upfront versus interval). The trial Vergote 2010 reported no significant difference between PDS and IDS for SVRD or NMRD (RD < 1cm including 0 cm) (HR 1.17, 95% CI 0.82 to 1.67). There were also no significant differences observed for SVRD (< 1 cm) (HR 1.22, 95% CI 0.84 to 1.77) and LVRD thresholds (HR 0.91, 95% CI 0.89 to 1.30) by type of surgery. Similarly, the authors of Kehoe 2015 reported a P value of 0.98 for the interaction

between treatment and extent of RD after debulking. It should be emphasised that these studies were RCTs designed to measure the effect of PDS versus IDS and were not designed to evaluate the intervention of differing degrees of surgical effort.

The results of the SOCQER-2 study assessing quality of life and progression-free survival found that patients with late-stage ovarian cancer had no important differences in EORTC QLQ-C30 global scores measured across six weeks, six months and 12 months post-surgery when undergoing surgery of varying complexity, despite a higher preoperative disease burden in patients undergoing more radical surgical procedures (Sundar 2022). The authors of the study found that patients who underwent low-complexity surgery had higher rates of residual disease and lower survival compared with those with a similar disease burden undergoing surgery of intermediate complexity. However, no statistical adjustment was performed in these analyses. Postoperative residual disease was associated with poorer overall survival, particularly in patients undergoing low-complexity surgery, but again no statistical adjustment was made and, as this was not an intervention study, it is not able to determine the causal effect of this relationship.

Women with FIGO stage IIIC disease with extra pelvic metastases smaller than 5 cm have been shown to have better progression-free survival after upfront debulking (Vergote 2018). An investigation of NMRD during the initial treatment of epithelial ovarian cancer comparing PDS versus IDS has been investigated in a TRUST (Trial of Radical Upfront Surgical Therapy in advanced ovarian cancer (ENGOT ov33/AGO-OVAR OP7)) trial, which is due to report in 2024 (Reuss 2019).

## AUTHORS' CONCLUSIONS

### Implications for practice

In a primary debulking surgery (PDS) setting, this review provides moderate-certainty evidence that residual disease (RD) after primary surgery is a strong prognostic factor for overall and progression-free survival in women with advanced ovarian cancer. The certainty of the evidence for these outcomes was very low for studies involving interval debulking surgery (IDS). We conclude that there should be three distinct categories of RD after PDS, including no macroscopic residual disease (NMRD) (labelled as optimal), < 1 cm (labelled as small-volume residual disease (SVRD) and strictly meaning 0.1 cm to 1 cm) and > 1 cm (large-volume residual disease (LVRD)).

After IDS, there may be only two categories required, although this is based on very low-certainty evidence and it would be unwise to make any firm inferences or conclusions until further studies are added to the evidence base.

It is acknowledged that there is considerable variation in achieving NMRD or SVRD between different surgeons and centres. Predicting the achievement of NMRD or SVRD prior to surgery will be dependent on this variation, resulting in difficulties in developing models of prediction, so deciding on whether to perform PDS or IDS at present is down to clinician preference.

NMRD remains a key prognosticator of survival in advanced ovarian cancer. Whether PDS or IDS is the primary treatment, the surgical goal should be to completely remove all visible disease, although SVRD should still be regarded as a favourable outcome after PDS, as shown in this systematic review, although this is not clear following IDS.

The evidence on the ability of different thresholds of RD to distinguish between a good and bad prognosis can aid decision-making for clinicians and diagnosed individuals, where the survival advantage can be considered alongside any potential morbidity or adverse event trade-offs.

### Implications for research

The purpose of this systematic review was to assess RD as a prognostic factor in women who received primary surgery (PDS and IDS) for advanced epithelial ovarian cancer (stages III and IV). The results should encourage the surgical community to make trials in this area a priority. Future research should focus on investigations that determine whether increasing attempts at achieving NMRD have a direct effect in improving survival outcomes using methodologies and trial designs that reduce or eliminate confounding effects, such as the women's performance status, disease spread and tumour biology.

Greater emphasis should be made in future studies to investigate IDS to raise the certainty of the evidence profiles. In both PDS and IDS settings, quality of life parameters and adverse effects and complications of the surgery need to be adequately addressed as there are significant deficiencies in previous studies in evaluating these outcome measures. It is unlikely that studies on prognosis will measure or report adverse events, so our focus in this review was on survival as an outcome. These additional evaluations should be given high priority, as this systematic review has identified large differences in survival outcomes associated with LVRD compared to when NMRD is achieved. The results of the SOCQER-2 study suggest that quality of life may still be reasonable even after more extensive surgery, which is reassuring, although this was an observational study (Sundar 2022). An investigation of cytoreductive surgery during the initial treatment of epithelial ovarian cancer comparing PDS versus IDS has also been investigated in a TRUST (Trial of Radical Upfront Surgical Therapy in advanced ovarian cancer (ENGOT ov33/AGO-OVAR OP7)) trial and we await the results in 2024 (Reuss 2019).

To avoid continuous confounding of results, observational studies should report the following to better assess the effect of surgical treatment in advanced ovarian cancer:

- Structural selection – the specific setting in which women are referred/seek care and which sample of the population (or population) has been chosen.
- To what extent the population of women with ovarian cancer are accounted for (selection of patients macro level).
- Institutional selection – how women were selected for surgery (choice of surgeon, patient, etc.).
- The extent of surgery needed to achieve complete resection, i.e. procedures and surgical complexity scores (surgical proficiency).
- Complete resection rates.

## ACKNOWLEDGEMENTS

We thank Jo Morrison, Tracey Harrison and Gail Quinn from Cochrane Gynaecological, Neuro-oncology and Orphan Cancers

(GNOC) for their advice support and contribution to the editorial process. We thank Nicole Skoetz from the Prognostic Methods Group for her valuable input. We also thank Jo Platt, Information Manager for GNOC, for designing the search strategies.

The authors and GNOC are grateful to the following peer reviewers for their time and comments; Simon Butler-Manuel, Jennifer Hare, Sonali Kaushik, Hans Nagar and Sahar Salehi.

This project was supported by the National Institute for Health Research (NIHR), via Cochrane infrastructure funding to Cochrane Gynaecological, Neuro-oncology and Orphan Cancers. The views and opinions expressed therein are those of the review authors and do not necessarily reflect those of the Systematic Reviews Programme, NIHR, NHS, or the Department of Health.

# REFERENCES

## References to studies included in this review

**Akahira 2001** *{published data only}*

Akahira JI, Yoshikawa H, Shimizu Y, Tsunematsu R, Hirakawa T, Kuramoto H. Prognostic factors of stage IV epithelial ovarian cancer: a multicenter retrospective study. *Gynecologic Oncology* 2001;**81**(3):398-403.

**Aletti 2006** *{published data only}*

Aletti GD, Dowdy SC, Gostout BS, Jones MB, Stanhope CR, Wilson TO, et al. Aggressive surgical effort and improved survival in advanced-stage ovarian cancer. *Obstetrics and Gynecology* 2006;**107**:77–85.

Aletti GD, Dowdy SC, Podratz KC, Cliby WA. Relationship among surgical complexity, short-term morbidity, and overall survival in primary surgery for advanced ovarian cancer. *American Journal of Obstetrics and Gynecology* 2007;**197**(6):1-7.

Aletti GD, Dowdy SC, Podratz KC, Cliby WA. Surgical treatment of diaphragm disease correlates with improved survival in optimally debulked advanced stage ovarian cancer. *Gynecologic Oncology* 2006;**100**(2):283-7.

Aletti GD, Podratz KC, Jones MB, Cliby WA. Role of rectosigmoidectomy and stripping of pelvic peritoneum in outcomes of patients with advanced ovarian cancer. *Journal of the American College of Surgeons* 2006;**203**(4):521-6.

**Ataseven 2016** *{published data only}*

Ataseven B, Grimm C, Harter P, Heitz F, Traut A, Prader S, et al. Prognostic impact of debulking surgery and residual tumor in patients with epithelial ovarian cancer FIGO stage IV. *Gynecologic Oncology* 2016;**140**(2):215-20.

**Bixel 2020** *{published data only}*

Bixel K, Vetter M, Davidson B, Berchuck A, Cohn D, Copeland L, et al. Intraperitoneal chemotherapy following neoadjuvant chemotherapy and optimal interval tumor reductive surgery for advanced ovarian cancer. *Gynecologic Oncology* 2020;**156**:530–4.

**Bristow 2011** *{published data only}*

Bristow RE, Ueda S, Gerardi MA, Ajiboye OB, Ibeanu OA. Analysis of racial disparities in stage IIIC epithelial ovarian cancer care and outcomes in a tertiary gynecologic oncology referral center. *Gynecologic Oncology* 2011;**122**(2):319-23.

**Chan 2003** *{published data only}*

Chan JK, Loizzi V, Lin YG, Osann K, Brewster WR, DiSaia PJ. Stages III and IV invasive epithelial ovarian carcinoma in younger versus older women: what prognostic factors are important? *Obstetrics and Gynecology* 2003;**102**(1):156-61.

**Chang 2012a** *{published data only}*

Chang SJ, Bristow RE, Ryu HS. Impact of complete cytoreduction leaving no gross residual disease associated with radical cytoreductive surgical procedures on survival in advanced ovarian cancer. *Annals of Surgical Oncology* 2012;**19**(13):4059-67.

**Chang 2012b** *{published data only}*

Chang SJ, Bristow RE, Ryu HS. Prognostic significance of systematic lymphadenectomy as part of primary debulking surgery in patients with advanced ovarian cancer. *Gynecologic Oncology* 2012;**126**(3):381-6.

**Chi 2001** *{published data only}*

Chi DS, Liao JB, Leon LF, Venkatraman ES, Hensley ML, Bhaskaran D, et al. Identification of prognostic factors in advanced epithelial ovarian carcinoma. *Gynecologic Oncology* 2001;**82**(3):532-7.

**Chi 2006** *{published data only}*

Chi DS, Eisenhauer EL, Lang J, Huh J, Haddad L, Abu-Rustum NR, et al. What is the optimal goal of primary cytoreductive surgery for bulky stage IIIC epithelial ovarian carcinoma (EOC)? *Gynecologic Oncology* 2006;**103**:559–64.

Eisenhauer EL, Abu-Rustum NR, Sonoda Y, Aghajanian C, Barakat RR, Chi DS. The effect of maximal surgical cytoreduction on sensitivity to platinum-taxane chemotherapy and subsequent survival in patients with advanced ovarian cancer. *Gynecologic Oncology* 2008;**108**(2):276-81.

**Cioffi 2018** *{published data only}*

Cioffi R, Bergamini A, Rabaiotti E, Petrone M, Pella F, Ferrari D, et al. Neoadjuvant chemotherapy in high-risk ovarian cancer patients: role of age. *Tumori Journal* 2018;**105**(2):168-73.

**Cuylan 2018** *{published data only}*

Cuylan ZF, Meydanli MM, Sari ME, Akbayir O, Celik H, Dede M, et al. Prognostic factors for maximising or optimally cytoreduced stage III non serous epithelial ovarian carcinoma treated with carboplatin/paclitaxel chemotherapy. *Journal of Obstetrics and Gynaecology Research* 2018;**44**(7):1284-93.

**Davidson 2019** *{published data only}*

Davidson BA, Broadwater G, Crim A, Boccacio R, Bixel K, Backes F, et al. Surgical complexity score and role of laparoscopy in women with advanced ovarian cancer treated with neoadjuvant chemotherapy. Gynecologic Oncology 2019;**152**(3):554-9.

**Eisenkop 2003** *{published data only}*

Eisenkop SM, Friedman RL, Wang HJ. Complete cytoreductive surgery is feasible and maximizes survival in patients with advanced epithelial ovarian cancer: a prospective study. *Gynecologic Oncology* 1998;**69**(2):103-8.

Eisenkop SM, Spirtos NM, Friedman RL, Lin WC, Pisani AL, Perticucci S. Relative influences of tumor volume before surgery and the cytoreductive outcome on survival for patients with advanced ovarian cancer: a prospective study. *Gynecologic Oncology* 2003;**90**(2):390-6.

**Feng 2016** *{published data only}*

Feng Z, Wen H, Bi R, Yang W, Wu X. Prognostic impact of the time interval from primary surgery to intravenous chemotherapy

in high grade serous ovarian cancer. *Gynecologic Oncology* 2016;**141**:466-70.

**Hofstetter 2013** *{published data only}*

Hofstetter G, Concin N, Braicu I, Chekerov R, Sehouli J, Cadron I. The time interval from surgery to start of chemotherapy significantly impacts prognosis in patients with advanced serous ovarian carcinoma - analysis of patient data in the prospective OVCAD study. *Gynecologic Oncology* 2013;**131**(1):15-20.

**Iwase 2015** *{published data only}*

Iwase H, Takada T, Iitsuka C, Nomura H, Abe A, Taniguchi T, et al. Clinical significance of systematic retroperitoneal lymphadenectomy during interval debulking surgery in advanced ovarian cancer patients. *Journal of Gynecologic Oncology* 2015;**26**(4):303-10.

**Kaban 2017** *{published data only}*

Kaban A, Topuz S, Saip P, Sozen H, Celebi K, Salihoglu Y. Poor prognostic factors in patients undergoing surgery after neoadjuvant chemotherapy for ovarian, tubal, or peritoneal cancer. *Journal of Obstetrics and Gynaecology* 2017;**39**(12):1163-70.

**Kahl 2017** *{published data only}*

Kahl A, du Bois A, Harter P, Prader S, Schneider S, Heitz F, et al. Prognostic value of the age-adjusted Charlson comorbidity index (ACCI) on short- and long-term outcome in patients with advanced primary epithelial ovarian cancer. *Annals of Surgical Oncology* 2017;**24**:3692-9.

**Klar 2016** *{published data only}*

du Bois A, Lück H, Meier W, Adams HP, Möbus V, Costa S, et al. A randomized clinical trial of cisplatin/paclitaxel versus carboplatin/paclitaxel as first-line treatment of ovarian cancer. *Journal of the National Cancer Institute* 2003;**95**(17):1320-9.

du Bois A, Weber B, Rochon J, Meier W, Goupil A, Olbricht S, , et al. Addition of epirubicin as a third drug to carboplatin/paclitaxel in first-line treatment of advanced ovarian cancer: a prospectively randomized gynecologic cancer intergroup trial by the Arbeitsgemeinschaft Gynaekologische Onkologie Ovarian Cancer Study Group and the Groupe d'Investigateurs Nationaux pour l'Etude des Cancers Ovariens. *Journal of Clinical Oncology* 2006;**24**(7):1127-35.

du Bois A HJ, Hardy-Bessard AC, Muller HH, Harter P, Kristensen G. Phase III trial of carboplatin plus paclitaxel with or without gemcitabine in first-line treatment of epithelial ovarian cancer. *Journal of Clinical Oncology* 2010;**28**(27):4162-9.

Klar M, Hasenburg A, Hasenov M, Hilpert F, Meier W, Pfisterer J, et al. Prognostic factors in young ovarian cancer patients: an analysis of four prospective phase III intergroup trials of the AGO Study Group, GINECO and NSGO. *European Journal of Cancer* 2016;**66**:114-24.

Mahner S, Eulenburg C, Staehle A, Wegscheider K, Reuss A, Pujade-Lauraine E, et al. Prognostic impact of the time interval between surgery and chemotherapy in advanced ovarian cancer: analysis of prospective randomised phase III trials. *European Journal of Cancer* 2013;**49**(1):142-9.

Pfisterer JWB, Reuss A, Kimmig R, du Bois A, Wagner U, et al. Randomized phase III trial of topotecan following carboplatin and paclitaxel in first-line treatment of advanced ovarian cancer: a gynecologic cancer intergroup trial of the AGO-OVAR and GINECO. *Journal of the National Cancer Institute* 2006;**98**(15):1036-45.

**Langstraat 2011** *{published data only}*

Langstraat C, Aletti GD, Cliby WA. Morbidity, mortality and overall survival in elderly women undergoing primary surgical debulking for ovarian cancer: a delicate balance requiring individualization. *Gynecologic Oncology* 2011;**123**(2):187-91.

**Lecointre 2020** *{published data only}*

Lecointre L, Velten M, Lodi M, Saadeh R, Lavoué V, Ouldamer L, et al. Impact of neoadjuvant chemotherapy cycles on survival of patients with advanced ovarian cancer: a French national multicenter study (FRANCOGYN). *European Journal of Obstetrics & Gynecology and Reproductive Biology* 2020;**245**:64–72.

**Lecuru 2019** *{published data only}*

Lecuru F, Pujade-Lauraine E, Hamizi S, Caumont-Prim A, Raban N, Malaurie E, et al. 1016P - Surrogate endpoint of progression-free (PFS) and overall survival (OS) for advanced ovarian cancer (AOC) patients (pts) treated with neo-adjuvant chemotherapy (NACT): results of the CHIVA randomized phase II GINECO study. *Abstract Book of the 44th ESMO Congress (ESMO 2019) 27 September - 1 October 2019, Barcelona, Spain* 2019;**30**:v415.

**Liu 2020** *{published data only}*

Liu Y, Cao L, Chen W, Wang J, Wang W, Liang Z. Feasibility of neoadjuvant and adjuvant intraperitoneal chemotherapy in patients with advanced epithelial ovarian cancer: a single-center experience. *Medicine* 2020;**99**(36):e22100.

**Lorusso 2016** *{published data only}*

Lorusso D, Bogani G, Matteucci L, Ditto A, Tamberi S, Arcangeli V, et al. Number of cycles of neoadjuvant chemotherapy might influence survival of patients undergoing interval debulking surgery for non-cytoreducible ovarian cancer: results from a multi-institutional study. *International Journal of Gynecological Cancer* 2016;**26**:647.

**Luger 2020** *{published data only}*

Luger AK, Steinkohl F, Aigner F, Jaschke W, Marth C, Zeimet AG, et al. Enlarged cardiophrenic lymph nodes predict disease involvement of the upper abdomen and the outcome of primary surgical debulking in advanced ovarian cancer. *Acta Obstetricia et Gynecologica Scandinavica* 2020;**99**:1092–9.

**McGuire 1995** *{published data only}*

Hoskins WJ, McGuire WP, Brady MF, Homesley HD, Creasman WT, Berman M, et al. The effect of diameter of largest residual disease on survival after primary cytoreductive surgery in patients with suboptimal residual epithelial ovarian carcinoma. *American Journal of Obstetrics and Gynecology* 1994;**170**(4):974-80.

McGuire WP, Hoskins WJ, Brady MF, Homesley HD, Creasman WT, Berman ML, et al. Assessment of dose-intensive therapy in suboptimally debulked ovarian cancer: a Gynecologic Oncology Group Study. *Journal of Clinical Oncology* 1995;**13**:1589-99.

**Melamed 2017a** *{published data only}*

Melamed A, Manning-Geist B, Bregar AJ, Diver EJ, Goodman A, del Carmen MG, et al. Associations between residual disease and survival in epithelial ovarian cancer by histologic type. *Gynecologic Oncology* 2017;**14**(7):250-6.

**Melamed 2017b** *{published data only}*

Melamed A, Manning-Geist B, Bregar AJ, Diver EJ, Goodman A, del Carmen MG, et al. Associations between residual disease and survival in epithelial ovarian cancer by histologic type. *Gynecologic Oncology* 2017;**14**(7):250-6.

**Paik 2018** *{published data only}*

Paik ES, Kim JH, Kim TJ, Lee JW, Kim BG, Bae DS, et al. Prognostic significance of normal-sized ovary in advanced serous epithelial ovarian cancer. *Journal of Gynecologic Oncology* 2018;**29**(1):e13.

**Peiretti 2010** *{published data only}*

Peiretti M, Zanagnolo V, Aletti GD, Bocciolone L, Colombo N, Landoni F, et al. Role of maximal primary cytoreductive surgery in patients with advanced epithelial ovarian and tubal cancer: surgical and oncological outcomes. *Gynecologic Oncology* 2010;**119**(2):259-64.

**Peiretti 2012** *{published data only}*

Peiretti M, Bristow RE, Zapardiel I, Gerardi M, Zanagnolo V, Biffi R, et al. Rectosigmoid resection at the time of primary cytoreduction for advanced ovarian cancer. A multi-center analysis of surgical and oncological outcomes. *Gynecologic Oncology* 2012;**126**(2):220-3.

**Petrillo 2014** *{published data only}*

Petrillo M, Zannoni GF, Tortorella L, Pedone AL, Salutari V, Ercoli A, et al. Prognostic role and predictors of complete pathologic response to neoadjuvant chemotherapy in primary unresectable ovarian cancer. *American Journal of Obstetrics and Gynecology* 2014;**211**(6):632.e1-8.

**Phillips 2018** *{published data only}*

Phillips A, Sundar S, Singh K, Nevin J, Elattar A, Kehoe S, et al. Complete cytoreduction after five or more cycles of neo-adjuvant chemotherapy confers a survival benefit in advanced ovarian cancer. *European Journal of Surgical Oncology* 2018;**44**(6):760-5.

**Polterauer 2012** *{published data only}*

Polterauer S, Vergote I, Concin N, Braicu I, Chekerov R, Mahner S. Prognostic value of residual tumor size in patients with epithelial ovarian cancer FIGO stages IIA-IV: analysis of the OVCAD data. *International Journal of Gynecological Cancer* 2012;**22**(3):380-5.

**Shibutani 2020** *{published data only}*

Shibutani T, Nagao S, Suzuki K, Kaneda M, Yamamoto K, Jimi T. `Dose-dense` paclitaxel and carboplatin vs. conventional paclitaxel and carboplatin as neoadjuvant chemotherapy for advanced epithelial ovarian, fallopian tube, or primary peritoneal cancer: a retrospective study. *International Journal of Clinical Oncology* 2020;**25**:502–7.

**Shim 2016** *{published data only}*

Shim SH, Kim DY, Jung PS, Kim JH, Kim YM, Suh DS, et al. Prognostic impact of the time interval from surgery to chemotherapy in patients with advanced ovarian cancer. *Gynecologic Oncology* 2016;**141**(S1):49.

**Stoeckle 2014** *{published data only}*

Stoeckle E, Bourdarias L, Guyon F, Croce S, Brouste V, Thomas L, et al. Progress in survival outcomes in patients with advanced ovarian cancer treated by neo-adjuvant platinum/taxane-based chemotherapy and late interval debulking surgery. *Annals of Surgical Oncology* 2014;**21**(2):629-36.

**Tewari 2016** *{published data only}*

Tewari KS, Java JJ, Eskander RN, Monk BJ, Burger RA. Early initiation of chemotherapy following complete resection of advanced ovarian cancer associated with improved survival: NRG Oncology/Gynecologic Oncology Group study. *Annals of Oncology* 2016;**27**:114-21.

**Tseng 2018** *{published data only}*

Tseng JH, Cowan RA, Zhou Q, Lasonos A, Byrne M, Polcino T, et al. Continuous improvement in primary debulking surgery for advanced ovarian cancer: do increased complete Gross resection rates independently lead to increased progression-free and overall survival? *Gynecologic Oncology* 2018;**151**(1):24-31.

**Van Geene 1996** *{published data only}*

Van Geene P, Varma R, Dunn J, Chan KK, Luesley DM. The prognostic significance of intraperitoneal growth characteristics in epithelial ovarian carcinoma. *International Journal of Gynecological Cancer* 1996;**6**(3):219-24.

**Wimberger 2010** *{published data only}*

Wimberger P, Wehling M, Lehmann N, Kimmig R, Schmalfeldt B, Burges A. Influence of residual tumor on outcome in ovarian cancer patients with FIGO stage IV disease: an exploratory analysis of the AGO-OVAR (Arbeitsgemeinschaft Gynaekologische Onkologie Ovarian Cancer Study Group). *Annals of Surgical Oncology* 2010;**17**(6):1642-8.

**Winter 2007** *{published data only}*

Armstrong DK, Bundy BW, Wenzel L, Huang HQ, Baergen RL, Shashikant L, et al, the Gynecologic Oncology Group. Intraperitoneal cisplatin and paclitaxel in ovarian cancer. *New England Journal of Medicine* 2006;**354**(1):34-43.

Markman M, Bundy BM, Alberts DS, Fowler JM, Clark-Pearson DL, Carson LF, et al. Phase III trial of standard-dose intravenous cisplatin plus paclitaxel versus moderately high-dose carboplatin followed by intravenous paclitaxel and intraperitoneal cisplatin in small-volume stage III ovarian

carcinoma: An Intergroup Study of the Gynecologic Oncology Group, Southwestern Oncology Group, and Eastern Cooperative Oncology Group. *Journal of Clinical Oncology* 2001;**19**(4):1001-7.

McGuire WP, Hoskins WJ, Brady MF, Kucera PR, Partridge EE, Look KY, et al. Cyclophosphamide and cisplatin compared with paclitaxel and cisplatin in patients with stage III and stage IV ovarian cancer. *New England Journal of Medicine* 1996;**334**(1):1-6.

Muggia FM, Braly PS, Brady MF, Sutton GN, Theodore HL, Samuel LA, et al. Phase III randomized study of cisplatin versus paclitaxel versus cisplatin and paclitaxel in patients with suboptimal stage III or IV ovarian cancer: a Gynecologic Oncology Group Study. *Journal of Clinical Oncology* 2000;**18**(1):106-15.

Ozols RF, Bundy BN, Greer BE, Fowler JM, Clarke-Pearson DB, Robert AM, et al. Phase III trial of carboplatin and paclitaxel compared with cisplatin and paclitaxel in patients with optimally Resected stage III ovarian cancer: a Gynecologic Oncology Group Study. *Journal of Clinical Oncology* 2003;**21**(17):3194-200.

Rose PG, Nerenstone S, Brady MF, Clarke-Pearson D, Olt G, Rubin SC, et al. Secondary surgical cytoreduction for advanced ovarian carcinoma. *New England Journal of Medicine* 2004;**351**(24):2489-97.

Winter WE, Maxwell III GL, Tian C, Carlson JW, Ozols RF, Rose PG, et al. Prognostic factors for stage III epithelial ovarian cancer: a Gynecologic Oncology Group Study. *Journal of Clinical Oncology* 2007;**25**(24):3621-7.

**Winter 2008** *{published data only}*

McGuire WP, Hoskins WJ, Brady MF, Kucera PR, Partridge EE, Look KY, et al. Cyclophosphamide and cisplatin compared with paclitaxel and cisplatin in patients with stage III and stage IV ovarian cancer. *New England Journal of Medicine* 1996;**334**(1):1-6.

Muggia FM, Braly PS, Brady MF, Sutton G, Niemann HL, Lentz SL, et al. Phase III randomized study of cisplatin versus paclitaxel versus cisplatin and paclitaxel in patients with suboptimal stage III or IV ovarian cancer: a Gynecologic Oncology Group Study. *Journal of Clinical Oncology* 2000;**18**(1):106-15.

Rose PG, Nerenstone S, Brady MF, Clarke-Pearson D, Olt G, Rubin SC, et al. Secondary surgical cytoreduction for advanced ovarian carcinoma. *New England Journal of Medicine* 2004;**351**(24):2489-97.

Spriggs DR, Brady MF, Vaccarello L, Clarke-Pearson DL, Burger RA, Mannel R, et al. Phase III randomized trial of intravenous cisplatin plus a 24- or 96-hour infusion of paclitaxel in epithelial ovarian cancer: a Gynecologic Oncology Group Study. *Journal of Clinical Oncology* 2007;**25**(28):4466-71.

Winter WE, Maxwell GL, Tian C, Sundborg MJ, Rose GS, Rose PG, et al. Tumor residual after surgical cytoreduction in prediction of clinical outcome in stage IV epithelial ovarian cancer: a Gynecologic Oncology Group Study. *Journal of Clinical Oncology* 2008;**26**(1):83-9.

**Zhang 2018** *{published data only}*

Zhang J, Ning L, Zhang A, Xiangxiang B. Potential risk factors associated with prognosis of neoadjuvant chemotherapy followed by interval debulking surgery in stage IIIc-IV high-grade serous ovarian carcinoma patients. *Journal of Obstetrics and Gynaecology Research* 2018;**44**(9):1808-16.

**Zhu 2016** *{published data only}*

Zhu J, Wang H, Cheng-Cheng L, Lu Y, Tang H. The Glasgow Prognostic Score (GPS) is a novel prognostic indicator in advanced epithelial ovarian cancer: a multicenter retrospective study. *Journal of Cancer Research and Clinical Oncology* 2016;**142**(11):2339-45.

## References to studies excluded from this review

**Alberts 1993** *{published data only}*

Alberts DS, Dahlberg S, Green SJ, Garcia D, Hannigan EV, O'Toole R, et al. Analysis of patient age as an independent prognostic factor for survival in a phase III study of cisplatin-cyclophosphamide versus carboplatin-cyclophosphamide in stages III (suboptimal) and IV ovarian cancer. A Southwest Oncology Group study. *Cancer* 1993;**71**(2 Suppl):618-27.

**Alberts 1996** *{published data only}*

Alberts DS, Liu PY, Hannigan EEV, O'Toole RW, Stephen DY, James AF, et al. Intraperitoneal cisplatin plus intravenous cyclophosphamide versus intravenous cisplatin plus intravenous cyclophosphamide for stage III ovarian cancer. *New England Journal of Medicine* 1996;**335**(26):1950-5.

**Alphs 2006** *{published data only}*

Alphs HH, Zahurak ML, Bristow RE, Diaz-Montes TP. Predictors of surgical outcome and survival among elderly women diagnosed with ovarian and primary peritoneal cancer. *Gynecologic Oncology* 2006;**103**(3):1048-53.

**Altman 2012** *{published data only}*

Altman AD, Nelson G, Chu P, Nation J, Ghatage P. Optimal debulking targets in women with advanced stage ovarian cancer: a retrospective study of immediate versus interval debulking surgery. *Journal of Obstetrics and Gynaecology Canada* 2012;**34**(6):558-66.

**Andersen Soegaard 2005** *{published data only}*

Andersen ES, Knudsen A, Svarrer T, Lund B, Nielsen K, Grove A, et al. The results of treatment of epithelial ovarian cancer after centralisation of primary surgery. Results from North Jutland, Denmark. *Gynecologic Oncology* 2005;**99**(3):552-6.

**Anuradha 2016** *{published data only}*

Anuradha S, Donovan PJ, Webb PM, Brand AH, Goh J, Friedlander M, et al. Variations in adjuvant chemotherapy and survival in women with epithelial ovarian cancer–a population-based study. *Acta Oncologica* 2016;**55**(2):226-33.

**Bailey 2006** *{published data only}*

Bailey J, Murdoch J, Anderson R, Weeks J, Foy C. Stage III and IV ovarian cancer in the South West of England: five-year outcome

analysis for cases treated in 1998. *International Journal of Gynecological Cancer* 2006;**1**:25-9.

**Baker 1994** *{published data only}*

Baker TR, Piver MS, Hempling RE. Long term survival by cytoreductive surgery to less than 1 cm, induction weekly cisplatin and monthly cisplatin, doxorubicin, and cyclophosphamide therapy in advanced ovarian adenocarcinoma. *Cancer* 1994;**74**(2):656-63.

**Barda 2004** *{published data only}*

Barda G, Menczer J, Chetrit A, Lubin F, Beck D, Piura B, et al. Comparison between primary peritoneal and epithelial ovarian carcinoma: a population-based study. *American Journal of Obstetrics and Gynecology* 2004;**190**(4):1039-45.

**Benedetti-Panici 1996** *{published data only}*

Benedetti-Panici P, Maneschi F, Scambia G, Cutillo G, Greggi S, Mancuso S. The pelvic retroperitoneal approach in the treatment of advanced ovarian carcinoma. *Obstetrics & Gynecology* 1996;**87**(4):532-8.

**Bertelsen 1990** *{published data only}*

Bertelsen K. Tumor reduction surgery and long-term survival in advanced ovarian cancer: a DACOVA study. *Gynecologic Oncology* 1990;**38**(2):203-9.

**Bertelsen 1993** *{published data only}*

Bertelsen K, Jakobsen A, Strøyer J, Nielsen K, Sandberg E, Andersen JE, et al. A prospective randomized comparison of 6 and 12 cycles of cyclophosphamide, adriamycin, and cisplatin in advanced epithelial ovarian cancer: a Danish Ovarian Study Group Trial (DACOVA). *Gynecologic Oncology* 1993;**49**:30-6.

**Bian 2016** *{published data only}*

Bian C, Yao K, Li L, Yi T, Zhao X. Primary debulking surgery vs. neoadjuvant chemotherapy followed by interval debulking surgery for patients with advanced ovarian cancer. *Archives of Gynecology and Obstetrics* 2016;**293**(1):163-8.

**Bonnefoi 1999** *{published data only}*

Bonnefoi H, A'Hern RP, Fisher C, Macfarlane V, Barton D, Blake P, et al. Natural history of stage IV epithelial ovarian cancer [see comment]. *Journal of Clinical Oncology* 1999;**17**(3):767-75.

**Brinkhuis 1996a** *{published data only}*

Brinkhuis M, Baak JPA, Van Diest PJ, Lund B, Wils J. In Dutch and Danish patients with FIGO III ovarian carcinoma, geographic survival differences are associated with differences in quantitative pathologic features. *International Journal of Gynecological Cancer* 1996;**6**(2):108-14.

**Brinkhuis 1996b** *{published data only}*

Brinkhuis M, Lund B, Meijer GA, Baak JPA. Quantitative pathological variables as prognostic factors for overall survival in Danish patients with FIGO stage III ovarian cancer. *International Journal of Gynecological Cancer* 1996;**6**(3):168-74.

**Bristow 1999** *{published data only}*

Bristow RE, Montz FJ, Lagasse LD, Leuchter RS, Karlan BY. Survival impact of surgical cytoreduction in stage IV epithelial ovarian cancer. *Gynecologic Oncology* 1999;**72**(3):278-87.

**Cai 2007** *{published data only}*

Cai HB, Zhou YF, Chen HZ, Hou HY. The role of bowel surgery with cytoreduction for epithelial ovarian cancer. *Clinical Oncology* 2007;**19**(10):757-62.

**Ceresoli 2018** *{published data only}*

Ceresoli M, Verrengia A, Montori G, Busci L, Coccolini F, Ansaloni L, et al. Effect of cytoreductive surgery and hyperthermic intraperitoneal chemotherapy on relapse pattern in primary epithelial ovarian cancer: a propensity score based case-control study. *Journal of Gynecologic Oncology* 2018;**29**(3):e53.

**Chekman 2015** *{published data only}*

Chekman C, Layoune R, Hocine O, Raissi N, Ferhat HA, Ali Khodja H, et al. An open prospective randomized trial comparing primary complete cytoreduction surgery to debulking surgery after chemotherapy in advanced stage (FIGO's IIIC) ovarian carcinoma. In: 19th International Meeting of the European Society of Gynaecological Oncology, ESGO 2015; 2015 Oct 24-27; Nice, France. 2015:1316. [12185451]

**Clamp 2018** *{published data only}*

Clamp AR, McNeish IA, Dean A, Gallardo-Rincon D, Kim JW, O'Donnell DM, et al. Response to neoadjuvant chemotherapy in ICON8: a GCIG phase III randomised trial evaluating weekly dose-dense chemotherapy integration in first-line epithelial ovarian/fallopian tube/primary peritoneal carcinoma (EOC) treatment. In: Annals of Oncology. Vol. 29. 2018:viii336.

**Colozza 1997** *{published data only}*

Colozza M, Mosconi AM, Gori S, Belsanti V, Basurto C, De Angelis V, et al. Long-term results in patients with advanced epithelial ovarian carcinoma treated with a combination of cisplatin, doxorubicin, and cyclophosphamide. *American Journal of Clinical Oncology* 1997;**20**(5):522-6.

**Conte 1991** *{published data only}*

Conte PF, Bruzzone M, Carnino F, Chiara S, Donadio M, Facchini V, et al. Carboplatin, doxorubicin, and cyclophosphamide versus cisplatin, doxorubicin, and cyclophosphamide: a randomized trial in stage III-IV epithelial ovarian carcinoma. *Journal of Clinical Oncology* 1991;**9**(4):658-63.

**Conte 1996** *{published data only}*

Conte PF, Bruzzone M, Carnino F, Gadducci A, Algeri R, Bellini A, et al. High-dose versus low-dose cisplatin in combination with cyclophosphamide and epidoxorubicin in suboptimal ovarian cancer: a randomized study of the Gruppo Oncologico Nord-Ovest. *Journal of Clinical Oncology* 1996;**14**(2):351-6.

**Crawford 2005** *{published data only}*

Crawford SC, Vasey PA, Paul J, Hay A, Davis JA, Kaye SB. Does aggressive surgery only benefit patients with less advanced ovarian cancer? Results from an international comparison

within the SCOTROC-1 trial. *Journal of Clinical Oncology* 2005;**23**(34):8802-11.

**Creasman 1990** *{published data only}*

Creasman WT, Omura GA, Brady MF, Yordan E, DiSaia PJ, Beecham J. A randomized trial of cyclophosphamide, doxorubicin, and cisplatin with or without bacillus Calmette-Guerin in patients with suboptimal stage III and IV ovarian cancer: a Gynecologic Oncology Group Study. *Gynecologic Oncology* 1990;**39**(3):239-43.

**Cummins 2019** *{published data only}*

Cummins C, Hannah P, Long J, Kumar S, Sundar S, Bramley G. VP100 Ultraradical ovarian cancer surgery comparative clinical effectiveness. *International Journal of Technology Assessment in Health Care* 2019;**35**(S1):97.

**Dao 2016** *{published data only}*

Dao F, Schlappe BA, Tseng J, Lester J, Nick AM, Lutgendorf SK, et al. Characteristics of 10-year survivors of high-grade serous ovarian carcinoma. *Gynecologic Oncology* 2016;**141**:260-3.

**Del Campo 1994** *{published data only}*

Del Campo JM, Felip E, Rubio D, Bermejo B, Colomer R, Zanon V. Long-term survival in advanced ovarian cancer after cytoreduction and chemotherapy treatment. *Gynecologic Oncology* 1994;**53**(1):27-32.

**de Oliviera 1990** *{published data only}*

de Oliveira CF, Lacave AJ, Villani C, Wolff JP, di Re F, Namer M, et al. Randomized comparison of cyclophosphamide, doxorubicin and cisplatin (CAP) versus cyclophosphamide and doxorubicin (CA) for the treatment of advanced ovarian cancer (ADOVCA). A EORTC Gynecological Cancer Cooperative Group Study. *European Journal of Gynaecologic Oncology* 1990;**11**(5):323-30.

**di Re 1996** *{published data only}*

di Re F, Baiocchi G, Fontanelli R, Grosso G, Cobellis L, Raspagliesi F, et al. Systematic pelvic and paraaortic lymphadenectomy for advanced ovarian cancer: prognostic significance of node metastases. *Gynecologic Oncology* 1996;**62**(3):360-5.

**Elgamal 2019** *{published data only}*

Elgamal M, Saha S, Cherry M, Saharan V, Buttar R, Wiese D, et al. Prognostic implications of lymph node metastasis in advanced ovarian cancer: analysis of the National Cancer Database 2006 to 2014. *Journal of Clinical Oncology* 2019;**0**:e17042.

**Fagotti 2020** *{published data only}*

Fagotti A, Ferrandina G, Vizzielli G, Fanfani F, Gallotta V, Chiantera V, et al. Phase III randomised clinical trial comparing primary surgery versus neoadjuvant chemotherapy in advanced epithelial ovarian cancer with high tumour load (SCORPION trial): final analysis of peri-operative outcome. *European Journal of Cancer* 2016;**59**:22-33. [12185453]

Fagotti A, Ferrandina MG, Vizzielli G, Pasciuto T, Fanfani F, Gallotta V, et al. Randomized trial of primary debulking surgery versus neoadjuvant chemotherapy for advanced epithelial ovarian cancer (SCORPION-NCT01461850). *International Journal of Gynecological Cancer* 2020;**30**(11):1657-64. [PMID: 10.1136/ijgc-2020- 001640]

Fagotti A, Vizzielli G, Ferrandina G, Fanfani F, Gallotta V, Chiantera V, et al. Survival analyses from a randomized trial of primary debulking surgery versus neoadjuvant chemotherapy for advanced epithelial ovarian cancer with high tumor load (SCORPION trial). *Journal of Clinical Oncology* 2018;**36**(15):5516. [12185454]

**Gao 2001** *{published data only}*

Gao J, Zheng A, Chen W, Peng Z, Cao Z. A study of prognostic factors of stage IV epithelial ovarian cancer [Chinese]. *Journal of West China University of Medical Sciences* 2001;**32**(2):309-12.

**Gasimli 2016** *{published data only}*

Gasimli K, Braicu EI, Nassir M, Richter R, Babayeva A, Chekerov R, et al. Lymph node involvement pattern and survival differences of FIGO IIIC and FIGO IIIA1 ovarian cancer patients after primary complete tumor debulking surgery: a 10-year retrospective analysis of the tumor bank ovarian cancer network. *Annals of Surgical Oncology* 2016;**23**(4):1279-86.

**Geisler 2004** *{published data only}*

Geisler JP, Tammela JE, Manahan KJ, Geisler HE, Miller GA, Zhou Z, et al. HSP27 in patients with ovarian carcinoma: still an independent prognostic indicator at 60 months follow-up. *European Journal of Gynaecological Oncology* 2004;**25**(2):165-8.

**Gershenson 1989** *{published data only}*

Gershenson DM, Wharton J, Taylor C, Stringer LJ, Edwards CA, Creighton L, et al. Treatment of advanced epithelial ovarian cancer with cisplatin and cyclophosphamide. *Gynecologic Oncology* 1989;**32**(3):336-41.

**Gershenson 1992** *{published data only}*

Gershenson DM, Mitchell MF, Atkinson N, Elvio G, Kavanagh J, Burke M, et al. The effect of prolonged cisplatin-based chemotherapy on progression-free survival in patients with optimal epithelial ovarian cancer: maintenance therapy reconsidered. *Gynecologic Oncology* 1992;**47**(1):7-13.

**Gershenson 1995** *{published data only}*

Gershenson DM, Morris M, Burke TW, Levenback C, Kavanagh JJ, Fromm GL. Combined cisplatin and carboplatin chemotherapy for treatment of advanced epithelial ovarian cancer. *Gynecologic Oncology* 1995;**58**(3):349-55.

**Greggi 2016** *{published data only}*

Greggi S, Falcone F, Carputo R, Raspagliesi F, Scaffa C, Laurelli G, et al. Primary surgical cytoreduction in advanced ovarian cancer: an outcome analysis within the MITO (Multicentre Italian Trials in Ovarian Cancer and Gynecologic Malignancies) Group. *Gynecologic Oncology* 2016;**140**(3):425-9.

**Grem 1991** *{published data only}*

Grem J, O'Dwyer P, Elson P, Simon N, Trump D, Frontiera M, et al. Cisplatin, carboplatin, and cyclophosphamide combination chemotherapy in advanced-stage ovarian carcinoma: an Eastern Cooperative Oncology Group pilot study. *Journal of Clinical Oncology* 1989;**9**(10):1793-800.

**Hainsworth 1990** *{published data only}*

Hainsworth JD, Burnett LS, Jones HW 3rd, Grosh WW, Johnson DH, Greco FA. High-dose cisplatin combination chemotherapy in the treatment of advanced epithelial ovarian carcinoma. *Journal of Clinical Oncology* 1990;**8**(3):502-8.

**Hakes 1992** *{published data only}*

Hakes TB, Chalas E, Hoskins WJ, Jones WB, Markman M, Rubin SC, et al. Randomized prospective trial of 5 versus 10 cycles of cyclophosphamide, doxorubicin, and cisplatin in advanced ovarian carcinoma. *Gynecologic Oncology* 1992;**45**(3):284-9.

**Hamid 2002** *{published data only}*

Hamid D, Rohr S, Baldauf JJ, Ritter J, Kurtz E, Dufour P, et al. Interest of intestinal resection for treatment of advanced ovarian carcinoma [Intérêt des exéréses digestives dans le traitement des cancers écolués de l'ovaire]. *Annales de Chirurgie* 2002;**127**(1):40-7.

**Hardy 1991** *{published data only}*

Hardy JR, Wiltshaw E, Blake PR, Harper P, Slevin M, Perren TJ, et al. Cisplatin and carboplatin in combination for the treatment of stage IV ovarian carcinoma. *Annals of Oncology* 1991;**2**(2):131-6.

**Heitz 2016** *{published data only}*

Heitz F, Harter P, Alesina PF, Walz MK, Lorenz D, Groeben H, et al. Pattern of and reason for postoperative residual disease in patients with advanced ovarian cancer following upfront radical debulking surgery. *Gynecologic Oncology* 2016;**141**(2):264-70.

**Hoskins 1992** *{published data only}*

Hoskins WJ, Bundy BN, Thigpen JT, Omura GA. The influence of cytoreductive surgery on recurrence-free interval and survival in small-volume stage III epithelial ovarian cancer: a Gynecologic Oncology Group study. *Gynecologic Oncology* 1992;**47**(2):159-66.

**Hoskins 1996** *{published data only}*

Hoskins PJ, Swenerton KD, Pike JA, McMurtrie EM, Lee N. "MECCA": a developmental, dose-intensive, non-cross-resistant platinum-based chemotherapy for advanced ovarian cancer. *Gynecologic Oncology* 1996;**63**(3):345-51.

**Hoskins 1997** *{published data only}*

Hoskins WJ, McGuire WP, Brady MF, Kucera PR, Partridge EE, Look KY, et al. Combination paclitaxel (Taxol(R))-cisplatin vs cyclophosphamide-cisplatin as primary therapy in patients with suboptimally debulked advanced ovarian cancer. *International Journal of Gynecological Cancer* 1997;**7**:9-13.

**Itamochi 2002** *{published data only}*

Itamochi H, Kigawa J, Sugiyama T, Kikuchi Y, Suzuki M, Terakawa N. Low proliferation activity may be associated with chemoresistance in clear cell carcinoma of the ovary. *Obstetrics & Gynecology* 2002;**100**(2):281-7.

**Kaern 2005** *{published data only}*

Kaern J, Aghmesheh M, Nesland JM, Danielsen HE, Sandstad B, Friedlander M, et al. Prognostic factors in ovarian carcinoma stage III patients. Can biomarkers improve the prediction of short- and long-term survivors? *International Journal of Gynecological Cancer* 2005;**15**(6):1014-22.

**Kehoe 2015** *{published data only}*

Kehoe S, Hook J, Nankivell M, Jayson GC, Kitchener H, Lopes T, et al. Primary chemotherapy versus primary surgery for newly diagnosed advanced ovarian cancer (CHORUS): an open-label, randomised, controlled, non-inferiority trial. *Lancet* 2015;**386**(9990):249-57. [12185455]

Kehoe S, Hook J, Nankivell M. Chemotherapy or upfront surgery for newly diagnosed advanced ovarian cancer: results from the MRC CHORUS trial. *Journal of Clinical Oncology* 2013;**31 Suppl**(15):Abstract 5500. [12185456]

Kehoe S, Wheeler S. CHORUS (Chemotherapy or Upfront Surgery). A randomised feasibility trial to determine the impact of timing of surgery and chemotherapy in newly diagnosed patients with advanced epithelial ovarian, primary peritoneal or fallopian tube carcinoma. www.ctu.mrc.ac.uk/plugins/StudyDisplay/protocols/CHORUS protocol Version 2.0 - 05 June 2008.pdf; and http://www.ctu.mrc.ac.uk/research_areas/study_details.aspx?s=9 (accessed 18/6/2012). [12185457]

Law K, Murray C, Kehoe S. CHORUS - a randomised study to determine the impact of timing of surgery and chemotherapy in newly diagnosed patients with advanced epithelial ovarian, primary peritoneal or fallopian tube carcinoma. In: Annual Meeting of the British Gynaecological Cancer Society; 2006 Nov 30-Dec 1; Manchester, UK. 2006:90. [12185459]

Vergote I, Coens C, Nankivell M, Kristensen GB, Parmar MK, Ehlen T, et al. Neoadjuvant chemotherapy versus debulking surgery in advanced tubo-ovarian cancers: pooled analysis of individual patient data from the EORTC 55971 and CHORUS trials. *Lancet Oncology* 2018;**19**:1680-7. [12185460]

**Kessous 2017** *{published data only}*

Kessous R, Laskov I, Abitbol J, Bitharas J, Yasmeen A, Salvador S, et al. Clinical outcome of neoadjuvant chemotherapy for advanced ovarian cancer. *Gynecologic Oncology* 2017;**144**(3):474-9.

**Keyver-Paik 2016** *{published data only}*

Keyver-Paik MD, Abramian A, Domröse C, Döser A, Höller T, Friedrich M, et al. Integrated care in ovarian cancer "IgV Ovar": results of a German pilot for higher quality in treatment of ovarian cancer. *Journal of Cancer Research and Clinical Oncology* 2016;**142**(2):481-7.

**Kirmani 1994** *{published data only}*

Kirmani S, Braly PS, McClay EF, Saltzstein SL, Plaxe SC, Kim S et al. A comparison of intravenous versus intraperitoneal chemotherapy for the initial treatment of ovarian cancer. *Gynecologic Oncology* 1994;**54**(3):338-44.

**Kristensen 1995** *{published data only}*

Kristensen GB, Baekelandt M, Vergote IB, Trope C. A phase II study of carboplatin and hexamethylmelamine as induction chemotherapy in advanced epithelial ovarian carcinoma. *European Journal of Cancer* 1995;**31**(11):1778-80.

**Le 1997** *{published data only}*

Le T, Krepart GV, Lotocki R J, Heywood MS. Does debulking surgery improve survival in biologically aggressive ovarian carcinoma? *Gynecologic Oncology* 1997;**67**(2):208-14.

**Lee 2018** *{published data only}*

Lee YY, Lee JW, Lu L, Xu W, Kollara A, Brown T, et al. Impact of interval from primary cytoreductive surgery to initiation of adjuvant chemotherapy in advanced epithelial ovarian cancer. *International Journal of Gynaecology and Obstetrics* 2018;**143**(3):325-32.

**Loizzi 2016** *{published data only}*

Fumarulo VV, Loizzi V, Cormio G, Murgia F, Vecchio V, Minicucci V, et al. Neoadjuvant chemotherapy in advanced ovarian cancer: a single institution experience and a review of the literature. In: International Journal of Gynecological Cancer. Vol. 26. 2016:1-1193.

* Loizzi V, Leone L, Camporeale A, Resta L, Selvaggi L, Cicinelli E, et al. Neoadjuvant chemotherapy in advanced ovarian cancer: a single-institution experience and a review of the literature. *Oncology* 2016;**91**(4):211-6.

**Lorusso 1998** *{published data only}*

Lorusso V, Leone B, Di Vagno G, Manzione L, Palmeri S, Vallejo C. Combined carboplatin plus ifosfamide and cisplatin in patients with advanced ovarian carcinoma. A phase I-II study. *Gynecologic Oncology* 1998;**68**(2):172-7.

**Malik 1998** *{published data only}*

Malik IM, Khan ZK, Khan WA, Hussain M, Moid I, Rizvi J. Continuous infusion of ifosfamide and cisplatin as first-line therapy of patients with suboptimally debulked stage III/IV epithelial ovarian cancer. *International Journal of Gynecological Cancer* 1998;**8**(2):138-43.

**Marchetti 1993** *{published data only}*

Marchetti DL, Lele SB, Priore RL, McPhee ME, Hreshchyshyn MM. Treatment of advanced ovarian carcinoma in the elderly. *Gynecologic Oncology* 1993;**49**(1):86-91.

**McGuire 1996** *{published data only}*

McGuire WP, Hoskins WJ, Brady MF, Kucera PR, Partridge EE, Look KY, et al. Cyclophosphamide and cisplatin compared with paclitaxel and cisplatin in patients with stage III and stage IV ovarian cancer. *New England Journal of Medicine* 1996;**334**(1):1-6.

**Michaan 2018** *{published data only}*

Michaan N, Chong WY, Han NY, Lim MC, Park SY. Prognostic value of pathologic chemotherapy response score in patients with ovarian cancer after neoadjuvant chemotherapy. *International Journal of Gynecologic Cancer* 2018;**28**(9):1676-82.

**Ngan 1989** *{published data only}*

Ngan HY, Choo YC, Cheung M, Wong LC, Ma HK, Collins R. A randomized study of high-dose versus low-dose cis-platinum combined with cyclophosphamide in the treatment of advanced ovarian cancer. *Chemotherapy* 1989;**35**(3):221-7.

**Omura 1989** *{published data only}*

Omura GA, Bundy BN, Berek JS, Curry S, Delgado G Mortel R. Randomized trial of cyclophosphamide plus cisplatin with or without doxorubicin in ovarian carcinoma: a Gynecologic Oncology Group Study. *Journal of Clinical Oncology* 1989;**7**(4):457-65.

**Onda 2020** *{published data only}*

Onda T, Matsumoto K, Shibata T, Sato A, Fukuda H, Konishi I, et al. Phase III trial of upfront debulking surgery versus neoadjuvant chemotherapy for stage III/IV ovarian, tubal and peritoneal cancers: Japan Clinical Oncology Group Study JCOG0602. *Japanese Journal of Clinical Oncology* 2008;**38**(1):74-7. [12185462]

* Onda T, Satoh T, Ogawa G, Saito T, Kasamatsu T, Nakanishi T, et al, Japan Clinical Oncology Group. Comparison of survival between primary debulking surgery and neoadjuvant chemotherapy for stage III/IV ovarian, tubal and peritoneal cancers in phase III randomised trial. *European Journal of Cancer* 2020;**130**:114-25.

Onda T, Satoh T, Saito T, Kasamatsu T, Nakanishi T, Nakamura K, et al. Comparison of treatment invasiveness between upfront debulking surgery versus interval debulking surgery following neoadjuvant chemotherapy for stage III/IV ovarian, tubal, and peritoneal cancers in a phase III randomised trial: Japan Clinical Oncology Group Study JCOG0602. *European Journal of Cancer* 2016;**64**:22-31. [12185463]

Onda T, Satoh T, Saito T, Kasamatsu T, Nakanishi T, Takehara K, et al. Comparison of survival between upfront primary debulking surgery versus neoadjuvant chemotherapy for stage III/IV ovarian, tubal and peritoneal cancers in phase III randomized trial: JCOG0602. Journal of Clinical Oncology 2018;**36**(15 Suppl). [12185464]

**Palmer 1992** *{published data only}*

Palmer MC, Shepert E, Schepansky A, MacLean GD. Novel, dose intensive, single-agent cisplatin in the first-line management of advanced stage ovarian cancer. *International Journal of Gynecological Cancer* 1992;**2**(6):301-6.

**Piver 1991** *{published data only}*

River MS, Fanning J, Sprance HE. Five-year survival for cisplatin-based chemotherapy versus single-agent melphalan in patients with advanced ovarian cancer and optimal debulking surgery. *Journal of Surgical Oncology* 1991;**48**(1):39-44.

**Raspagliesi 2018** *{published data only}*

Raspagliesi F, Bogani G, Matteucci L, Casarin J, Sabatucci I, Tamberi S, et al. Surgical efforts might mitigate difference in response to neoadjuvant chemotherapy in stage IIIC–IV unresectable ovarian cancer: a case-control multi-institutional study. *International Journal of Gynecologic Cancer* 2018;**28**(9):1706-13.

**Redman 1986** *{published data only}*

Redman JR, Petroni GR, Saigo PE, Geller NL, Hakes TB. Prognostic factors in advanced ovarian carcinoma. *Journal of Clinical Oncology* 1986;**4**(4):515-23.

**Risum 2012** *{published data only}*

Risum S, Loft A, Engelholm SA, Høgdall E, Berthelsen AK, Nedergaard L, et al. Positron emission tomography/computed tomography predictors of overall survival in stage IIIC/IV ovarian cancer. *International Journal of Gynecologic Cancer* 2012;**22**(7):1163-9.

**Rodriguez 2013** *{published data only}*

Rodriguez N, Miller A, Richard SD, Rungruang B, Hamilton CA, Bookman MA, et al. Upper abdominal procedures in advanced stage ovarian or primary peritoneal carcinoma patients with minimal or no gross residual disease: an analysis of Gynecologic Oncology Group (GOG). *Gynecologic Oncology* 2013;**130**(3):487-92.

**Rose 2004** *{published data only}*

Rose PG, Nerenstone S, Brady MF, Clarke-Pearson D, Olt G, Rubin SC, et al and the Gynecologic Oncology Group. Secondary surgical cytoreduction for advanced ovarian carcinoma. *New England Journal of Medicine* 2004;**351**(24):89-97.

**Ruscito 2016** *{published data only}*

Ruscito I, Gasparri ML, Marchetti C, Crispino S, La Russa C, Petriglia G, et al. Obesity is a key prognostic factor in stage IIIC-IV ovarian cancer diagnosed prior to 65 years of age: a 10-year survival analysis. *Gynecologic Oncology* 2016;**141**:204.

**Rutten 2014** *{published data only}*

* Rutten MJ, Boldingh JH, Schuit E, Trum H, Van Driel W, Mol BW, et al. Development and internal validation of a prognostic model for survival after debulking surgery for epithelial ovarian cancer. *Gynecologic Oncology* 2014;**35**(1):13-8.

Rutten MJ, Boldingh JH, Schuit E, Trum H, Van Driel W, Mol BW, et al. Prognostic model for survival of epithelial ovarian cancer patients treated with primary or interval debulking surgery. *InInternational Journal Of Gynecological Cancer* 2013;**23**:8.

**Salani 2007** *{published data only}*

Salani R, Zahurak ML, Santillan A, Giuntoli RL, Bristow RE. Survival impact of multiple bowel resections in patients undergoing primary cytoreductive surgery for advanced ovarian cancer: a case-control study. *Gynecologic Oncology* 2007;**107**(3):495-9.

**Sessa 1991** *{published data only}*

Sessa C, Colombo N, Bolis G, Marsoni S, Mangioni C. Randomized comparison of hexamethylmelamine, adriamycin, cyclophosphamide (HAC) vs. cisplatin, adriamycin, cyclophosphamide (PAC) in advanced ovarian cancer: long-term results. *Cancer Treatment Reviews* 1991;**18**(18 Suppl A):37-46.

**Shapiro 1998** *{published data only}*

Shapiro JD, Rothenberg ML, Sarosy GA, Steinberg SM, Adamo DO, Reed E, et al. Dose intensive combination platinum and cyclophosphamide in the treatment of patients with advanced untreated epithelial ovarian cancer. *Cancer* 1998;**83**(9):1980-8.

**Shinozuka 1999** *{published data only}*

Shinozuka T, Miyamoto T, Muramatsu T, Hirasawa T, Murakami M, Makino T, et al. High dose chemotherapy with autologous stem cell support in the treatment of patients with ovarian carcinoma: long term results for 105 patients. *Cancer* 1999;**85**(7):1555-64.

**Sioulas 2017** *{published data only}*

Sioulas VD, Schiavone MB, Kadouri D, Zivanovic O, Roche KL, O'Cearbhaill R, et al. Optimal primary management of bulky stage IIIC ovarian, fallopian tube and peritoneal carcinoma: are the only options complete gross resection at primary debulking surgery or neoadjuvant chemotherapy? *Gynecologic Oncology* 2017;**145**(1):15-20.

**Skarlos 1996** *{published data only}*

Skarlos DV, Aravantinos G, Kosmidis P, Pavlidis N, Gennatas K, Beer M, et al. Carboplatin alone compared with its combination with epirubicin and cyclophosphamide in untreated advanced epithelial ovarian cancer: a Hellenic Co-operative Oncology Group study. *European Journal of Cancer* 1996;**32**(3):421-8.

**Smits 2015** *{published data only}*

Smits A, Smits E, Lopes A, Das N, Hughes G, Talaat A, et al. Body mass index, physical activity and quality of life of ovarian cancer survivors: time to get moving? *Gynecologic Oncology* 2015;**139**(1):148-54.

**Solmaz 2015** *{published data only}*

Solmaz U, Mat E, Dereli ML, Turan V, Peker N, Tosun G, et al. Does neoadjuvant chemotherapy plus cytoreductive surgery improve survival rates in patients with advanced epithelial ovarian cancer compared with cytoreductive surgery alone? *Journal of Balkan Union of Oncology* 2015;**20**(3):847-54.

**Son 2017** *{published data only}*

Son JH, Kong TW, Paek J, Song KH, Chang SJ, Ryu HS. Clinical characteristics and prognostic inflection points among long-term survivors of advanced epithelial ovarian cancer. *International Journal of Gynaecology and Obstetrics* 2017;**193**(3):352-7.

**Stewart 2015** *{published data only}*

Stewart JM, Tone AA, Bernardini MQ, Ferguson SE, Dodge J, Laframboise S, et al. Surgical factors do not impact survival in high-grade serous ovarian carcinoma patients treated with neoadjuvant chemotherapy. *Gynecologic Oncology* 2015;**137**(Suppl 1):43.

**Stewart 2016** *{published data only}*

Stewart JM, Tone AA, Jiang H, Bernardini MQ, Ferguson S, Laframboise S, et al. The optimal time for surgery in women with serous ovarian cancer. *Canadian Journal of Surgery* 2016;**59**(4):223-32.

**Strauss 1996** *{published data only}*

Strauss G, Lund B, Hansen M, Hansen OP, Hansen HH. Combined high-dose platinum and etoposide in previously untreated ovarian cancer patients with residual disease. *International Journal of Gynecological Cancer* 1996;**6**(5):410-4.

**Suidan 2015** *{published data only}*

Suidan RS, Zhou Q, Iasonos A, O'Cearbhaill RE, Chi DS, Roche KC, et al. Prognostic significance of the number of postoperative intraperitoneal chemotherapy cycles for patients with advanced epithelial ovarian cancer. *International Journal of Gynecologic Cancer* 2015;**25**(4):599-606.

**Sun 2000** *{published data only}*

Sun T, Feng Y, ZhuY, Zheng Y. Therapeutic strategy in the management of stage II - IV epithelial ovarian carcinoma. *Chinese Medical Journal* 2000;**113**(7):625-7.

**Sutton 1989** *{published data only}*

Sutton GP, Stehman FB, Einhorn LH, Roth LM, Blessing JA, Ehrlich CE. Ten-year follow-up of patients receiving cisplatin, doxorubicin, and cyclophosphamide chemotherapy for advanced epithelial ovarian carcinoma. *Journal of Clinical Oncology* 1989;**7**:223-9.

**Takano 2006** *{published data only}*

Takano MJ, Kikuchi Y, Yaegashi N, Suzuki M, Tsuda H, Sagae S, et al. Adjuvant chemotherapy with irinotecan hydrochloride and cisplatin for clear cell carcinoma of the ovary. *Oncology Reports* 2006;**16**(6):1301-6.

**Takano 2007** *{published data only}*

Takano M, Sugiyama T, Yaegashi N, Suzuki M, Sagae S, Udagawa, et al. Progression-free survival and overall survival of patients with clear cell carcinoma of the ovary treated with paclitaxel-carboplatin or irinotecan-cisplatin: retrospective analysis. *International Journal of Clinical Oncology* 2007;**12**(4):256-60.

**Tay 1996** *{published data only}*

Tay SK, Chang TC. An evaluation of the policy of routine treatment of advanced epithelial ovarian carcinoma by debulking surgery and combined platinum-cyclophosphamide chemotherapy in Singapore. *International Journal of Gynecological Cancer* 1996;**6**(1):44-8.

**Taylor 1994** *{published data only}*

Taylor AE, Wiltshaw E, Gore ME, Fryatt I, Fisher C. Long-term follow-up of the first randomized study of cisplatin versus carboplatin for advanced epithelial ovarian cancer. *Journal Of Clinical Oncology* 1994;**12**(10):2066-70.

**Tingulstad 2003** *{published data only}*

Tingulstad S, Skjeldestad FE, Hagen B. The effect of centralization of primary surgery on survival in ovarian cancer patients. *Obstetrics and Gynecology* 2003;**102**(3):499-505.

**Todo 2003** *{published data only}*

Todo Y, Sakuragi N, Oikawa M, Negishi H, Yamamoto R, Yoshiaki K, et al. Cytoreductive surgery combined with organ resection for advanced ovarian carcinoma. *International Journal of Clinical Oncology* 2003;**8**(2):90-6.

**Trhlík 2013** *{published data only}*

Trhlík M, Soumarová R, Bartoš P, Koláček J, Horová I, Těžká M. Neoadjuvant chemotherapy for primary advanced ovarian cancer. In: Biomedical Papers. Vol. 157. 2013:S98.

**Uyar 2005** *{published data only}*

Uyar D, Frasure HE, Markman M, Von Gruenigen VE. Treatment patterns by decade of life in elderly women ([greater-than or equal to]70 years of age) with ovarian cancer. *Gynecologic Oncology* 2005;**98**(3):403-8.

**Vallejos 1997** *{published data only}*

Vallejos C, Solidoro A, Castellano C, Barriga C, Galdos O, Casanova R, et al. Ifosfamide plus cisplatin as primary chemotherapy of advanced ovarian cancer. *Gynecologic Oncology* 1997;**67**(2):168-71.

**Van Der Burg 1996** *{published data only}*

Van Der Burg MEL, Van Lent M, Buyse M, Kobierska A, Maggioni A, Favalli G, et al. The role of intervention debulking surgery in advanced epithelial ovarian cancer: an EORTC Gynecological Cancer Cooperative Group study. *International Journal of Gynecological Cancer* 1996;**6**:30-8.

**Van Driel 2017** *{published data only}*

Van Driel W, Sikorska K, Schagen van Leeuwen J, Schreuder H, Hermans R, de Hingh I, et al. A phase 3 trial of hyperthermic intraperitoneal chemotherapy (HIPEC) for ovarian cancer. *Journal of Clinical Oncology* 2017;**35**(15 Suppl):5519.

**van Vliet 2015** *{published data only}*

van Vliet MML, Schreuder HWR, Pasker-de Jong PCM, Duk MJ. Centralisation of epithelial ovarian cancer surgery: results on survival from a peripheral teaching hospital. *European Journal of Obstetrics, Gynecology, and Reproductive Biology* 2015;**192**:72-8.

**Vergote 2010** *{published data only}*

Greimel E, Kristensen G, Vergote I, Hoskins P, van der Burg ME, Casado Herraez A, et al. Quality of life in advanced ovarian cancer patients: a randomized phase III study comparing upfront debulking surgery versus neo-adjuvant chemotherapy. *International Journal of Gynaecological Cancer* 2011;**21**:S620. [12185466]

Greimel E, Kristensen GB, van der Burg ME, Coronado P, Rustin G, del Rio AS, et al. Quality of life of advanced ovarian cancer patients in the randomized phase III study comparing primary debulking surgery versus neo-adjuvant chemotherapy. *Gynecologic Oncology* 2013;**131**(2):437-44. [12185467]

Vergote I, Amant F, Kristensen G, Ehlen T, Reed NS, Casado A. Primary surgery or neoadjuvant chemotherapy followed by interval debulking surgery in advanced ovarian cancer. *European Journal of Cancer* 2011;**47**(Suppl 3):S88-91. [12185468]

Vergote I, Coens C, Nankivell M, Kristensen GB, Parmar MK, Ehlen T, et al. Neoadjuvant chemotherapy versus debulking surgery in advanced tubo-ovarian cancers: pooled analysis of individual patient data from the EORTC 55971 and CHORUS trials. *Lancet Oncology* 2018;**19**:1680-7. [12185469]

Vergote I, Pecorelli S, Stuart G. Intergroup Study (EORTC 55971/NCIC OV13). Randomized phase III study comparing upfront debulking surgery versus neo-adjuvant chemotherapy in patients with stage IIIc or IV epithelial ovarian carcinoma. Trial

protocol: http://www.cancer.gov/clinicaltrials/EORTC-55971 2003 (accessed 17 June 2012). [12185470]

Vergote I, Tropé CG, Amant F, Ehlen T, Reed NS, Casado A. Neoadjuvant chemotherapy is the better treatment option in some patients with stage IIIc to IV ovarian cancer. *Journal of Clinical Oncology* 2011;**29**(31):4076-8. [12185471]

Vergote I, Trope CG, Amant F, Kristensen GB, Ehlen T, Johnson N, et al. Neoadjuvant chemotherapy or primary surgery in stage IIIC or IV ovarian cancer. *New England Journal of Medicine* 2010;**363**(10):943-53. [12185472] [Incl. Supplementary Appendix and Protocol]

Verleye L, Ottevanger PB, Kristensen GB, Ehlen T, Johnson N, van der Burg ME, et al. Quality of pathology reports for advanced ovarian cancer: are we missing essential information? An audit of 479 pathology reports from the EORTC-GCG 55971/NCIC-CTG OV13 neoadjuvant trial. *European Journal of Cancer* 2011;**47**(1):57-64. [12185473]

**Vergote 2018** *{published data only}*

Vergote I, Coens C, Nankivell M, Kristensen G, Parmer M, Ehlen T, et al. Meta-analysis of the randomized EORTC and chorus neoadjuvant versus primary debulking trials in advanced tubo-ovarian cancer. *International Journal of Gynecological Cancer* 2016;**26**(Suppl 3):24-5.

Vergote I, Coens C, Nankivell M, Kristensen GB, Parmar MK, Ehlen T, et al. Neoadjuvant chemotherapy versus debulking surgery in advanced tubo-ovarian cancers: pooled analysis of individual patient data from the EORTC 55971 and CHORUS trials. *Lancet Oncology* 2018;**19**(12):1680-7.

* Vergote I, Coens C, Nankivell M, Kristensen GB, Parmar MK, Ehlen T, et al. Neoadjuvant chemotherapy versus debulking surgery in advanced tubo-ovarian cancers: pooled analysis of individual patient data from the EORTC 55971 and CHORUS trials. *Lancet Oncology* 2018;**19**(12):1680-7.

**Vidal 2016** *{published data only}*

Vidal F, Guerby P, Luyckx M, Haddad P, Stoeckle E, Morice P, et al. Are early relapses in advanced-stage ovarian cancer doomed to a poor prognosis? *PlOS One* 2016;**11**(1):e0147787. [DOI: 10.1371/journal.pone.0147787]

**Wadler 1996** *{published data only}*

Wadler S, Yeap B, Vogl S, Carbone P. Randomized trial of initial therapy with melphalan versus cisplatin-based combination chemotherapy in patients with advanced ovarian carcinoma: initial and long term results - Eastern Cooperative Oncology Group study E2878. *Cancer* 1996;**77**(4):733-42.

**Wallace 2017** *{published data only}*

Wallace S, Kumar A, McGree M, Weaver A, Mariani A, Langstraat C, et al. Efforts at maximal cytoreduction improve survival in ovarian cancer patients, even when complete gross resection is not feasible. *Gynecologic Oncology* 2017;**145**(1):21-6.

**Warwick 1995** *{published data only}*

Warwick J, Kehoe S, Earl H, Luesley D, Redman C, Chan KK. Long-term follow-up of patients with advanced ovarian cancer treated in randomised clinical trials. *British Journal of Cancer* 1995;**72**(6):1513-7.

**Willemse 1992** *{published data only}*

Willemse PHB, De Vries EGE, Kloppenburg M, Fontein DL, Aalders JG, Boonstra H, et al. Carboplatin with cyclophosphamide in patients with advanced ovarian cancer: an efficacy and quality-adjusted survival analysis. *International Journal of Gynecological Cancer* 1992;**2**(5):236-43.

**Wils 1990** *{published data only}*

Wils JA. Long-term follow-up of patients with advanced ovarian carcinoma treated with debulking surgery and chemotherapy consisting of cisplatin, doxorubicin, and cyclophosphamide. Gynecologic Oncology Group of the Comprehensive Cancer Center. *Oncology* 1990;**47**(2):115-20.

**Wimberger 2007** *{published data only}*

Wimberger P, Lehmann N, Kimmig R, Burges A, Meier W, Du Bois A. Prognostic factors for complete debulking in advanced ovarian cancer and its impact on survival. An exploratory analysis of a prospectively randomized phase III study of the Arbeitsgemeinschaft Gynaekologische Onkologie Ovarian Cancer Study Group (AGO-OVAR). *Gynecologic Oncology* 2007;**106**(1):69-74.

**Yamamoto 2007** *{published data only}*

Yamamoto S, Tsuda H, Yoshikawa T, Kudoh K, Kita T, Furuya K, et al. Clear cell adenocarcinoma associated with clear cell adenofibromatous components: a subgroup of ovarian clear cell adenocarcinoma with distinct clinicopathologic characteristics. *American Journal of Surgical Pathology* 2007;**31**(7):999-1006.

**Zang 1999** *{published data only}*

Zang RY, Zhang ZY, Cai SM, Li ZT, Chen J, Tang MQ, et al. Cytoreductive surgery for stage IV epithelial ovarian cancer. *Journal of Experimental & Clinical Cancer Research* 1999;**18**(4):449-54.

**Zhang 2015** *{published data only}*

Zhang LY, Li PL, Wang TZ, Zhang XC. Prognostic values of 5-hmC, 5-mC and TET2 in epithelial ovarian cancer. *Archives of Gynecology and Obstetrics* 2015;**292**(4):891-7.

## Additional references

**Aldin 2020**

Aldin A, Umlauff L, Estcourt LJ, Collins G, Moons KG, Engert A, et al. Interim PET-results for prognosis in adults with Hodgkin lymphoma: a systematic review and meta-analysis of prognostic factor studies. *Cochrane Database of Systematic Reviews* 2020, Issue 1. Art. No: CD012643. [DOI: 10.1002/14651858.CD012643.pub3]

**Aletti 2007**

Aletti GD, Dowdy SC, Podratz KC, Cliby WA. Relationship among surgical complexity, short-term morbidity, and overall survival in primary surgery for advanced ovarian cancer. *American Journal of Obstetrics and Gynecology* 2007;**187**(6):676.

**Armstrong 2006**

Armstrong DK, Bundy B, Wenzel L, Huang HQ, Baergen R, Lele S, et al. Intraperitoneal cisplatin and paclitaxel in ovarian cancer. *New England Journal of Medicine* 2006;**354**(1):34-43.

**Berek 2018**

Berek JS, Kehoe ST, Kumar L, Friedlander M. Cancer of the ovary, fallopian tube, and peritoneum. *International Journal of Gynaecology and Obstetrics* 2018;**143**:59-78.

**BGCS 2017**

BGCS. https://www.bgcs.org.uk/wp-content/uploads/2019/05/BGCS-Guidelines-Ovarian-Guidelines-2017.pdf 2017.

**Bookman 2009**

Bookman MA, Brady MF, McGuire WP, Harper PG, Alberts DS, Friedlander M, et al. Evaluation of new platinum-based treatment regimens in advanced-stage ovarian cancer: a phase III trial of the Gynecologic Cancer Intergroup. *Journal of Clinical Oncology* 2009;**27**(9):1419-25.

**Borley 2012**

Borley J, Wilhelm-Benartzi C, Brown R, Ghaem-Maghami S. Does tumour biology determine surgical success in the treatment of epithelial ovarian cancer? A systematic literature review. *British Journal of Cancer* 2012;**107**(7):1069-74.

**Bristow 2002**

Bristow RE, Tomacruz RS, Armstrong DK, Trimble EL, Montz FJ. Survival effect of maximal cytoreductive surgery for advanced ovarian carcinoma during the platinum era: a meta-analysis. *Journal of Clinical Oncology* 2002;**20**(5):1248-59.

**Bristow 2009**

Bristow RE, Zahurak ML, Diaz-Montes TP, Giuntoli RL, Armstrong DK. Impact of surgeon and hospital ovarian cancer surgical case volume on in hospital mortality and related short-term outcomes. *Gynecologic Oncology* 2009;**115**(3):334-8.

**Chang 2013**

Chang SJ, Hodeib M, Chang J, Bristow RE. Survival impact of complete cytoreduction to no gross residual disease for advanced-stage ovarian cancer: a meta-analysis. *Gynecologic Oncology* 2013;**130**(3):493–8.

**Coleridge 2021**

Coleridge SL, Bryant A, Kehoe S, Morrison J. Neoadjuvant chemotherapy before surgery versus surgery followed by chemotherapy for initial treatment in advanced ovarian epithelial cancer. *Cochrane Database of Systematic Reviews* 2021, Issue 7. Art. No: CD005343. [DOI: 10.1002/14651858.CD005343]

**Colombo 2019**

Colombo N, Sessa C, du Bois A, Ledermann J, McCluggage WG, McNeish I, et al. ESMO-ESGO consensus conference recommendations on ovarian cancer: pathology and molecular biology, early and advanced stages, borderline tumours and recurrent disease. *Annals of Oncology* 2019;**30**(5):672–705.

**Covens 2000**

Covens AL. A critique of surgical cytoreduction in advanced ovarian cancer. *Gynecologic Oncology* 2000;**78**:269-74.

**Debray 2018**

Debray TPA, Moons KGM, Riley RD. Detecting small-study effects and funnel plot asymmetry in meta-analysis of survival data: a comparison of new and existing tests. *Research Synthesis Methods* 2018;**9**(1):41-50.

**Deeks 2001**

Deeks JJ, Altman DG, Bradburn MJ. Statistical methods for examining heterogeneity and combining results from several studies in meta-analysis. In: Egger M, Davey Smith G, Altman DG, editors(s). Systematic Reviews in Health Care: Meta-Analysis in Context (2nd edition). London: BMJ Publication Group, 2001.

**DerSimonian 1986**

DerSimonian R, Laird N. Meta-analysis in clinical trials. *Controlled Clinical Trials* 1986;**7**:177-88.

**du Bois 2009**

du Bois A, Reuss A, Pujade-Lauraine E, Harter P, Ray-Coquard I, Pfisterer J. Role of surgical outcome as prognostic factor in advanced epithelial ovarian cancer: a combined exploratory analysis of 3 prospectively randomized phase 3 multicenter trials: by the Arbeitsgemeinschaft Gynaekologische Onkologie Studiengruppe Ovarialkarzinom (AGO-OVAR) and the Groupe d'Investigateurs Nationaux Pour les Etudes des Cancers de l'Ovaire (GINECO). *Cancer* 2009;**115**(6):1234–44.

**Elzarkaa 2018**

Elzarkaa AA, Shaalan W, Elemam D, Mansour H, Melis M, Malik E, et al. Peritoneal cancer index as a predictor of survival in advanced stage serous epithelial ovarian cancer: a prospective study. *Journal of Gynecologic Oncology* 2018;**29**(4):e47.

**EUROCARE 2015**

Sant M, Chirlaque Lopez MD, Agresti R, Sánchez Pérez MJ, Holleczek B, Bielska Lasota M, et al, EUROCARE-5 Working Group. Survival of women with cancers of breast and genital organs in Europe 1999-2007: results of the EUROCARE-5 study. *European Journal of Cancer* 2015;**51**(15):2191-205.

**Fader 2007**

Fader AN, Rose PG. Role of surgery in ovarian carcinoma. *Journal of Clinical Oncology* 2007;**25**(20):2873-83.

**Falconer 2020**

Falconer H, Joneborg U, Krawiec K, Palsdottir K, Bottai M, Salehi S. Ultra-radical upfront surgery does not improve survival in women with advanced epithelial ovarian cancer; a natural experiment in a complete population. *Gynecologic Oncology* 2020;**159**(1):58-65.

**Ferron 2019**

Ferron G, De Rauglaudre G, Chevalier A, Combe P, Joly F, Lortholary A, et al. Impact of adding nintedanib to neoadjuvant chemotherapy (NACT) for advanced epithelial ovarian cancer (EOC) patients: the CHIVA double-blind randomized phase II GINECO study. *Journal of Clinical Oncology* 2019;**37**(15 Suppl):5512.

**Foroutan 2020**

Foroutan F, Guyatt G, Zuk V, Vandvik PO, Alba AC, Mustafa R, et al. GRADE Guidelines 28: Use of GRADE for the assessment of evidence about prognostic factors: rating certainty in identification of groups of patients with different absolute risks. *Journal of Clinical Epidemiology* 2020;**121**:62-70. [DOI: 10.1016/j.jclinepi.2019.12.023; PubMed: 31982539]

**Ghirardi 2020**

Ghirardi V, Moruzzi MC, Bizzarri N, Vargiu V, D'Indinosante M, Garganese G, et al. Minimal residual disease at primary debulking surgery versus complete tumor resection at interval debulking surgery in advanced epithelial ovarian cancer: a survival analysis. *Gynecologic Oncology* 2020;**157**(1):209-13.

**Girling 1996**

Girling JC, Soutter WP. Cytoreductive surgery in the primary management of advanced epithelial ovarian carcinoma: a topic for debate. *International Journal of Gynecological Cancer* 1996;**1**:81-4.

**GLOBOCAN 2018**

(GLOBOCAN 2018) Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: A Cancer Journal for Clinicians. Available from: http://gco.iarc.fr/today/home 2018;**68**(394-424). [DOI: 10.3322/caac.21492]

**Grabowski 2016**

Grabowski JP, Harter P, Heitz F, Pujade-Lauraine E, Reuss A, Kristensen G, et al. Operability and chemotherapy responsiveness in advanced low-grade serous ovarian cancer. An analysis of the AGO Study Group metadatabase. *Gynecologic Oncology* 2016;**140**(3):457-62.

**GRADE Working Group**

GRADE Working Group. Grading quality of evidence and strength of recommendations. *BMJ* 2004;**328**:1490-4.

**Griffiths 1975**

Griffiths. Surgical resection of tumor bulk in the primary treatment of ovarian carcinoma. *Journal of the National Cancer Institute Monographs* 1975;**42**:1014.

**Guyatt 2008**

Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;**336**(7650):924-6.

**Higgins 2003**

Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;**327**:557-60.

**Hoskins 1994**

Hoskins WJ, McGuire WP, Brady MF, Homesley HD, Creasman WT, Berman M, et al. The effect of diameter of largest residual disease on survival after primary cytoreductive surgery in patients with suboptimal residual epithelial ovarian carcinoma: a Gynecologic Oncology Group study. *American Journal of Obstetrics and Gynecology* 1994;**170**:974-80.

**Hui 2022**

Hui S, Bryant A, Kunonga P, Gajjar K, Naik R. Ultra-radical (extensive) surgery versus standard surgery for the primary cytoreduction of advanced epithelial ovarian cancer. *Cochrane Database of Systematic Reviews* 2022, Issue In Press.

**Hunter 1992**

Hunter RW, Alexander ND, Soutter WP. Meta-analysis of surgery in advanced ovarian carcinoma: Is maximum cytoreductive surgery an independent determinant of prognosis? *American Journal of Obstetrics and Gynecology* 1992;**166**:504-11.

**Jemal 2017**

Jemal A, Ward E, Johnson C, Cronin K, Ma J, Ryerson B, et al. Annual Report to the Nation on the Status of Cancer, 1975–2014, Featuring Survival Cancer Statistics. *Journal of the National Cancer Institute* 2017;**109**:https://doi.org/10.1093/jnci/djx030.

**Kobal 2018**

Kobal B, Noventa M, Cvjeticanin B, Barbic M, Meglic L, Herzog M, et al. Primary debulking surgery versus primary neoadjuvant chemotherapy for high grade advanced stage ovarian cancer: comparison of survivals. *Radiology and Oncology* 2018;**52**(3):307-19.

**Kumar 2017**

Kumar A, Langstraat CL, DeJong SR, McGree ME, Bakkum-Gamez JN, Weaver AL, et al. Functional not chronologic age: frailty index predicts outcomes in advanced ovarian cancer. *Gynecologic Oncology* 2017;**147**(1):104-9.

**Kurman 2010**

Kurman RJ, Shih IeM. The origin and pathogenesis of epithelial ovarian cancer: a proposed unifying theory. *American Journal of Surgical Pathology* 2010;**34**(3):433-43.

**Kurman 2011**

Kurman RJ, Shih IeM. Molecular pathogenesis and extra ovarian origin of epithelial ovarian cancer-shifting the paradigm. *Human Pathology* 2011;**42**(7):918-31.

**Kurman 2014**

Kurman RJ, Carcangiu ML, Herrington CS. WHO Classification of Tumours of Female Reproductive Organs. 4 edition. Lyon: WHO Press, 2014.

**Markar 2016**

Makar AP, Tropé CG, Tummers P, Denys H, Vandecasteele K. Advanced ovarian cancer: primary or interval debulking? Five

categories of patients in view of the results of randomized trials and tumor biology: Primary ebulking surgery and interval debulking surgery for advanced ovarian cancer. *Oncologist* 2016;**21**(6):745-54.

**Markman 2001**

Markman M, Bundy BN, Alberts DS, Fowler JM, Clark-Pearson DL, Carson LF, et al. Phase III trial of standard-dose intravenous cisplatin plus paclitaxel versus moderately high-dose carboplatin followed by intravenous paclitaxel and intraperitoneal cisplatin in small-volume stage III ovarian carcinoma: an Intergroup Study of the Gynecologic Oncology Group, Southwestern Oncology Group, and Eastern Cooperative Oncology Group. *Journal of Clinical Oncology* 2001;**19**(4):1001-7.

**Moons 2014**

Moons KGM, de Groot JAH, Bouwmeester W, Vergouwe Y, Mallett S, Altman DG, et al. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. *PLOS Medicine* 2014;**11**(10):e1001744.

**Muggia 2000**

Muggia FM, Braly PS, Brady MF, Sutton GN, Theodore HL, Samuel AL, et al. Phase III randomized study of cisplatin versus paclitaxel versus cisplatin and paclitaxel in patients with suboptimal stage III or IV ovarian cancer: a Gynecologic Oncology Group Study. *Journal of Clinical Oncology* 2000;**18**(1):106-15.

**Napoletano 2010**

Napoletano C, Bellati F, Landi R, Pauselli S, Marchetti C, Visconti V, et al. Ovarian cancer cytoreduction induces changes in T cell population subsets reducing immunosuppression. *Journal of Cellular and Molecular Medicine* 2010;**14**(12):2748-59.

**National Comprehensive Cancer Network 2020**

National Comprehensive Cancer Network. NCCN clinical practice guidelines in oncology, ovarian cancer including fallopian tube cancer and primary peritoneal cancer. http://www.nccn.org/professionals/physician_gls/pdf/ovarian.pdf 2020.

**NICE 2013**

National Institute for Health and Care Excellence. Ultra-radical (extensive) surgery for advanced ovarian cancer. NICE Interventional procedure guidance [IPG 470]. https://www.nice.org.uk/guidance/ipg470 2013.

**Ogundimu 2016**

Ogundimu EO, Altman DG, Collins GS. Adequate sample size for developing prediction models is not simply related to events per variable. *Journal of Clinical Epidemiology* 2016;**76**:175-82. [DOI: 10.1016/j.jclinepi.2016.02.031]

**Ozols 2003**

Ozols RF, Bundy BN, Greer BE, Fowler JM, Clarke-Pearson DB, Mannel RA, et al. Phase III trial of carboplatin and paclitaxel compared with cisplatin and paclitaxel in patients with optimally resected stage III ovarian cancer: a Gynecologic Oncology Group Study. *Journal of Clinical Oncology* 2003;**21**(17):3194-200.

**Parmar 1998**

Parmar MK, Torri V, Stewart L. Extracting summary statistics to perform meta-analyses of the published literature for survival endpoints. *Statistics in Medicine* 1998;**17**(24):2815-34.

**Perets 2016**

Perets R, Drapkin R. It's totally tubular. Riding the new wave of ovarian cancer research. *Cancer Research* 2016;**76**(1):10-7.

**Redman 1994**

Redman CW, Warwick J, Luesley DM, Varma R, Lawton FG, Blackledge GR. Intervention debulking surgery in advanced epithelial ovarian cancer. *British Journal of Obstetrics and Gynaecology* 1994;**101**:142-6.

**Reuss 2019**

Reuss A, du Bois A, Harter P, Fotopoulou C, Sehouli J, Aletti G, et al. TRUST: trial of radical upfront surgical therapy in advanced ovarian cancer (ENGOT ov33/AGO-OVAR OP7). *International Journal of Gynecologic Cancer* 2019;**29**:1327-31.

**Review Manager 2020 [Computer program]**

Review Manager 5 (RevMan 5). Version 5.4. Copenhagen: The Cochrane Collaboration, 2020.

**Riley 2019**

Riley RD, Moons KGM, Snell KIE, Ensor J, Hooft L, Altman DG, et al. A guide to systematic review and meta-analysis of prognostic factor studies. *BMJ* 2019;**364**:k4597.

**Shafi 2018**

Shafi M, Bolton H, Gajjar K, editor(s). Gynaecological Oncology for the MRCOG. Cambridge: Cambridge University Press, 2018. [DOI: 10.1017/9781316986844]

**Spriggs 2007**

Spriggs DR, Brady M, Vaccarello L, Clarke-Pearson DL, Burger RA, Mannel R, et al. Phase III randomized trial of intravenous cisplatin plus a 24- or 96-hour infusion of paclitaxel in epithelial ovarian cancer: a Gynecologic Oncology Group Study. *Journal of Clinical Oncology* 2007;**25**(28):4466-71.

**Sterne 2011**

Sterne JAC, Egger M, Moher D (editors). Chapter 10: Addressing reporting biases. In: Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from training.cochrane.org/handbook.

**Stuart 2011**

Stuart GC, Kitchener H, Bacon M, duBois A, Friedlander M, Ledermann J, et al. Gynecologic Cancer InterGroup (GCIG) consensus statement on clinical trials in ovarian cancer: report from the Fourth Ovarian Cancer Consensus Conference. *International Journal of Gynecological Cancer* 2011;**21**(4):750–5.

**Sundar 2022**

Sundar S, Cummins C, Kumar S, Long J, Arora V, Balega J, et al. Quality of life from cytoreductive surgery in advanced ovarian cancer: Investigating the association between disease burden and surgical complexity in the international,

prospective, SOCQER-2 cohort study. *British Journal of Obstetrics and Gynaecology* 2022;**129**(7):1122-32. [DOI: 10.1111/1471-0528.17041] [PMCID: PMC9306902] [PMID: 34865316]

**Sørensen 2019**

Sørensen SM, Schnack TH, Høgdall C. Impact of residual disease on overall survival in women with federation of gynecology and obstetrics stage IIIB-IIIC vs stage IV epithelial ovarian cancer after primary surgery. *Acta Obstetricia et Gynecologica Scandinavica* 2019;**98**(1):34–43.

**Tangjitgamol 2016**

Tangjitgamol S, Manusirivithaya S, Laopaiboon M, Lumbiganon P, Bryant A. Interval debulking surgery for advanced epithelial ovarian cancer. *Cochrane Database of Systematic Reviews* 2016, Issue 1. Art. No: CD006014. [DOI: 10.1002/14651858.CD006014]

**Torre 2018**

Torre LA, Trabert B, DeSantis CE, Miller K, Samimi G, Runowicz C, et al. Ovarian cancer statistics. *CA: a Cancer Journal for Clinicians* 2018;**68**(4):284–96.

**Van der Burg 1995**

van der Burg ME, van Lent M, Buyse M, Kobierska A, Colombo N, Favalli G, et al. The effect of debulking surgery after induction chemotherapy on the prognosis in advanced epithelial ovarian cancer. Gynecological cancer co-operative group of the European organization for research and treatment of cancer. *New England Journal of Medicine* 1995;**332**(10):629-34.

**Vergote 2013**

Vergote I, du Bois A, Amant F, Heitz F, Leunen K, Harter P. Neoadjuvant chemotherapy in advanced ovarian cancer: on what do we agree and disagree? *Gynecologic Oncology* 2013;**128**:6-11.

**Vergote 2016**

Vergote I, Vlayen J, Heus P, Hoogendam JP, Damen JA, Van de Wetering FT, et al. Ovarian cancer: diagnosis, treatment and follow-up; KCE Report 28Cs; 2016. Available at: kce.fgov.be/sites/default/files/atoms/files/KCE_268Cs_Ovarian_cancer_summary.pdf 2016.

**Webb 2017**

Webb PM, Jordan SJ. Epidemiology of epithelial ovarian cancer. *Clinical Obstetrics & Gynaecology* 2017;**41**:3-14.

**Whiting 2011**

Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks, JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine* 2011;**155**(8):529-36.

**Wolff 2019**

Wolff RF, Moons KGM, Riley RD, Whiting PF, Westwood M, Collins GS, et al. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. *Annals of Internal Medicine* 2019;**170**(1):51-8.

**Woo 2012**

Woo YL, Kyrgiou M, Bryant A, Everett T, Dickinson HO. Centralisation of services for gynaecological cancer. *Cochrane Database of Systematic Reviews* 2012, Issue 3. Art. No: CD007945. [DOI: 10.1002/14651858.CD007945.pub2]

* Indicates the major publication for the study

## CHARACTERISTICS OF STUDIES

### Characteristics of included studies *[ordered by study ID]*

**Akahira 2001**

| ***Study characteristics*** | |
|---|---|
| Methods | Multicentre retrospective analysis:<br><br>24 Japanese institutions received questionnaires regarding stage IV epithelial ovarian cancer women |
| Participants | 225 women with stage IV ovarian cancer whose disease had been confirmed by exploration and only women with complete medical records were included. Stage IV disease was defined according to FIGO. Only women who underwent an initial attempt at surgical debulking were analysed.<br><br>The median age in the study was 54 years (range: 26 to 85 years)<br><br>All 225 women had FIGO stage IV disease<br><br>Histological cell type: serous: 136 (60.5%), mucinous: 16 (7%), clear cell 26 (11.5%), endometrioid 27 (12%), transitional 4 (2%), undifferentiated 12 (5%), other 4 (2%)<br><br>Extent of disease: pleural effusion: 89 (39.5%), liver: 34 (15%), lung: 8 (3.5%), lymph node: 44 (19.5%), other: 15 (6.5%), multiple sites: 35 (15%) |

**Akahira 2001** *(Continued)*

| | |
|---|---|
| | Performance status: 0: 26 (11%), 1: 76 (34%), 2: 49 (22%), 3: 67 (30%), 4: 7 (3%) |
| Residual disease details | **Intervention group:**<br>**'Optimal' cytoreduction** was defined as no gross residual tumour greater than 2 cm in diameter<br>**Comparison group:**<br>**LVRD** was defined as any gross residual disease remaining greater than 2 cm in diameter |
| Outcomes | Overall survival: HR adjusted for histology and performance status:<br>• < 2 cm versus > 2 cm; HR 0.42 (95% CI 0.31 to 0.64), or > 2 cm vs < 2c m; HR 2.39 (95% CI 1.68 to 3.40) so that reference group is consistent throughout review<br>Adverse events; median blood loss, blood transfusions |
| Risk of bias (QUIPS) | 1. Study participation (a-f): low risk<br>Adequate number of participants and description of target population. Baseline characteristics, eligibility criteria, sampling frame and period/place study took place presented clearly.<br>2. Study attrition (a-e): unclear risk<br>Unclear if patients with incomplete follow-up were excluded before arriving at the stated sample size. Insufficient information to permit judgement.<br>3. Prognostic factor measurement (a-f): unclear risk<br>Valid and reliable measurement of RD. Multicentre design may introduce heterogeneity in measurement of RD<br>**Outcome level assessment:**<br>**Outcome: overall survival**<br>4. Outcome measurement (a-c): low risk<br>Valid and reliable measurement of outcome<br>5. Adjustment for other prognostic factors (a-g): high risk<br>HR for OS was adjusted for histology, performance status and RD in multivariable Cox model<br>6. Statistical analysis and reporting (a-d): unclear risk<br>In methods, authors reported that significant variables from the univariate analysis were included in the multivariable model<br>**Outcome: progression-free survival**<br>Not reported |
| Notes | There were 70 women (31.1%) in the 'optimal' group and 155 (68.9%) in the LVRD group<br>The median follow-up time was 47.5 months (range: 13 to 112 months)<br>The median survival for all women with stage IV ovarian cancer was 20 months, with an estimated 5-year survival rate of 19.6%<br>Mean survival in the optimal group was 32 months and 16 months in the suboptimal group (P < 0.0001)<br>MV analysis included the histology and performance status as covariates in the model<br>The median duration of the debulking surgery was 240 minutes (range: 40 to 780 minutes. |

**Akahira 2001** *(Continued)*

| | |
|---|---|
| | The median estimated blood loss was 1085 mL (range 40 to 11,000 mL), and 112 women (50%) received blood transfusions intra- and postoperatively |

**Aletti 2006**

| ***Study characteristics*** | |
|---|---|
| Methods | Retrospective cohort study of consecutive women identified from surgical records |
| Participants | Women with FIGO stage IIIC ovarian cancer, where disease status was extracted from surgical exploration notes<br><br>The mean and median age at study entry was 64.4 and 64 years respectively (range: 24 to 87)<br><br>All women presented with FIGO stage IIIC - 194 (100%)<br><br>Tumour cell type: serous 126 (64.9%), mucinous: 4 (2.1%), endometrioid: 18 (9.3%), clear cell: 7 (3.6%), mixed: 17 (8.8%), seroanaplastic: 17 (8.8%), mullerian origin: 2 (1%)<br><br>Tumour grade: 1: 1 (0.5%), 2: 13 (6.7%), 3: 180 (92.8%)<br><br>ASA score: 1: 7 (3.6%), 2: 87 (44.8%), 3: 88 (45.4%), 4: 7 (3.6%), unknown: 5 (2.6%)<br><br>Ascites: mean: 2076 mL, median 1000 mL, (range: 0 to 12,000 mL)<br><br>Extent of disease: carcinomatosis: 144 (74.2%), diaphragm involvement: 137 (70.6%), mesentery: 138 (71.1%), cul-de-sac: 163 (84%), omentum 168: (86.6%), ascites 160: (82.5%) |
| Residual disease details | Residual disease was noted as follows:<br><br>1. NMRD: 46 (23.7%)<br>2. SVRD: 85 (43.8%)<br>3. Residual disease of 1 cm to 2 cm: 22 (11.3%)<br>4. Residual disease larger than 2cm: 41 (21.1%)<br><br>Optimal cytoreduction was defined as residual disease < 1 cm<br><br>All women were scheduled for treatment with first-line postoperative platinum-based chemotherapy (paclitaxel or cyclophosphamide for 6 to 8 courses, every 3 to 4 weeks) |
| Outcomes | • Overall survival, HR adjusted for several prognostic categories:<br>  ○ SVRD vs NMRD: HR 3.89 (95% CI 2.27 to 7.11)<br>  ○ 1 cm to 2 cm vs NMRD: HR 6.25 (95% CI 3.16 to 12.61)<br>  ○ > 2 cm vs NMRD: HR 13.00 (95% CI 7.14 to 24.87)<br>• Adverse events:<br>  ○ Perioperative mortality rate, defined as the percentage of women who died within 30 days of surgery, was 1.5% (3/194; 95% CI 0.5 to 4.4%). However, there was no breakdown by treatment arm. |
| Risk of bias (QUIPS) | 1. Study participation (a-f): low risk<br><br>Adequate number of participants and description of target population. Baseline characteristics, eligibility criteria, sampling frame and period/place study took place presented clearly.<br><br>2. Study attrition (a-e): unclear risk<br><br>Unclear if patients with incomplete follow-up were excluded before arriving at the stated sample size. Insufficient information to permit judgement.<br><br>3. Prognostic factor measurement (a-f): low risk |

**Aletti 2006** *(Continued)*

| | |
|---|---|
| | Valid and reliable measurement of RD |
| | **Outcome level assessment:** |
| | **Outcome: overall survival** |
| | 4. Outcome measurement (a-c): high risk |
| | Overall survival not used as outcome. Rather, disease-specific survival was used. |
| | 5. Adjustment for other prognostic factors (a-g): low risk |
| | HR for disease-specific survival was adjusted for residual disease, age, American Society of Anesthesiology (ASA) score, histological grade, operative time and aggressive surgery in multivariable Cox model |
| | 6. Statistical analysis and reporting (a-d): unclear risk |
| | In methods, authors reported that significant variables from the univariate analysis were included in the multivariable model |
| | **Outcome: progression-free survival** |
| | Not reported |
| Notes | Median length of follow-up: 2.7 years |
| | Mean length of follow-up: 3.5 years (range 0.02 to 10.5 years) |
| | 5-year disease-specific death rate: |
| | Optimal group: 70/131 (53.4%) |
| | Suboptimal group: 56/63 (88.9%) |

**Ataseven 2016**

| | |
|---|---|
| ***Study characteristics*** | |
| Methods | Prospective cohort study |
| Participants | 326 consecutive women with FIGO IV |
| | Median age in the study was 61 years (range: 19 to 88 years) |
| | All 326 women presented with FIGO stage IV disease |
| | Histological cell type: high grade serous: 287 (88.0%), others: 39 (12.0%) |
| | Ascites: ≤ 500 mL: 149 (45.7%), > 500 mL: 177 (54.3%) |
| | Performance status: ECOG 0: 248 (76.1%), ECOG > 0: 78 (23.9%) |
| | Localization of metastasis:<br>• Pleural effusion/involvement: 134 (41.1%)<br>• Abdominal wall: 133 (40.8%)<br>• Extraregional lymph node: 63 (19.3%)<br>• Liver: 45 (13.8%)<br>• Spleen: 22 (6.7%)<br>• Others: 19 (5.8%) |

**Ataseven 2016** *(Continued)*

| | |
|---|---|
| | Germany |
| Residual disease details | Surgery was performed by accredited gynaecological oncologists<br>Cohort 1 included 286 women who underwent primary debulking surgery<br>Postoperative chemotherapy was administered in 92% (263/286)<br>Cohort 2 included 40 women who underwent either no surgery or only diagnostic procedures without cytoreductive intention (NoCS - no cytoreductive surgery)<br>In cohort 2, platinum-based chemotherapy was given to 87.5% (35/40) of women<br>Residual disease for total cohort was noted as follows, n (%):<br>• NMRD: 157 (48.2%)<br>• SVRD: 88 (27.0%)<br>• LVRD (> 10 mm): 41 (12.6%)<br>• No cytoreduction: 40 (12.3%) |
| Outcomes | Overall survival: HR adjusted for age, performance status, residual tumour, tumour stage and ascites<br>NMRD: HR 1<br>SVRD: HR 1.50 (95% CI 1.01 to 2.23)<br>LVRD (> 10 mm): HR 2.20 (95% CI 1.36 to 3.55) |
| Risk of bias (QUIPS) | 1. Study participation (a-f): low risk<br>Adequate number of participants and description of target population. Baseline characteristics, eligibility criteria, sampling frame and period/place study took place presented clearly.<br>2. Study attrition (a-e): unclear risk<br>Unclear if patients with incomplete follow-up were excluded before arriving at the stated sample size. Insufficient information to permit judgement.<br>3. Prognostic factor measurement (a-f): low risk<br>Valid and reliable measurement of RD<br>**Outcome level assessment:**<br>**Outcome: overall survival**<br>4. Outcome measurement (a-c): low risk<br>Valid and reliable measurement of outcome<br>5. Adjustment for other prognostic factors (a-g): low risk<br>HR for OS was adjusted for age, performance status, residual tumour, tumour stage and ascites in a multivariable Cox model<br>6. Statistical analysis and reporting (a-d): high risk<br>No conceptual framework; unclear of variable selection criteria in multivariate model<br>**Outcome: progression-free survival**<br>Not reported |
| Notes | Follow-up time: up to 4 years (mean: 46 months; median: 34 months; interquartile range: 12 to 70 months) |

**Ataseven 2016** *(Continued)*

| | |
|---|---|
| | In total, 28 women (8.6%) did not receive chemotherapy<br><br>30-day mortality was observed in: 12/326 (3.68%)<br><br>Median OS for all women was 50.3 months<br><br>In cohort 1, complete resection was achieved in 54.9% (n = 157/286; RD0), cytoreduction to 1 mm to 10 mm in 30.7% (n = 88/286; RD1-10) and bulky residual disease exceeding 10 mm in 14.3% (n = 41/286; RD > 10)<br><br>Risk factors for residual disease after debulking surgery in women with EOC FIGO stage IV included:<br><br>• Age (OR 1.85, 95% CI 1.13 to 3.03; P = 0.015)<br>• Poor performance status (OR 3.46, 95% CI 1.67 to 7.18; P = 0.001)<br>• Large volume ascites > 500 mL (OR 2.37, 95% CI 1.06 to 4.22; P = 0.035)<br>• Presence of liver metastasis (OR 6.17, 95% CI 2.78 to 13.7; P < 0.001)<br><br>Length of hospital stay not reported |

**Bixel 2020**

| ***Study characteristics*** | |
|---|---|
| Methods | Retrospective analysis of past medical data from The Ohio State University Wexner Medical Center and Duke University Health System between January 2004 and April 2017<br><br>Multicentre study<br><br>USA |
| Participants | 134 patients diagnosed with stage III to IV ovarian, fallopian tube or primary peritoneal cancer<br><br>Median age (range): 64.3 (21 to 87)<br>Median BMI (range): 28.1 (16 to 52.5)<br>Ethnicity: 110 white (82%)<br>FIGO III: 49 (36%)<br>FIGO IV: 54 (40%)<br>FIGO stage not otherwise specified but considered advanced: 31 (24%)<br>Serous histology: 112 (83%)<br>Tumour grade 1: 3 (2%)<br>Tumour grade 2: 123 (92%)<br>Tumour grade unknown: 8 (6%) |
| Residual disease details | Women underwent interval debulking surgery<br><br>Optimal RD defined as RD ≤ 1 cm<br><br>NMRD: 89 (66%)<br><br>SVRD: 45 (34%) |
| Outcomes | Overall survival<br><br>Median OS: 35.3 (95% CI 28.6 to 42.9)<br><br>There was no multivariate model for overall survival despite there being progression-free survival<br><br>Progression-free survival<br><br>Disease recurrence: 117 (87%)<br>Median PFS: 12.2 (95% CI 11.3 to 13.7) |

**Bixel 2020** *(Continued)*

| | |
|---|---|
| | After controlling for NACT cycles, route of postoperative chemotherapy administration (intraperitoneal or intravenous), maintenance therapy (yes/no); residual disease (SVRD vs NMRD) (adjusted HR 1.564 (1.055 to 2.287))<br><br>2 (1%) patients died during treatment: 1 patient in the IP group died from a myocardial infarction and 1 patient in the IV group died as a result of sepsis with resulting complications |
| Risk of bias (QUIPS) | 1. Study participation (a-f): low risk<br><br>Adequate number of participants and description of target population. Baseline characteristics, eligibility criteria, sampling frame and period/place study took place presented clearly.<br><br>2. Study attrition (a-e): unclear risk<br><br>Unclear if patients with incomplete follow-up were excluded before arriving at the stated sample size. Insufficient information to permit judgement.<br><br>3. Prognostic factor measurement (a-f): unclear risk<br><br>Adequate cut-off for residual disease used (< 1 cm). Multicentre design may introduce heterogeneity in measurement of RD<br><br>**Outcome level assessment:**<br><br>**Outcome: overall survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Valid and reliable measurement of OS<br><br>5. Adjustment for other prognostic factors (a-g): high risk<br><br>OS was reported in KM curve but was not used in any multivariable modelling<br><br>6. Statistical analysis and reporting (a-d): high risk<br><br>There was only a multivariate model for PFS but not OS<br><br>**Outcome: progression-free survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Valid and reliable measurement of PFS<br><br>5. Adjustment for other prognostic factors (a-g): high risk<br><br>Model for PFS adjusted for NACT cycles, route of administration (IP or IV), maintenance therapy. However, none deemed to be critically important prognostic factors.<br><br>6. Statistical analysis and reporting (a-d): high risk<br><br>No conceptual framework; data driven based on P values of univariate associations. Unclear if multivariate Cox was used as logistic regression mentioned in methods but hazard ratios reported. There was only a multivariate model for PFS but not OS. |
| Notes | 37 (28%) patients receiving IP and 97 (72%) patients receiving IV chemotherapy<br><br>Median NACT cycles: 3 (range 1 to 6)<br><br>NACT regime<br><br>Platinum/taxane: 133 (99%)<br>Platinum/other: 1 (1%)<br><br>Adjuvant chemotherapy regime |

**Bixel 2020** *(Continued)*

| | |
|---|---|
| | Platinum/taxane: 122 (91%)<br>Platinum/other: 3 (2%)<br>Non platinum: 9 (7%)<br>Adjuvant chemotherapy cycles:<br>Intraperitoneal group: median 4 (range 2 to 6)<br>Intravenous group: median 3 (range 1 to 6)<br><br>Maintenance therapy following completion of planned chemotherapy: 10 (7%)<br><br>At the time of surgery, 32 (24%) patients underwent a bowel resection and 15 (11%) underwent extensive upper abdominal debulking procedures |

**Bristow 2011**

| ***Study characteristics*** | |
|---|---|
| Methods | Retrospective chart review at Johns Hopkins Hospital, USA<br><br>Women enrolment was between January 1995 and December 2008 |
| Participants | 405 women with FIGO stage IIIC epithelial ovarian cancer based on intraoperative findings or radiographic imaging coupled with fine-needle biopsy diagnosis. All epithelial histological subtypes were included. Borderline ovarian tumours of low malignant potential were excluded.<br><br>Women characteristics reported as Whites (n = 366) vs African-Americans (n = 39)<br><br>Median age: 59 vs 59 years<br><br>ASA class, I/II/III/IV: 5/124/232/5 vs 0/4/31/4<br><br>Histology, serous/non-serous: 314/52 vs 31/8<br><br>Tumour grade, 1/2/3: 39/33/294 vs 2/4/33<br><br>Optimal RD (defined as ≤ 1 cm)/no gross RD: 267/188 vs 18/21 |
| Residual disease details | All women underwent attempted surgical cytoreduction either primarily<br><br>Residual disease was defined as:<br><br>SVRD (RD 0.1 cm to 1.0 cm)<br><br>NMRD (no gross RD)<br><br>Residual disease was noted as follows:<br><br>Optimal (≤ 1 cm): White, n (%): 178 (44%); African-American; n (%): 18 (4.5%)<br><br>NMRD: White, n (%): 188 (46.5%); African-American; n (%):21 (5%) |
| Outcomes | SVRD vs NMRD: HR for OS 2.74 (95% CI 1.98 to 3.71) (HR adjusted for age, race, tumour grade, histology, ASA score, surgical complexity score, serum albumin, administration of platinum-based chemotherapy and significant peri-operative morbidity)<br><br>OS was calculated from the date of diagnosis using Kaplan–Meier curves and compared using the log-rank test and Cox proportional hazards model |
| Risk of bias (QUIPS) | 1. Study participation (a-f): low risk<br><br>Adequate number of participants and description of target population. Baseline characteristics, eligibility criteria, sampling frame and period/place study took place presented clearly. |

**Bristow 2011** *(Continued)*

2. Study attrition (a-e): unclear risk

Unclear if patients with incomplete follow-up were excluded before arriving at the stated sample size. Insufficient information to permit judgement.

3. Prognostic factor measurement (a-f): low risk

Valid and reliable measurement of RD

**Outcome level assessment:**

**Outcome: overall survival**

4. Outcome measurement (a-c): low risk

Valid and reliable measurement of outcome

5. Adjustment for other prognostic factors (a-g): high risk

HR for OS was adjusted for race, tumour grade 3, non-serous histology, ASA score >3, surgical complexity score, serum albumin < 3.0 g/dL, platinum-based therapy, residual disease and perioperative morbidity in multivariable Cox model

6. Statistical analysis and reporting (a-d): high risk

No conceptual framework; unclear of variable selection criteria in multivariate model

**Outcome: progression-free survival**

Not reported

| | |
|---|---|
| Notes | A total of 433 ovarian cancer women were identified with stage IIIC disease. Of these, 28 women were variously classified as either Asian-Pacific Islander, Hispanic, unknown or other and were excluded from further study.<br><br>Source of funding: the Queen of Hearts Foundation for Ovarian Cancer Research<br><br>Declaration of interest: none declared<br><br>Median follow-up: 33.0 months<br><br>The 30-day mortality rate for all 405 women was 1.5%<br><br>Retrospective non-randomised study. Blinding not reported (but not applicable). Adjusted HRs are derived from a prognostic model. No details on how modelling was performed, but this seems to have been done based on significance testing (and not on including putative confounders in the analysis, irrespective of statistical significance).<br><br>Women and disease characteristics not reported according to debulking status. NB: study only included women with stage IIIC ovarian cancer/possible overlap with Peiretti 2012. |

**Chan 2003**

| ***Study characteristics*** | |
|---|---|
| Methods | Retrospective cohort study |
| Participants | All consecutive cases of advanced-stage epithelial ovarian carcinoma diagnosed in younger women (range 22 to 45 years) were identified from tumour registry databases and a comparable group of 52 women who averaged 21 years older (range 46 to 85 years) was selected as controls. One-to-one matching from the same database was performed based on the date of diagnosis and stage of disease |

**Chan 2003** *(Continued)*

| | |
|---|---|
| | during the same period in the same institution. Thus, the controls were similarly distributed across 17 years.<br><br>The mean age at study entry was 50.5 years with a range between 22 and 85 years (40 (SD 5.7) and 61 years (SD 8.7) for younger and older women respectively)<br><br>5 (4.8%) women had FIGO stage IIIA, 5 (4.8%) had stage IIIB, 74 (71.1%) women had stage IIIC and 20 (19.2%) had stage IV disease<br><br>Tumour cell type: papillary serous 72 (63.16%), mucinous: 3 (2.63%), endometrioid: 17 (14.9%), clear cell: 1 (0.88%), small cell: 3 (2.63%), undifferentiated: 8 (7%)<br><br>Tumour grade: 1: 8 (7%), 2: 24 (21.1%), 3: 72 (63.2%)<br><br>Performance status: 0: 65 (57%), 1 to 2: 35 (30.7%), unknown: 4 (3.51%) |
| Residual disease details | Residual disease was noted as follows:<br><br>1. SVRD: 71 (62.3%)<br>2. LVRD (> 1 cm): 43 (37.7%)<br><br>Women were divided into SVRD (defined as optimal) and 1 cm or more (defined as suboptimal) groups based on residual disease after initial surgery. Optimal debulking was achieved in 36 (69%) and 35 (67%) women in younger in older groups respectively.<br><br>All women received either a platinum/paclitaxel or a platinum/cyclophosphamide regimen for primary chemotherapy and women who underwent neoadjuvant chemotherapy with interval debulking were removed from the study.<br><br>Gynaecology oncologists from the academic institution surgically staged all women. |
| Outcomes | A multivariable analysis which included older versus younger age, stage (IV vs III), performance status (1 to 2 vs 0) and residual disease (LVRD (> 1 cm) vs SVRD) was performed to evaluate all factors that were significant in the univariate analysis<br><br>Overall survival: HR adjusted for prognostic categories (see above):<br><br>• LVRD (> 1 cm) vs SVRD HR 1.67 (95% CI 1.03 to 2.72) |
| Risk of bias (QUIPS) | 1. Study participation (a-f): low risk<br><br>Adequate number of participants and description of target population. Baseline characteristics, eligibility criteria, sampling frame and period/place study took place presented clearly.<br><br>2. Study attrition (a-e): unclear risk<br><br>Unclear if patients with incomplete follow-up were excluded before arriving at the stated sample size. Insufficient information to permit judgement.<br><br>3. Prognostic factor measurement (a-f): unclear risk<br><br>Valid and reliable measurement of RD. Multicentre design may introduce heterogeneity in measurement of RD.<br><br>**Outcome level assessment:**<br><br>**Outcome: overall survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Definition of OS not provided but no reason to doubt they used standard definition<br><br>5. Adjustment for other prognostic factors (a-g): low risk |

**Chan 2003** *(Continued)*

| | |
|---|---|
| | HR for OS was adjusted for residual disease, age (older versus younger), stage (IV versus III) and performance status (1 to 2 versus 0) in a multivariable Cox model<br><br>6. Statistical analysis and reporting (a-d): high risk<br><br>No conceptual framework; unclear of variable selection criteria in multivariate model<br><br>**Outcome: progression-free survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Definition of PFS not provided but no reason to doubt they used standard definition<br><br>5. Adjustment for other prognostic factors (a-g): high risk<br><br>PFS was reported in table comparing younger vs older patients but was not used in any multivariable modelling<br><br>6. Statistical analysis and reporting (a-d): high risk<br><br>There was only a multivariate model for OS but not PFS |
| Notes | The median follow-up after surgery was 33 months (range 6 to 142 months)<br><br>5-year survival: of younger and older women: SVRD: 59% and 21% in young and old women respectively, LVRD (> 1 cm): 28% and 22% in young and old women respectively<br><br>Median survival: SVRD: 66 months and 45 in young and old women respectively, LVRD (> 1 cm): 37 and 19 months in young and old women respectively, P = 0.003<br><br>Other variables in Cox model:<br><br>Older versus younger age (HR 1.82, 95% CI 1.09 to 3.05), stage IV versus stage III disease (HR 3.00, 95% CI 1.71 to 5.25), performance status 1 to 2 versus 0 (HR 1.89, 95% CI 1.13 to 3.15)<br><br>Despite the higher prevalence of poorly differentiated tumours in the older group, tumour grade (3 versus 1 to 2) was not an important prognostic factor in multivariable analysis (HR 1.06, 95% CI 0.57 to 1.97) |

**Chang 2012a**

| ***Study characteristics*** | |
|---|---|
| Methods | Retrospective review of medical records |
| Participants | All women underwent primary cytoreductive surgery followed by platinum-based chemotherapy.<br><br>Consecutive women with stage IIIC and IV primary epithelial ovarian, fallopian tube or peritoneal cancer who underwent primary cytoreductive surgery at Ajou University Hospital between 1 January 2000 and 31 December 2011.<br><br>Women received neoadjuvant chemotherapy, operated in other institution, stage IIIC due to nodal involvement were excluded<br><br>N = 203<br><br>Median age was 54 years (range 30 to 78)<br><br>Median BMI 23.3 (range 11.7 to 35.2)<br><br>ASA 1 to 2: 114 (56.2%), 3 to 4: 80 (39.4%) |

**Chang 2012a** *(Continued)*

| | |
|---|---|
| | Stage IIIC: 189 (93.1%), IV: 14(6.9%)<br><br>Tumour grade 1: 26 (12.8%), grade 2: 72 (35.5%), grade 3: 100 (49.3%)<br><br>Histological subtype: serous: 167 (82.3%), mucinous: 4 (2.0%), endometrioid: 5 (2.5%), clear cell: 9 (4.4%), mixed: 18 (8.9%)<br><br>Median pre-operative CA-125: 603.8 (range 4.5 to 21,677)<br><br>Ascites < 1000 mL (54.7%), > 1000 mL (45.3%)<br><br>Carcinomatosis: yes (73.4), no (26.6%)<br><br>Simple procedure (58.6%), radical procedure (41.4%). Cohort was divided into simple procedures and radical procedures group for statistical analysis. |
| Residual disease details | Residual disease were defined:<br><br>• NMRD (31.0%)<br>• SVRD 0.1 cm to 1.0 cm (37.9%)<br>• LVRD (> 1 cm) (31.0%) |
| Outcomes | Median follow-up was 43 months (range of 1 to 124)<br><br>**Kaplan-Meier**<br><br>Median unadjusted OS LVRD > 1 cm 37 months; SVRD 0.1 cm to 1 cm 46 months; NMRD 86 months<br><br>Median unadjusted PFS LVRD > 1 cm 9 months; SVRD 0.1 cm to 1 cm 15 months; NMRD 35 months<br><br>**Multivariate analysis for OS:**<br><br>HR (LVRD > 1 cm vs NMRD) 3.24 (95% CI 1.90 to 5.53)<br><br>HR (SVRD 0.1 cm to 1 cm vs NMRD): 2.22 (95% CI 1.25 to 3.94)<br><br>**Multivariate analysis for PFS:**<br><br>HR (LVRD > 1 cm vs NMRD): 3.40 (95% CI 2.00 to 5.77)<br><br>HR (SVRD 0.1 cm to 1 cm vs NMRD): 2.20 (95% CI 1.26 to 3.84)<br><br>HRs adjusted for age, FIGO stage and type of surgery (radical vs simple)<br><br>**Morbidity**<br><br>Operative time (minutes): simple: 235 (range 85 to 570), radical: 307 (range 150 to 810)<br><br>Estimated blood loss: simple: 500 (range 200 to 4000), radical: 800 (range 300 to 7500)<br><br>Intraoperative blood transfusion: simple (17.6%), radical (25.0%)<br><br>Postoperative blood transfusion: simple (26.1%), radical (39.3%)<br><br>Length of stay in ICU: simple: 0.8 (0 to 6), radical: 1.5 (0 to 6)<br><br>Postoperative morbidity: simple (11.8%), radical (38.1%)<br><br>Postoperative death < 30 days: simple = 0, radical = 1 |
| Risk of bias (QUIPS) | 1. Study participation (a-f): low risk<br><br>Adequate number of participants and description of target population. Baseline characteristics, eligibility criteria, sampling frame and period/place study took place presented clearly.<br><br>2. Study attrition (a-e): unclear risk |

**Chang 2012a** *(Continued)*

| | |
|---|---|
| | Unclear if patients with incomplete follow-up were excluded before arriving at the stated sample size. Insufficient information to permit judgement.<br><br>3. Prognostic factor measurement (a-f): low risk<br><br>Valid and reliable measurement of RD<br><br>**Outcome level assessment:**<br><br>**Outcome: overall survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Definition of OS not provided but it usually has a standard definition<br><br>5. Adjustment for other prognostic factors (a-g): low risk<br><br>HR for OS was adjusted for stage (IV), surgical procedure, residual disease and age in a multivariable Cox model<br><br>6. Statistical analysis and reporting (a-d): high risk<br><br>No conceptual framework; unclear of variable selection criteria in multivariate model<br><br>**Outcome: progression-free survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Definition of PFS not provided but it usually has a standard definition.<br><br>5. Adjustment for other prognostic factors (a-g): low risk<br><br>HR for PFS was adjusted for stage (IV), surgical procedure, residual disease and age in a multivariable Cox model<br><br>6. Statistical analysis and reporting (a-d): high risk<br><br>No conceptual framework; unclear of variable selection criteria in multivariate model |
| Notes | Subgroup analysis for 139 women with peritoneal carcinomatosis, the median unadjusted OS LVRD > 1 cm 39 months, SVRD 0.1 cm to 1 cm 50 months, NMRD 86 months |

**Chang 2012b**

| ***Study characteristics*** | |
|---|---|
| Methods | Retrospective review of medical records |
| Participants | Consecutive women with stage IIIC primary epithelial ovarian, fallopian tube or peritoneal cancer who underwent primary cytoreductive surgery at Ajou University Hospital between 1 January 2000 and 31 December 2011<br><br>After primary surgery, all women received adjuvant chemotherapy consisting of cisplatin (75 mg/$m^2$) or carboplatin (area under the curve; 5 to 7) and paclitaxel (135 mg/$m^2$) based systemic combination chemotherapy (every 3 weeks for 6 to 9 cycles)<br><br>Exclusion: primary cytoreduction at an outside institution, neoadjuvant chemotherapy, stage IIIC disease based on lymph node metastasis only or borderline malignancy<br><br>N = 191<br><br>Median age was 54 years (range 30 to 78) |

**Chang 2012b** *(Continued)*

| | |
|---|---|
| | Median BMI 23.2 (18.1 to 35.2)<br><br>ASA 1 or 2: 107 (56.6%), 3 or 4: 74 (39.2%)<br><br>Median pre-op CA-125 173.1 (range 4.5 to 21,677)<br><br>Histological subtypes: serous: 155 (82%), mucinous: 4 (2.1%), endometrioid: 4 (2.1%), clear cell: 9 (4.8%), mixed: 17 (9.0%)<br><br>Grade 1: 26 (13.8%), grade 2: 67 (35.4%), grade3: 5 (2.6%)<br><br>Ascites < 1000 mL (57.7%), > 1000 mL (42.3%)<br><br>Peritoneal carcinomatosis: yes:139 (73.5%), no: 50 (26.5%)<br><br>Systematic lymphadenectomy (n = 135), no lymphadenectomy (n = 54)<br><br>Lymphadenectomy; pelvic only (22.2%), pelvic and para-aortic (77.8%) |
| Residual disease details | Residual disease were defined:<br><br>• NMRD: 61 (32.3%)<br>• SVRD (0.1 to 1.0 cm): 67 (35.4%)<br>• LVRD (> 1.0 cm): 61 (32.3%)<br><br>Overall surgical morbidity - blood transfusion, deep vein thrombosis, sepsis, intestinal obstruction, ileus, lymphocyst or wound dehiscence was significantly higher in women who had lymphadenectomy |
| Outcomes | **Multivariate analysis for OS:**<br><br>SVRD 0.1 cm to 1 cm vs NMRD: HR 2.25 (95% CI 1.25 to 4.03)<br><br>LVRD > 1 cm vs NMRD: HR 3.09 (95% CI 1.80 to 5.30)<br><br>HRs adjusted for age, performance of radical surgery and performance of lymphadenectomy |
| Risk of bias (QUIPS) | 1. Study participation (a-f): low risk<br><br>Adequate number of participants and description of target population. Baseline characteristics, eligibility criteria, sampling frame and period/place study took place presented clearly.<br><br>2. Study attrition (a-e): unclear risk<br><br>Unclear if patients with incomplete follow-up were excluded before arriving at the stated sample size. Insufficient information to permit judgement.<br><br>3. Prognostic factor measurement (a-f): low risk<br><br>**Valid and reliable measurement of RD**<br><br>**Outcome level assessment:**<br><br>**Outcome: overall survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Definition of OS not provided but it usually has a standard definition<br><br>5. Adjustment for other prognostic factors (a-g): low risk<br><br>HR for OS was adjusted for residual disease, type of surgery, performance of lymphadenectomy and age in a multivariable Cox model<br><br>6. Statistical analysis and reporting (a-d): high risk |

**Chang 2012b** *(Continued)*

| | |
|---|---|
| | No conceptual framework; unclear of variable selection criteria in multivariate model<br><br>**Outcome: progression-free survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Definition of PFS not provided but it usually has a standard definition.<br><br>5. Adjustment for other prognostic factors (a-g): low risk<br><br>HR for PFS was adjusted for residual disease, type of surgery, performance of lymphadenectomy and age in a multivariable Cox model<br><br>6. Statistical analysis and reporting (a-d): high risk<br><br>No conceptual framework; unclear of variable selection criteria in multivariate model |
| Notes | Systematic lymphadenectomy was performed in 135 (71.4%) of whom 105 had both pelvic and para-aortic lymphadenectomy. The mean number of dissected pelvic and para-aortic nodes were 25 (range 11 to 57) and 11 (range 3 to 35), respectively. 53.4% were found to have grossly enlarged lymph nodes during surgery.<br><br>Of 135 women who underwent systematic lymphadenectomy, positive lymph nodes were found in 59%.<br><br>The median unadjusted OS; lymphadenectomy 66 months, no lymphadenectomy 40 months. Sub-group analysis of NMRD: median OS 86 month versus no lymphadenectomy 46 months<br><br>Of 189 women, tumour recurred in 110 women (58.2%) and 90 (47.6%) died of disease. 65 women with lymphadenectomy and 45 without lymphadenectomy had disease recurrence and there is no significant difference in the site of disease recurrence. |

**Chi 2001**

| ***Study characteristics*** | |
|---|---|
| Methods | Retrospective cohort study |
| Participants | 282 women with stage III and IV epithelial ovarian cancer. Women with ovarian tumours of low-malignant potential were excluded from this study.<br><br>All women were treated between 1987 and 1994 at Memorial Sloan-Kettering Cancer Center (MSKCC)<br><br>The median age at study entry was 59 years with a range between 22 and 87 years<br><br>22 (8%) women had FIGO stage IIIA/IIIB, 194 (69%) had stage IIIC and 66 (23%) had stage IV disease<br><br>Tumour cell type: serous 199 (71%), endometrioid: 46 (16%), clear cell: 19 (7%), mucinous: 10 (4%), mixed: 8 (3%)<br><br>Tumour grade: 1: 13 (5%), 2: 69 (24%), 3: 184 (65%)<br><br>Ascites: yes: 238 (84%), no: 43 (15%), unknown: 1 (1%) |
| Residual disease details | Women were treated with primary surgery followed by chemotherapy<br><br>**Type of surgeon**<br><br>Residual disease was noted as follows:<br><br>1. SVRD: 71 (25.2%)<br>2. Residual disease between 1 cm and 2 cm: 73 (26%) |

**Chi 2001** *(Continued)*

| | |
|---|---|
| | 3. LVRD greater than 2 cm: 137 (48.7%)<br><br>The following types of chemotherapy were given to women in the study: cisplatin/cyclophosphamide: 143 (51%), carboplatin/cyclophosphamide: 65 (23%), carboplatin/paclitaxel: 31 (11%), cisplatin/paclitaxel 24 (8%), carboplatin: 7 (3%), cisplatin 1 (< 1%), none or unknown 10 (4%)<br><br>Gynaecology oncologists from the academic institution surgically staged all women |
| Outcomes | A multivariable analysis which included age, stage (IIIC and IV vs IIIA/IIIB), ascites (yes vs no) and residual disease (1 cm to 2cm and > 2 cm vs < 1 cm) was performed to evaluate important prognostic factors<br><br>Overall survival: HR adjusted for prognostic categories (see above):<br><br>• 1 cm to 2 cm vs SVRD: HR 1.7 (95% CI 1.1 to 2.6)<br>• LVRD (> 2 cm) vs SVRD: HR 2.0 (95% CI 1.3 to 2.9)<br><br>**Direct surgical morbidity**<br><br>8 women (2.83%) died within 1 month of surgery |
| Risk of bias (QUIPS) | 1. Study participation (a-f): low risk<br><br>Adequate number of participants and description of target population. Baseline characteristics, eligibility criteria, sampling frame and period/place study took place presented clearly.<br><br>2. Study attrition (a-e): unclear risk<br><br>Unclear if patients with incomplete follow-up were excluded before arriving at the stated sample size. Insufficient information to permit judgement.<br><br>3. Prognostic factor measurement (a-f): low risk<br><br>**Valid and reliable measurement of RD**<br><br>**Outcome level assessment:**<br><br>**Outcome: overall survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Survival was calculated as the number of months from initial surgery to death or the date of last follow-up<br><br>5. Adjustment for other prognostic factors (a-g): low risk<br><br>HR for OS was adjusted for residual disease, age, stage (IIIC and IV versus IIIA/IIIB) and ascites (yes versus no) in a multivariable Cox model<br><br>6. Statistical analysis and reporting (a-d): high risk<br><br>No conceptual framework; unclear of variable selection criteria in multivariate model<br><br>**Outcome: progression-free survival**<br><br>Not reported |
| Notes | Of the 295 women who were treated for FIGO stage III and IV epithelial ovarian cancer at this centre over the period of the study, 13 (5%) were lost to follow-up, and the remaining 282 form the study group for this analysis<br><br>Median follow-up in the study was 32 months (range: 1 to 139 months)<br><br>The chemotherapy was platinum-based and when women who had initially had single agent therapy or combinations with cyclophosphamide recurred they were often given paclitaxel |

**Chi 2001** *(Continued)*

Survival was calculated as the number of months from initial surgery to death or the date of last follow-up.

214 of the 282 (76%) women were dead from disease or other causes at the time of census.

**Multivariate analysis**:

Only women age at diagnosis (P = 0.001), presence of ascites (P = 0.001) and the size of residual disease after primary cytoreductive surgery (1 cm vs 1 cm to 2cm vs > 2 cm (P = 0.02 and 0.001, respectively)) retained prognostic significance

**Kaplan-Meier curve**

Women with no more than 1 cm of residual disease after primary surgery have a 5-year survival of 50% and a median survival of 55 months. There is no statistically significant difference in survival between those women with 1 cm to 2 cm of residual disease and those with greater than 2 cm residual (P = 0.40). This combined group of women have a 5-year survival of 22% with a median survival of 28 months.

**Impact of residual tumour volume for FIGO stage III**

A subgroup analysis of the 216 women with stage III disease was done to examine the impact of size of residual disease on survival

56 of these women had up to 1 cm of residual disease and had 5-year survival of 50% and median survival of 56 months

73 of these women had between 1 cm and 2cm of residual disease and had 5-year survival of 28% and median survival of 31 months

87 of these women had greater than 2 cm of residual disease after surgery and had 5-year survival of 21% and a median survival of 28 months

The differences in survival are statistically significant between the women with up to 1 cm of residual disease and the women in the other 2 groups (P = 0.001). There is no statistically significant difference in survival between the women who had more than 1 cm residual disease.

**Chi 2006**

| ***Study characteristics*** | |
|---|---|
| Methods | Retrospective study |
| Participants | Women with stage IIIC epithelial ovarian cancer<br>The median age at study entry was 60 years (range: 22 to 87)<br>All women presented with FIGO stage IIIC: 465 (100%)<br>Tumour cell type: serous 331 (72%), endometrioid: 57 (12%), clear cell: 22 (5%), mixed: 53 (11%)<br>Tumour grade: 1: 13 (3%), 2: 90 (19%), 3: 339 (73%), unknown: 23 (5%)<br>Ascites: median 1600 mL (range: 0 to 17,000 mL), presence of ascites (N = 429): no = 58 (14%); yes = 371 (86%) |
| Residual disease details | Type of surgeon: gynaecologic oncologist<br>Options for residual disease on the standardised operative form were as follows:<br>1. NMRD: 67 (14.4%)<br>2. Gross residual disease < 0.5 cm: 70 (15.1%) |

**Chi 2006** *(Continued)*

| | |
|---|---|
| | 3. SVRD of 0.6 cm to 1.0 cm: 99 (21.3%)<br>4. LVRD of 1 cm to 2 cm: 53 (11.4%)<br>5. LVRD > 2.0 cm: 176 (37.8%)<br><br>Optimal is defined in 2 ways as NMRD and SVRD (< 1 cm), suboptimal defined as LVRD (> 1 cm)<br><br>Postoperative chemotherapy records were available in 440/465 (95%) women. Of these 440 women, 426 (97%) were treated with primary platinum-based systemic chemotherapy with the intent to treat with at least 6 cycles. |
| Outcomes | Three women (0.6%) died within 30 days of surgery<br><br>Overall survival: HR adjusted for age and ascites using Cox model:<br><br>SVRD (< 1 cm) vs NMRD HR 2.07 (95% CI 1.23 to 3.46)<br><br>LVRD (> 1 cm) vs NMRD HR 3.70 (95% CI 2.27 to 6.04) |
| Risk of bias (QUIPS) | 1. Study participation (a-f): low risk<br><br>Adequate number of participants and description of target population. Baseline characteristics, eligibility criteria, sampling frame and period/place study took place presented clearly.<br><br>2. Study attrition (a-e): unclear risk<br><br>Unclear if patients with incomplete follow-up were excluded before arriving at the stated sample size. Insufficient information to permit judgement.<br><br>3. Prognostic factor measurement (a-f): low risk<br><br>**Valid and reliable measurement of RD**<br><br>**Outcome level assessment:**<br><br>**Outcome: overall survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Valid and reliable measurement of outcome<br><br>5. Adjustment for other prognostic factors (a-g): unclear risk<br><br>HR for OS was adjusted for residual disease, age and ascites in a multivariable Cox model<br><br>6. Statistical analysis and reporting (a-d): high risk<br><br>No conceptual framework; unclear of variable selection criteria in multivariate model<br><br>**Outcome: progression-free survival**<br><br>Not reported |
| Notes | Median follow-up: 38 months (range: 1 to 199 months)<br><br>17-year death rate:<br><br>'Optimal' group: 105/236<br><br>'Suboptimal' group: 188/229<br><br>Median overall survival in relation to the 5 residual disease categories was:<br><br>NMRD: 106 months; gross < 0.5 cm: 66 months; 0.6 cm to 1.0 cm: 48 months; 1 cm to 2 cm: 33 months; and > 2 cm: 34 months |

**Cioffi 2018**

| ***Study characteristics*** | |
|---|---|
| Methods | Single-centre retrospective study |
| Participants | N = 102 participants who received a diagnosis of International Federation of Gynecology and Obstetrics (FIGO) stage IIIC or IV EOC between 2000 and 2016, received neoadjuvant chemotherapy, and presented at least one of the following:<br>• High tumour dissemination (assessed by laparoscopic Fagotti score > 8 or Peritoneal Cancer Index > 15): 83 (81.4%)<br>• Stage IV: 38 (37.3%)<br>• Comorbidities (Charlson comorbidity score ≥ 1): 27 (26.5%)<br>• Poor performance status (ASA score ≥ 3): 58 (56.9%)<br>Participants were stratified according to their age: ≥ 70 vs < 70<br>Age (mean): 74.5 (≥ 70 years) and 58.3 (< 70 years)<br>FIGO: III - 64 (62.7%); IV - 38 (37.3%)<br>Histology: serous - 58 (56.9%); undifferentiated - 1 (1%); endometrioid - 14 (13.7%); sero-endometrioid - 21 (20.6%); clear cell - 3 (2.9%); unknown - 5 (4.9%)<br>Ascites (≥ 500 mL): 76 (74.5%)<br>Tumour grade: G1 - 0; G2 - 8 (7.8%); G3 - 80 (78.4%); unknown - 14 (13.7%)<br>CA-125 at diagnosis (median): 2934.1 (≥ 70 years) and 1462 (< 70 years) |
| Residual disease details | All women received platinum-based regimens, according to standard first-line protocols. After receiving 3 cycles of NACT, women were evaluated by computed tomography (CT) scan or positron emission tomography (PET)–CT scan; radiologic response was assessed according to RECIST 1.1. Women showing complete response (CR) or partial response (PR) to chemotherapy, and considered respectable by a gynaecologic oncologist team, underwent IDS. Women with either stable disease (SD) or progressive disease (PD) after 3 NACT cycles were re-evaluated after 3 further chemotherapy cycles. Women showing CR, PR or SD after 6 chemotherapy cycles underwent debulking surgery.<br>Carboplatin AUC5 and paclitaxel 175 mg/m$^2$ every 3 weeks: 58 (56.9%)<br>Carboplatin AUC5, paclitaxel 175 mg/m$^2$ and bevacizumab (15 mg/kg) on day 1 for 6 x 3-weekly courses followed by bevacizumab single-agent maintenance for 22 cycles or until toxicity or progression: 11 (10.8%)<br>Carboplatin AUC5 every 3 weeks: 25 (24.5%)<br>Carboplatin AUC2 and paclitaxel 60 mg/m$^2$ weekly: 5 (4.9%)<br>Carboplatin AUC2 weekly: 3 (2.9%)<br>Response to NAC (RECIST):<br>• Complete: 35 (34.3%)<br>• Partial: 33 (32.4%)<br>• Stable: 18 17.6%)<br>• Progressive: 13 (12.7%)<br>• Missing: 3 (2.9%)<br>**Optimal cytoreduction** defined as residual disease no greater than 1 cm (RD ≤ 1 cm) (n = 57; 67.1%)<br>• NMRD (described in study as RD0): 37/85 (43.5%)<br>• SVRD: 20 (23.5%)<br>• LVRD (RD > 1): 28 (32.9%) |

**Cioffi 2018** *(Continued)*

| | |
|---|---|
| Outcomes | Overall survival defined as interval from the date of initial diagnosis to the date of death or last follow-up<br><br>Median overall survival: 25 months<br><br>Multivariate Cox PH model for overall survival adjusted for age, number of chemotherapy courses, debulking surgery, ASA score, hypoalbuminaemia (defined as albuminaemia < 32 g/L), FIGO stage, presence of ascites, high tumour dissemination and Charlson comorbidity score:<br><br>• SVRD < 1 cm (including NMRD) (vs LVRD > 1): HR 0.29 (95% CI 0.127 - 0.662), P = 0.003<br><br>Progression-free survival defined as interval from the date of initial diagnosis to the date of first recurrence, death or last follow-up.<br><br>Median progression-free survival: 11 months<br><br>Multivariate Cox PH model for PFS adjusted for age, number of chemotherapy courses, debulking surgery, ASA score, hypoalbuminaemia (defined as albuminaemia < 32 g/L), FIGO stage, presence of ascites ≥ 500 mL, high tumour dissemination and Charlson comorbidity score:<br><br>• SVRD < 1 cm (including NMRD) (vs LVRD > 1 cm): HR 0.43 (95% CI 0.205 to 0.935), P = 0.03 |
| Risk of bias (QUIPS) | 1. Study participation (a-f): low risk<br><br>Adequate number of participants and description of target population. Baseline characteristics, eligibility criteria, sampling frame and period/place study took place presented clearly.<br><br>2. Study attrition (a-e): unclear risk<br><br>Unclear if patients with incomplete follow-up were excluded before arriving at the stated sample size Insufficient information to permit judgement.<br><br>3. Prognostic factor measurement (a-f): low risk<br><br>Valid and reliable measurement of RD<br><br>**Outcome level assessment:**<br><br>**Outcome: overall survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Valid and reliable measurement of outcome. Overall survival defined as interval from the date of initial diagnosis to the date of death or last follow-up.<br><br>5. Adjustment for other prognostic factors (a-g): low risk<br><br>HR for OS was adjusted for residual disease, age, number of neoadjuvant chemotherapy courses, debulking surgery, ASA score, hypoalbuminaemia (defined as albuminaemia < 32 g/L), FIGO stage, presence of ascites ≥ 500 mL, high tumour dissemination and Charlson comorbidity score<br><br>6. Statistical analysis and reporting (a-d): unclear risk<br><br>No conceptual framework; although appears all variables were used in the multivariate models<br><br>**Outcome: progression-free survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Valid and reliable measurement of outcome. Progression-free survival defined as interval from the date of initial diagnosis to the date of first recurrence, death, or last follow-up.<br><br>5. Adjustment for other prognostic factors (a-g): low risk |

**Cioffi 2018** *(Continued)*

| | |
|---|---|
| | HR for PFS was adjusted for residual disease, age, number of neoadjuvant chemotherapy courses, debulking surgery, ASA score, hypoalbuminaemia (defined as albuminaemia < 32 g/L), FIGO stage, presence of ascites ≥ 500 mL, high tumour dissemination and Charlson comorbidity score.<br><br>6. Statistical analysis and reporting (a-d): unclear risk<br><br>No conceptual framework; although appears all variables were used in the multivariate models |
| Notes | ASA score: 1: 5 (4.9%); 2: 36 (35.3%); 3: 51 (50%); 4: 7 (6.9%)<br><br>BMI (mean): 24.4 (≥ 70 years) and 25.5 (< 70 years)<br><br>Charlson comorbidity score ≥1: 27 (26.5%)<br><br>Procedures before NAC: diagnostic laparoscopy: 78 (27.7%); clinical exam/imaging: 196 (69.5%); unknown: 8 (2.8%) |

**Cuylan 2018**

| ***Study characteristics*** | |
|---|---|
| Methods | Retrospective study |
| Participants | 218 women with stage III non-serous EOC<br><br>Median age of women was 54 (range: 18 to 78) years<br><br>Stage, n (%):<br><br>• IIIA1: 55 (25.5%)<br>• IIIA2: 14 (6.4%)<br>• IIIB: 34 (15.6%)<br>• IIIC: 115 (52.8%)<br><br>55 (25.2%) women underwent maximal CRS, 163 (74.8%) had optimal debulking<br><br>Histopathology, n (%): endometrioid 64 (29.4%), mucinous 61 (28%), mixed 39 (17.9%), clear 54 (24.8%)<br><br>Ascites, n (%): present 122 (56%), absent 96 (44%)<br><br>Serum CA 125 (median, IU/mL): ≥ 240 IU/mL 109 (50%), < 240 IU/mL 109 (50%)<br><br>Grade 1: 31 (14.2%), Grade 2: 57 (26.1%), Grade 3: 76 (34.9%)<br><br>Turkey |
| Residual disease details | Speciality of surgeon: gynaecologic oncologist<br><br>All women underwent maximal or optimal primary CRS followed by 6 cycles of carboplatin plus paclitaxel chemotherapy<br><br>Residual disease was noted as follows:<br><br>• NMRD after primary CRS: 55 (25.2%)<br>• 'Optimal' cytoreduction, defined as SMRD (≤ 1 cm): 163 (74.8%) |
| Outcomes | HR for prognostic factors for OS:<br><br>• Age 51 to 69 years vs ≤ 50 years (HR 1.73, 95% CI 1.23 to 2.66)<br>• Age ≤ 50 vs ≥ 70 years (HR 2.6, 95% CI 1.215 to 5.591) |

**Cuylan 2018** *(Continued)*

- NMRD (HR 0.31, 95% CI 0.166 to 0.615)

HR for prognostic factors for PFS:

- Bilaterality (HR 1.44, 95% CI 1.01 to 2.056)
- Age (HR 2.25, 95% CI 1.176 to 4.323)
- NMRD (HR 0.34, 95% CI 0.202 to 0.58)

| | |
|---|---|
| Risk of bias (QUIPS) | 1. Study participation (a-f): low risk<br><br>Adequate number of participants and description of target population. Baseline characteristics, eligibility criteria, sampling frame and period/place study took place presented clearly.<br><br>2. Study attrition (a-e): unclear risk<br><br>Unclear if patients with incomplete follow-up were excluded before arriving at the stated sample size. Insufficient information to permit judgement.<br><br>3. Prognostic factor measurement (a-f): unclear risk<br><br>Valid and reliable measurement of RD. Multicentre design may introduce heterogeneity in measurement of RD<br><br>**Outcome level assessment:**<br><br>**Outcome: overall survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Valid and reliable measurement of outcome<br><br>5. Adjustment for other prognostic factors (a-g): low risk<br><br>HR for OS was adjusted for age, maximal cytoreduction and stage in multivariable Cox model<br><br>6. Statistical analysis and reporting (a-d): high risk<br><br>No conceptual framework; unclear of variable selection criteria in multivariate model<br><br>**Outcome: progression-free survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Valid and reliable measurement of outcome<br><br>5. Adjustment for other prognostic factors (a-g): low risk<br><br>HR for PFS was adjusted for age, maximal cytoreduction and stage in multivariable Cox model<br><br>6. Statistical analysis and reporting (a-d): high risk<br><br>No conceptual framework; unclear of variable selection criteria in multivariate model |
| Notes | Median duration of follow-up was 31.5 (range: 1 to 20) months<br><br>5-year PFS rate was 34.8%<br><br>5-year OS rate was 44.2%, median OS was 47 months (95% CI 36.12 to 57.88)<br><br>A univariate analysis showed an OS rate of 81.2% the maximal CRS group<br><br>Status: alive 109 (50%); dead 109 (50%) |

**Davidson 2019**

| *Study characteristics* | |
|---|---|
| Methods | Multicentre retrospective and single-centre prospective cohort<br><br>Prospective data collection was to explore minimally-invasive surgery following NACT |
| Participants | All participants received NACT followed by interval debulking surgery for an advanced ovarian, fallopian tube or primary peritoneal cancer<br><br>At Duke, information on women receiving NACT was collected retrospectively between January 2000 and September 2013 and prospectively (with subject informed consent after October 2013). At the Ohio State University and the University of Oklahoma, subjects were identified retrospectively. Women at all 3 institutions were included if they were diagnosed prior to 30 June 2017 to allow for at least 12 months of post-diagnosis follow-up.<br><br>N = 282 participants with advanced ovarian, fallopian tube or primary peritoneal cancer<br><br>Median age: 63.9 (range: 34.1 to 84.8)<br><br>Race: Caucasian – 229 (81.2%)<br><br>FIGO: IIIC – 114 (40.4%); IV – 101 (35.8%); presumed AOC – 57 (20.2%); unknown stage – 10 (3.5%)<br><br>Histology: serous – 227 (80.5%); undifferentiated – 4 (1.5%); endometrioid – 1 (0.4%); mixed – 5 (1.8%); clear cell – 5 (1.8%); NOS – 21 (7.5%); unknown – 15 (5.3%)<br><br>Ascites: 88 (31.2%) |
| Residual disease details | Carboplatin and paclitaxel: 87.2%<br><br>Median NACT cycles: 4 (range: 2 to 10)<br><br>Indication for NACT: disease volume - 80 (28.4%); comorbidities - 19 (6.7%); both - 29 (10.3%)<br><br>Median surgery duration, minutes: 194 (range: 45 to 459)<br><br>Determination of resectability: diagnostic laparoscopy – 78 (27.7%); clinical exam/imaging – 196 (69.5%); unknown – 8 (2.8%)<br><br>Surgical approach at IDS: laparoscopy only – 27 (9.6%); laparoscopy converted to laparotomy – 26 (9.2%); exploratory laparotomy only – 221 (78.4%)<br><br>Median surgical complexity score: 2<br><br>Surgical complexity score:<br><br>• Low (0 to 3): 193 (68.4%)<br>• Moderate (4 to 7): 80 (28.4%)<br>• Complex (8 to 9): 9 (3.2%)<br><br>Intraoperative complications: 23 women (8.7%). Bowel injuries (including serosal injuries) (n = 16); bladder (n = 6); vascular injuries (n = 6).<br><br>Postoperative complications were seen in 62 women (22%) prior to hospital discharge and included:<br><br>• Ileus/small bowel obstruction: 26 (9.2%)<br>• Pulmonary issues: 12 (4.3%)<br>• Altered mental state: 10 (3.6%)<br>• Wound cellulitis/haematoma, UTI and cardiac concerns: 5 (1.8%)<br>• Re-operation: 1 (0.4%)<br><br>32 (11.3%) experienced complications after discharge and within 30 days of surgery |

**Davidson 2019** *(Continued)*

| | |
|---|---|
| | 18 (6.4%) re-admitted. Data for reasons for re-admission available for n = 7: infectious complications (n = 3), gastrointestinal dysmotility (n = 3), acute renal failure related to urinary retention (n = 1). 2 required re-operation during re-admission. 1 underwent re-operation in outpatient setting for wound debridement.<br><br>**Optimal cytoreduction** defined using two methods: NMRD (described in study as RD0) (n = 165/271; 60.9%) or SVRD ≤ 1 (n = 228/271; 84.1%). The latter definition is used in multivariable analysis.<br><br>• NMRD: 165 (60.9%)<br>• SVRD: 63 (23.2%)<br>• LVRD 1 cm to 2 cm: 6 (2.2%)<br>• LVRD > 2 cm: 37 (13.7%)<br>• Missing (n = 11) |
| Outcomes | Disease-specific overall survival (DSS) defined as time from completion of adjuvant chemotherapy to death due to cancer<br><br>Median disease-specific overall survival (DSS): 24.8 months<br><br>Median DSS in RD ≤ 1: 25 months<br><br>Median DSS in RD > 1: 23.5 months<br><br>Multivariable Cox PH for DSS adjusted for ASA score, age, SCS and major morbidity:<br><br>• LVRD > 1 cm (vs SVRD ≤ 1 cm): HR 1.7 (95% CI 1.1 to 2.8), P = 0.03<br><br>No deaths within 30 days of IDS |
| Risk of bias (QUIPS) | 1. Study participation (a-f): low risk<br><br>Adequate number of participants and description of target population. Baseline characteristics, eligibility criteria, sampling frame and period/place study took place presented clearly.<br><br>2. Study attrition (a-e): unclear risk<br><br>Unclear if patients with incomplete follow-up were excluded before arriving at the stated sample size. Insufficient information to permit judgement.<br><br>3. Prognostic factor measurement (a-f): unclear risk<br><br>Adequate cut-off for residual disease used. Multicentre design may introduce heterogeneity in measurement of RD.<br><br>**Outcome level assessment:**<br><br>**Outcome: disease-specific survival**<br><br>4. Outcome measurement (a-c): high risk<br><br>Overall survival not used as outcome. Rather, disease-specific survival was used. Disease-specific survival (DSS) defined as time from completion of adjuvant chemotherapy to death due to cancer.<br><br>5. Adjustment for other prognostic factors (a-g): high risk<br><br>Age arbitrarily categorised; ASA score dichotomised. Model predicting DSS adjusted for ASA score, age, SCS, presence of major morbidity. Few of these were deemed important prognostic factors.<br><br>6. Statistical analysis and reporting (a-d): unclear risk<br><br>No conceptual framework; data driven based on P values of univariate associations<br><br>**Outcome: progression-free survival** |

**Davidson 2019** *(Continued)*

| | |
|---|---|
| | Not reported |
| Notes | — |

**Eisenkop 2003**

| ***Study characteristics*** | |
|---|---|
| Methods | This is a prospective study of women with FIGO stage IIIC ovarian cancer treated with primary cytoreductive surgery followed by platinum-based chemotherapy between 1990 and 2002 at a single North American institution |
| Participants | 408 consecutive women presenting with stage IIIC epithelial ovarian cancer form the study group<br><br>The median age at study entry was 62.8 years (range: 24 to 91)<br><br>All women presented with FIGO stage IIIC epithelial ovarian cancer: 408 (100%)<br><br>Tumour cell type: serous: 239 (58.5%), unspecified adenocarcinoma: 98 (24%), endometrioid: 32 (8%), clear cell: 10 (2.5%), mucinous: 18 (4.5%), mixed: 9 (2%), transitional cell: 2 (0.5%)<br><br>Tumour grade: 1: 21 (5%), 2: 82 (20%), 3: 304 (75%), unspecified: 1 woman<br><br>Volume of ascites: none: 20 (5%), ≤ 1000 mL: 114(28%), > 1000 mL: 249(61%), not recorded: 24(6%)<br><br>GOG performance score: 0: 17 (4%), 1: 88 (21.5%), 2: 177 (43.5%), 3: 59 (14.5%), 4: 2 (0.5%), unspecified: 65 (16%)<br><br>Preoperative tumour volume:<br><br>Location of the largest metastases: omentum and adjacent structures: 228 (56%), pelvis: 102 (25%), retroperitoneal lymph nodes: 34 (8%), diaphragm: 12 (3%), other (large bowel, small bowel, mesentery, etc): 32 (8%)<br><br>Largest metastatic disease: < 10 cm: 104 (26%), > 10 cm: 302 (74%) |
| Residual disease details | Residual disease was noted as follows:<br><br>1. NMRD: 351 (86%)<br>2. SVRD: 41 (10%)<br>3. LVRD (> 1 cm): 16 (6%)<br><br>Surgery was undertaken by a gynaecological oncologist and disease was assessed intraoperatively in each of the following 5 regions: the left and right upper abdominal quadrants, the pelvis, the retroperitoneum and the central abdomen. A specifically defined numerical rank of 0 to 3 was assigned to each of the 5 regions and the ranks for each of the 5 regions were summed to give a total score before cytoreduction.<br><br>'Optimal' cytoreduction was defined as complete cytoreduction with no visible residual disease. The authors have previously described in other publications how this can be achieved at different anatomical sites but recourse to bowel resection was routine as was pelvic and para-aortic nodal dissection.<br><br>Postoperative chemotherapy was platinum-based: cisplatin (50 to 100 mg/m$^2$) or carboplatin (300 to 400 mg/m$^2$) given in combination therapy with either cyclophosphamide or paclitaxel every 3 weeks for a planned 6 to 8 cycles. |
| Outcomes | Overall survival: HR adjusted for sum of rankings (a numerical ranking system was devised to reflect the continuum of progressively extensive tumour involvement for 5 anatomic regions) using a Cox model:<br><br>SVRD vs NMRD: HR 2.32 (95% CI 1.20 to 5.37) |

**Eisenkop 2003** *(Continued)*

| | |
|---|---|
| | LVRD (> 1 cm) vs NMRD HR: 2.98 (95% CI 1.74 to 5.23)<br><br>**Direct surgical morbidity and mortality**<br><br>Postoperative mortality occurred in 10 (2.5%) women<br><br>Other morbidity including surgically related systemic morbidity such as chest infection, thromboembolic disease and cardiovascular events have not been reported<br><br>**Recovery**<br><br>The median length of hospital stay was 10 days |
| Risk of bias (QUIPS) | 1. Study participation (a-f): low risk<br><br>Adequate number of participants and description of target population. Baseline characteristics, eligibility criteria, sampling frame and period/place study took place presented clearly.<br><br>2. Study attrition (a-e): unclear risk<br><br>Unclear if patients with incomplete follow-up were excluded before arriving at the stated sample size. Insufficient information to permit judgement.<br><br>3. Prognostic factor measurement (a-f): low risk<br><br>Valid and reliable measurement of RD<br><br>**Outcome level assessment:**<br><br>**Outcome: overall survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Valid and reliable measurement of outcome. Survival was measured in months from the date of primary surgery to the time of death or last follow-up appointment using life table analysis.<br><br>5. Adjustment for other prognostic factors (a-g): high risk<br><br>HR for OS was adjusted for residual disease and sum of rankings (a numerical ranking system was devised to reflect the continuum of progressively extensive tumour involvement for 5 anatomic regions) in a multivariable Cox model<br><br>6. Statistical analysis and reporting (a-d): high risk<br><br>No conceptual framework; unclear of variable selection criteria in multivariate model<br><br>**Outcome: progression-free survival**<br><br>Not reported |
| Notes | The median follow-up interval was 32.8 months<br><br>Survival was measured in months from the date of primary surgery to the time of death or last follow-up appointment using life table analysis. Survival outcomes were analysed based on the numerical ranking of disease in each anatomical region, the sum of the ranking and the cytoreductive outcome.<br><br>The median survival was 58.2 months (24% to 91%) and the estimated 5-year survival was 49%<br><br>**Ranking of disease load**<br><br>349 (85.5%) of women had ranking in all 5 designated regions. Ranking was not possible in the rest because lymph node dissection was deferred in 48 women (12%) or the pattern of spread was inconsistent with ranking criteria in 16 women (4%). |

**Eisenkop 2003** *(Continued)*

On univariate analysis, categorisation of the sum of ranking scores (0 to 5 vs 6 to 10, vs ≥ 11), as well as ranking in the left upper abdominal quadrant and in the central abdomen were statistically important determinants of survival.

Univariate analysis showed that any rank score over zero (any disease) in the left upper abdominal quadrant (P = 0.01) and in the central abdominal region (P = 0.04) adversely affected survival. An effect of the anatomical site of disease on survival was not confirmed on multivariate analysis.

On multivariate analysis, survival was most influenced by the completeness of cytoreduction (P = 0.001), and less influenced by the categorised sum of rankings (P = 0.05).

This study demonstrates that high rates of complete cytoreduction can be achieved within dedicated teams with suitable training. The independent effect of completeness of cytoreduction on survival is confirmed though the median length of follow-up in the report is modest.

**Feng 2016**

| ***Study characteristics*** | |
|---|---|
| Methods | Retrospective study |
| Participants | 625 women who underwent primary staging or debulking surgery for high-grade serous ovarian cancer (HGSC)<br><br>Age at diagnosis, median (range), years: 56 (30 to 84)<br><br>FIGO stage: early (I,II) - 58 (9.3%); advanced (III, IV) - 567 (90.7%)<br><br>Performance status: 0 to 379 (60.6%); 1 to 202 (32.3%); 2 to 44 (7.0%)<br><br>132 (21.1%) underwent bowel resection; 91 (14.6%) underwent upper abdominal surgery; 104 (16.6%) underwent lymphadenectomy<br><br>CA-125: < 500 U/mL - 144 (23.6%); ≥ 500 U/mL - 465 (76.5%)<br><br>Ascites: no - 75 (12%); < 500 mL - 104 (16.7%); ≥ 500 mL - 445 (71.3%)<br><br>China |
| Residual disease details | Speciality of surgeon not reported<br><br>After primary cytoreduction, all women received platinum-based intravenous chemotherapy<br><br>Chemotherapy regimen:<br><br>• Paclitaxel + carboplatin - 518 (82.9%)<br>• Other platinum and taxane agents - 91 (14.6%)<br>• Platinum and other agents - 16 (2.6%)<br><br>Majority (441, 70.6%) of women had completed 6 to 8 cycles at intervals of 3 weeks<br><br>R0 was defined as NMRD after surgery and was noted as follows:<br><br>• No - 209 (33.4%)<br>• Yes - 416 (66.6%) |
| Outcomes | PFS was defined as the time interval from the date of primary surgery to the date of disease progression or recurrence<br><br>Median PFS was 18 months; 2-year PFS was 38.4%; 5-year PFS was 21.4% |

**Feng 2016** *(Continued)*

| | |
|---|---|
| | OS was defined as the time interval from the date of the primary surgery to the date of death or last follow-up<br><br>2-year OS was 82.5%; 5-year OS was 51.4%<br><br>At the time of analysis, 355 (56.8%) women were still alive |
| Risk of bias (QUIPS) | 1. Study participation (a-f): low risk<br><br>Adequate number of participants and description of target population. Baseline characteristics, eligibility criteria, sampling frame and period/place study took place presented clearly.<br><br>2. Study attrition (a-e): unclear risk<br><br>Unclear if patients with incomplete follow-up were excluded before arriving at the stated sample size. Insufficient information to permit judgement.<br><br>3. Prognostic factor measurement (a-f): low risk<br><br>Valid and reliable measurement of RD<br><br>**Outcome level assessment:**<br><br>**Outcome: overall survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Valid and reliable measurement of OS. OS was defined as the time interval from the date of the primary surgery to the date of death or last follow-up<br><br>5. Adjustment for other prognostic factors (a-g): unclear risk<br><br>Multivariate models for OS adjusted for age, FIGO stage and time to chemotherapy<br><br>6. Statistical analysis and reporting (a-d): high risk<br><br>No conceptual framework; unclear of variable selection strategy into multivariate model. Unclear on reasoning behind inclusion of other prognostic factors in Cox models.<br><br>**Outcome: progression-free survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Valid and reliable measurement of PFS; PFS was defined as the time interval from the date of primary surgery to the date of disease progression or recurrence<br><br>5. Adjustment for other prognostic factors (a-g): unclear risk<br><br>Multivariate models for PFS adjusted for age, FIGO stage and time to chemotherapy.<br><br>6. Statistical analysis and reporting (a-d): high risk<br><br>No conceptual framework; unclear of variable selection strategy into multivariate model. Unclear on reasoning behind inclusion of other prognostic factors in Cox models. |
| Notes | The median (range) follow-up time was 29 (3 to 100) months<br><br>The median (range) of time to chemotherapy (TTC) was 15 (4 to 62) days. TTC was longer for women who underwent bowel resection (P < 0.001). There were no differences in PFS and OS between women initiating chemotherapy before and after 15 days (P = 0.604 and 0.826 respectively) or among 4 groups categorised by quartile values (< 10 days, 10 to 14 days, 15 to 20 days, or ≥ 21 days after surgery) (P = 0.471 and 0.516, respectively). The time interval between surgery and chemotherapy seemed to have no prognostic impact on women with HGSC within 6 weeks.<br><br>Length of hospital stay not reported |

**Hofstetter 2013**

| ***Study characteristics*** | |
|---|---|
| Methods | Prospective multicentre study |
| Participants | 191 women with stage IIIA to IV primary ovarian cancer. Stage IIIa: 3, IIIb: 8, IIIc: 147, IV: 33<br>ECOG performance status (only available for 183 women) 0: 113, 1: 60; 2/3: 10<br>Age < 57: 98, > 57: 93<br>Histological subtypes; serous: 182, mixed serous:1, serous/clear cell: 4, undifferentiated: 4<br>Tumour grade 1/2: 51, 3: 140 |
| Residual disease details | All women underwent primary surgery. All women received postoperative intravenous or intraperitoneal platinum-based chemotherapy.<br>Women that received neoadjuvant chemotherapy were excluded<br>**Postoperative residual disease defined as**<br>• NMRD (n = 121)<br>• Macroscopic or 'suboptimal' if residual tumour lesions of any size or number (n = 70) |
| Outcomes | Median follow-up was 42 months<br>3-year OS: HR of NMRD vs macroscopic RD: 2.95 (95% CI 1.87 to 4.67)<br>HR adjusted for interval between surgery and start of chemotherapy, tumour stage, age and extent of surgery<br>**Morbidity**<br>Intraoperative complications included bladder injury (2), ureteral injury (1), intestinal injury (1), vascular injury (2), other operative injury (1). 9 of 185 women required blood transfusions. Postoperative complications comprised surgical site complications (35), medical complications (42), infectious complications (22) and reoperation's (22).<br>**Adjuvant chemotherapy**<br>• Intravenous carboplatin/taxane 1 cycle (3), 3 cycles (3), 4 cycles (6), 5 cycles (9), 6 cycles (139), 7 cycles (5), 8 cycles (4), 9 cycles (1)<br>• Intraperitoneal platinum/taxane (13)<br>• 9 women had single agent carboplatin: 2 cycles (1), 3 cycles (1), 6 cycles (7)<br>• 1 women received carboplatin/liposomal doxorubicin |
| Risk of bias (QUIPS) | 1. Study participation (a-f): unclear risk<br>Adequate number of participants and description of target population. Baseline characteristics, sampling frame and period/place study took place presented clearly. Though, inclusion criteria not detailed.<br>2. Study attrition (a-e): unclear risk<br>Unclear if patients with incomplete follow-up were excluded before arriving at the stated sample size. Insufficient information to permit judgement.<br>3. Prognostic factor measurement (a-f): low risk<br>Valid and reliable measurement of RD |

**Hofstetter 2013** *(Continued)*

| | |
|---|---|
| | **Outcome level assessment:**<br><br>**Outcome: overall survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Valid and reliable measurement of outcomes<br><br>5. Adjustment for other prognostic factors (a-g): unclear risk<br><br>Interval from primary surgery to chemotherapy (continuous) arbitrarily dichotomised along the median. Multivariate model predicting OS adjusted for interval from surgery to chemotherapy, FIGO stage, age and extent of surgery<br><br>6. Statistical analysis and reporting (a-d): high risk<br><br>No conceptual framework; variable selection strategy into multivariate model unclear. HRs for centre not included in the results for multivariate analysis. There were other factors that were also significant at univariate analysis but were not included in multivariate model.<br><br>**Outcome: progression-free survival**<br><br>Not reported |
| Notes | The median time interval from primary surgery to the start of platinum-based chemotherapy was 28 days (range: 4 to 128). Women who received the first cycle of chemotherapy less than 28 days after surgery had a significantly improved 3-year survival rate of 70% as opposed to 60% in women with a later start of cytotoxic treatment. |

**Iwase 2015**

| ***Study characteristics*** | |
|---|---|
| Methods | Single-centre retrospective analysis of medical records |
| Participants | N = 124 women with advanced EOC who received NACT-IDS therapy at the Cancer Institute Hospital (Tokyo, Japan) between 2000 and 2008.<br><br>Median age: 58 (range: 29 to 83)<br><br>FIGO: IIIB – 6 (4.8%); IIIC – 77 (62.1%); IV – 41 (33.1%)<br><br>Histology: serous – 105 (84.6%); mixed adenocarcinoma or carcinosarcoma included serous component – 10 (8.1%); non-serous – 9 (7.3%)<br><br>Median CA-125 at pre-NACT, U/mL: 1569.4 (range: 13.5 to 24821)<br><br>Median CA-125 post-NACT, U/mL: 15.8 (range: 2.3 to 1965.1)<br><br>Lymph node metastasis: positive – 49 (39.5%); negative – 41 (33.1%); not evaluated – 34 (27.4%) |
| Residual disease details | Strategy for NACT-IDS therapy consisted of intensive chemotherapy (6 or more cycles) aimed at complete resection during IDS and pathological complete response followed by maximum debulking surgery included systematic retroperitoneal lymphadenectomy in principle. After about 6 cycles of NACT, we then performed IDS unless the disease had progressed. After IDS, ACT was generally administered for about 3 cycles using the same regimen. However, some women did not receive 3 cycles of ACT due to having undergone intensive chemotherapy before surgery or having undergone highly invasive surgery. Conversely, more than 3 cycles of ACT were necessary in the case of some women for whom complete resection was not achieved.<br><br>Method to diagnose: laparotomy - 62 (50%); non-laparotomy - 62 (50%) |

**Iwase 2015** *(Continued)*

| | |
|---|---|
| | Median NACT cycles: 6 (range: 2 to 9)<br><br>NACT regimen: ifosfamide, epirubicin and cisplatin (IEP) including cyclophosphamide, adriamycin and cisplatin (CAP) – 44 (35.5%); paclitaxel and carboplatin (TC) including docetaxel and carboplatin (DC) – 80 (64.5%); irinotecan (CPT) base – 3 (2.4%)<br><br>Surgical procedure at IDS: exploratory laparotomy – 11 (8.9%); total abdominal hysterectomy, bilateral salpingo-oophorectomy, and omentectomy (TAH + BSO + OM) – 10 (8.1%); TAH + BSO + OM + excision of other organs – 17 (13.7%); TAH + BSO + OM + retroperitoneal lymphadenectomy – 48 (38.7%); TAH + BSO + OM + excision of other organs + retroperitoneal lymphadenectomy – 38 (30.6%)<br><br>Median operative blood loss, mL: 1291 (range: 220 to 5640)<br><br>Blood transfusion: 72 women (70.6%)<br><br>Median adjuvant CT cycles: 3 (range: 1 to 8)<br><br>ACT regimen: ifosfamide, epirubicin and cisplatin (IEP) including cyclophosphamide, adriamycin and cisplatin(CAP) – 25 (20.2%); paclitaxel and carboplatin (TC) including docetaxel and carboplatin (DC) – 65 (52.4%); docetaxel and cisplatin (DP) including docetaxel (DTX) – 22 (17.7%); others – 7 (5.6%)<br><br>**'Optimal' cytoreduction** defined as SVRD < 1 cm (n = 113; 91.1%)<br><br>• NMRD: 98 (79%)<br>• SVRD (RD < 1): 15 (12.1%)<br>• LVRD (RD ≥ 1): 11 (8.9%)<br><br>* Note: in multivariable analysis, it is RD > 0 cm vs NMRD |
| Outcomes | 2-year OS: NMRD (88.8%); SVRD (40%); LVRD (≥ 1 cm) (36.3%)<br><br>5-year OS: NMRD (43.4%); SVRD (0%); LVRD (≥ 1 cm) (0%)<br><br>Multivariable Cox PH for overall survival adjusted for FIGO stage, histological subtype, NACT cycles, NACT regimen, systematic lymphadenectomy, excision of other organ(s), ascites cytology, lymph node metastasis:<br><br>• RD > 0 cm (vs NMRD): HR 4.03 (95% CI 2.39 to 7.16), P < 0.001<br><br>2-year PFS: NMRD (39.8%); SVRD (< 1 cm) (13.3%); LVRD (≥ 1) (0%) |
| Risk of bias (QUIPS) | 1. Study participation (a-f): unclear risk<br><br>Number of participants below the minimum cutoff of n = 100 for this meta-analysis. Adequate description of target population. Baseline characteristics, eligibility criteria, sampling frame and period/place study took place presented clearly.<br><br>2. Study attrition (a-e): unclear risk<br><br>Unclear if patients with incomplete follow-up were excluded before arriving at the stated sample size. Insufficient information to permit judgement.<br><br>3. Prognostic factor measurement (a-f): low risk<br><br>Valid and reliable measurement of RD<br><br>**Outcome level assessment:**<br><br>**Outcome: overall survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Definition of OS not provided but it usually has a standard definition<br><br>5. Adjustment for other prognostic factors (a-g): low risk |

**Iwase 2015** *(Continued)*

| | |
|---|---|
| | Adjustment for large number of important PFs (FIGO stage, histological subtype, NACT cycles, NACT regimen, systematic lymphadenectomy, excision of other organ(s), ascites cytology, lymph node metastasis)<br><br>6. Statistical analysis and reporting (a-d): high risk<br><br>No conceptual framework; criteria for variable selection for univariate and multivariate Cox PH for OS unspecified<br><br>**Outcome: progression-free survival**<br><br>Progression-free survival mentioned in methods but not reported in results |
| Notes | Median follow-up, months: 39.5 (range: 5 to 142)<br><br>Exclusion criteria: synchronous or metachronous (within 5 years) malignancies other than carcinoma in situ, missing data because women were referred to a different institution for initial treatment, received only palliative therapy after exploratory laparotomy, stage III disease without macroscopic peritoneal dissemination (e.g. pT1N1, pT2N1, pT3aN0 and pT3aN1), and received PDS-ACT therapy as initial treatment.<br><br>Finally, excluding women who were not able to undergo IDS because of disease progression during NACT. |

**Kaban 2017**

| ***Study characteristics*** | |
|---|---|
| Methods | Single-centre retrospective analysis of medical records |
| Participants | N = 203 women diagnosed with stage IIIC to IV ovarian, fallopian tube or primary peritoneal cancer (according to postoperative pathology reports) who underwent treatment with interval surgery after NACT at the Istanbul University Gynecological Oncology Department between January 2002 and December 2012.<br><br>Median age: 59 (range: 28 to 84)<br><br>FIGO staging not reported<br><br>Histology: serous – 171 (84.2%); undifferentiated – 1 (0.4%); endometrioid – 2 (0.9%); carcinosarcoma – 7 (3.4%); mixed – 2 (0.9%); clear cell – 4 (1.9%); mesothelioma – 2 (0.9%); Brenner tumour – 1 (0.4%); missing – 10 (4.9%)<br><br>Visible tumour in diaphragm/liver: 29 (14.3%)<br><br>Presence of tumour in omentum: macroscopic – 144 (70.9%); tumour-free – 44 (21.6%); no macroscopic – 14 (6.9%); missing – 1<br><br>Median lymph node count 10 (range: 2 to 24)<br><br>Nodal metastasis: 3 |
| Residual disease details | NAC consisted of a carbo-platinum (area under the curves 5 to 6) and paclitaxel (135 to 175 mg/m2) regimen every 3 weeks<br><br>Median NACT cycles: 6 (range: 1 to 10)<br><br>Pelvic +/- para-aortic lymphadenectomy performed in n = 25 women (12.3%)<br><br>Extra-surgical procedure: bowel resection (n = 4); splenectomy (n = 1) |

**Kaban 2017** *(Continued)*

| | |
|---|---|
| | Intraperitoneal port placement: 13 (6.4%)<br><br>After surgery, all women continued chemotherapy with 2 to 6 additional cycles<br><br>**'Optimal' cytoreduction** defined as SVRD:<br><br>• SVRD (RD ≤ 1): 165 (81.3%)<br>• LVRD (RD > 1): 36 (17.9%)<br>• Missing (n = 2) |
| Outcomes | Overall survival (OS) was defined as the time from initial treatment to death or to the last follow-up examination.<br><br>5-year OS: 33.4%<br><br>Median OS: 37.5 months<br><br>Median OS in RD ≤ 1 cm: 40.6 months<br><br>Median OS in RD > 1 cm: 21.3 months<br><br>Multivariable Cox PH for OS adjusted for age, lymphadenectomy, macroscopic tumour in omentum, number of chemotherapy cycles:<br><br>• LVRD (> 1 cm) (vs SVRD): HR 1.629 (95% CI 1.024 to 2.593), P = 0.039 |
| Risk of bias (QUIPS) | 1. Study participation (a-f): unclear risk<br><br>Adequate number of participants and description of target population. Baseline characteristics, sampling frame and period/place study took place presented clearly. Though, inclusion criteria not detailed.<br><br>2. Study attrition (a-e): unclear risk<br><br>Unclear if patients with incomplete follow-up were excluded before arriving at the stated sample size. Insufficient information to permit judgement.<br><br>3. Prognostic factor measurement (a-f): low risk<br><br>Valid and reliable measurement of RD; 201 (99%) with available RD data<br><br>**Outcome level assessment:**<br><br>**Outcome: overall survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Valid and reliable measurement of OS which was defined as the time from initial treatment to death or to the last follow-up examination.<br><br>5. Adjustment for other prognostic factors (a-g): unclear risk<br><br>Number of chemotherapy cycles dichotomised along arbitrary cut-off. Model predicting OS adjusted for age, lymphadenectomy, macroscopic tumour in omentum, number of chemotherapy cycles. Inclusion of other important PFs in model may alter results.<br><br>6. Statistical analysis and reporting (a-d): high risk<br><br>No conceptual framework; unclear how variables selected for multivariate models. But age was included even though it was not significant at univariate, suggesting some assessment of clinical judgment in selection of important PFs.<br><br>**Outcome: progression-free survival**<br><br>Not reported |

**Kaban 2017** *(Continued)*

| | |
|---|---|
| Notes | Median follow-up, months: 34.5 (range: 1 to 124) |

**Kahl 2017**

| ***Study characteristics*** | |
|---|---|
| Methods | Retrospective, multicentre cohort study |
| Participants | 793 women with FIGO stage IIIB to IV<br>Median age, years (range) (% < 55 years): 60 (19 to 88)<br>ECOG performance status (PS): 0 to 683 (86.1%); > 0 to 110 (13.9%)<br>FIGO stages, n (%):<br>• IIIB - 110 (13.9%)<br>• Stage IIIC - 318 (40.1%)<br>• Stage IV - 365 (46.0%)<br>Ascites, mL: ≤ 500 to 450 (56.7%); > 500 to 343 (43.3%)<br>Histology: high-grade serous - 660 (83.2%); others - 133 (16.8%)<br>Surgical complexity score: low/intermediate (≤ 7) - 165 (20.8%); high (≤ 8) - 628 (79.2%)<br>Lymph node dissection: systematic - 472 (59.5%); sampling - 111 (14%); no - 210 (26.5%)<br>CDC: 0 to 2 - 593 (74.8%); 3 to 4 - 176 (22.1%); 5 - 24 (3.0%)<br>Germany |
| Residual disease details | Procedure performed by accredited gynaecological oncologist<br>All women underwent primary cytoreductive surgery followed by postoperative systemic therapy with platinum-based chemotherapy<br>Residual disease was noted as follows, n (%):<br>• NMRD: 482 (60.8%)<br>• SVRD (1 mm to 10 mm): 226 (28.5%)<br>• LVRD (> 10 mm): 85 (10.7%)<br>Women were divided into 3 groups based on their age-adjusted Charlson Comorbidity Index (ACCI): low (0 to 1), intermediate (2 to 3), and high (≥ 4)<br>Postoperative surgical complications were graded according to the Clavien-Dindo classification (CDC) |
| Outcomes | Multivariate analysis of prognostic factors for OS:<br>Residual disease (versus NMRD):<br>• SVRD (1 mm to 10 mm): HR 1.96 (95% CI 1.55 to 2.46)<br>• LVRD (> 10 mm): HR 2.75 (95% CI 2.01 to 3.77)<br>Multivariate analysis of prognostic factors for high complications (CDC 3 to 5):<br>• Surgical complexity score: high (≤ 8): RR 1.70 (95% CI 1.01 to 2.85)<br>• Blood loss: ≥ 500: RR 1.0 (95% CI 0.64 to 1.44) |

**Kahl 2017** *(Continued)*

| | |
|---|---|
| | • Duration of surgery, minutes: ≥ 360: RR 1.84 (95% CI 1.24 to 2.72) |
| Risk of bias (QUIPS) | 1. Study participation (a-f): low risk<br><br>Adequate number of participants and description of target population. Baseline characteristics, eligibility criteria, sampling frame and period/place study took place presented clearly.<br><br>2. Study attrition (a-e): unclear risk<br><br>Unclear if patients with incomplete follow-up were excluded before arriving at the stated sample size. Insufficient information to permit judgement.<br><br>3. Prognostic factor measurement (a-f): unclear risk<br><br>Valid and reliable measurement of RD. Multicentre design may introduce heterogeneity in measurement of RD<br><br>**Outcome level assessment:**<br><br>**Outcome: overall survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Valid and reliable measurement of outcome<br><br>5. Adjustment for other prognostic factors (a-g): unclear risk<br><br>Ascites dichotomised along arbitrary cutoff of 500 mL. Multivariate model predicting OS adjusted for ACCI, ECOG, FIGO stage, histology and ascites.<br><br>6. Statistical analysis and reporting (a-d): high risk<br><br>No conceptual framework; criteria for variable selection into multivariate model is unclear. Dichotomisation of continuous variables also apparent.<br><br>**Outcome: progression-free survival**<br><br>Not reported |
| Notes | After a median follow-up was 47 months (interquartile range 18 to 87 months), 397 (50.1%) women had died.<br><br>Significant differences between the 3 ACCI groups were detected for performance status (ECOG 0: 95.7% vs 84.2% vs 65.9%) and residual disease (NMRD 70.7% vs 55.3% vs 49.6%).<br><br>Residual disease after debulking surgery was significantly more frequent in women with a high ACCI compared with women with an intermediate or low ACCI (50.4% vs 44.7% vs 29.3%)<br><br>The mortality rate in the low-ACCI group was 1.2%, in the intermediate-ACCI group it was 2.3% and it was 9.8% for the high-ACCI group |

**Klar 2016**

| ***Study characteristics*** | |
|---|---|
| Methods | Retrospective analysis of primary trials |
| Participants | 5055 participants with stages I to IV ovarian cancer from AGO Study groups were included in Klar 2016. A total of 4488/5130 (87.5%) were stage III/IV in the 4 reported trials that were included in Klar 2016 and n = 4850 were included in the RD analysis.<br><br>AGO-OVAR 3 trial: n = 798 |

**Klar 2016** *(Continued)*

| | |
|---|---|
| | • FIGO stage IIIA to IV: 717/798 (89.85%)<br>• Postoperative residual tumour size, n (%): unknown: 4; ≤ 1 cm: 488 (62.6%); > 1 cm: 291 (37.4%)<br><br>AGO-OVAR 5 trial: n = 1308<br><br>• FIGO stage IIIA to IV: 1191/1308 (91.06%)<br>• Postoperative residual tumour size, n (%): unknown: 122; ≤ 1 cm: 799 (67.4%); > 1 cm: 387 (32.6%)<br><br>AGO-OVAR 7 trial: n = 1282<br><br>• FIGO stage IIIA to IV: 1156/1282 (90.17%)<br>• Postoperative residual tumour size, n (%): unknown: 151; ≤ 1 cm: 773 (68.3%); > 1 cm: 358 (31.7%)<br><br>AGO-OVAR 9 trial: n = 1716<br><br>• FIGO stage IIIA to IV: 1424/1742 (81.75%)<br>• Postoperative residual tumour size, n (%): unknown: 156; ≤ 1 cm: 1.111 (70.1%); > 1 cm: 475 (29.9%)<br><br>Total cohort characteristics:<br><br>Overall mean age of all women was 57.4 years (standard deviation, 10.53)<br><br>FIGO 1A to IIA: 184 (3.6%); FIGO IIB to IIIB: 1182 (23.4%); FIGO IIIC to IV: 3684 (72.9%)<br><br>ECOG 0: 1999 (39.7%); ECOG 1: 2544 (50.5%); ECOG 2: 490 (9.7%); ECOG 3: 2 (0%); ECOG 4: 1 (0%)<br><br>BMI: underweight: 330 (6.5%); normal weight: 2099 (41.5%); overweight: 2626 (51.9%)<br><br>Residual tumour: NMRD: 1779 (36.7%); SVRD (1 mm to 10 mm): 1442 (29.7%); LVRD (> 10 mm): 1629 (33.6%)<br><br>Grading: G1: 399 (8.3%); G2: 1572 (32.9%); G3: 2574 (53.8%); G4: 225 (4.7%); GX: 10 (0.2%)<br><br>Histology: serous: 3656 (72.4%); endometrioid: 428 (8.5%); mucinous: 219 (4.3%); undifferentiated: 214 (4.2%); others: 533 (10.6%)<br><br>Death: tumour related: 2686 (94.8%); therapy associated: 24 (0.8%); other: 124 (4.4%)<br><br>Germany, Austria and France |
| Residual disease details | Speciality of surgeon not reported<br><br>All women underwent surgical cytoreduction followed by chemotherapy regimens:<br><br>AGO-OVAR 3 trial: comparison of the combination of carboplatin/paclitaxel with paclitaxel/cisplatin<br><br>AGO-OVAR 5 trial: comparison of carboplatin/paclitaxel and epirubicin with carboplatin/paclitaxel<br><br>AGO-OVAR 7 trial: comparison of carboplatin/paclitaxel followed by topotecan with carboplatin/paclitaxel<br><br>AGO-OVAR 9 trial: comparison of carboplatin and paclitaxel with or without gemcitabine |
| Outcomes | The effect of young age on PFS and OS in a multivariate analysis including all potential confounders<br><br>FIGO III to IV versus IIB to IIIB:<br><br>• PFS (HR 1.55, 95% CI 1.40 to 1.71)<br>• OS (HR 1.51, 95% CI 1.34 to 1.70)<br><br>Residual tumour NMRD versus SVRD:<br><br>• PFS (HR 0.47, 95% CI 0.43 to 0.52)<br>• OS (HR 0.43, 95% CI 0.38 to 0.49) |

**Klar 2016** *(Continued)*

| | |
|---|---|
| | Residual tumour LVRD (> 10 mm) versus SVRD (1 mm to 10 mm):<br>• PFS (HR 1.22, 95% CI 1.12 to 1.33)<br>• OS (HR 1.21, 95% CI 1.10 to 1.33) |
| Risk of bias (QUIPS) | 1. Study participation (a-f): low risk<br>Adequate number of participants and description of target population. Baseline characteristics, eligibility criteria, sampling frame and period/place of study presented clearly.<br>2. Study attrition (a-e): unclear risk<br>Unclear if patients with incomplete follow-up were excluded before arriving at the stated sample size. Insufficient information to permit judgement.<br>3. Prognostic factor measurement (a-f): unclear risk<br>Adequate cut-off for residual disease used. As data come from different trials, this may introduce heterogeneity in measurement of RD<br>**Outcome level assessment:**<br>**Outcome: overall survival**<br>4. Outcome measurement (a-c): low risk<br>Valid and reliable measurement of outcome<br>5. Adjustment for other prognostic factors (a-g): unclear risk<br>Age and BMI dichotomised. Tumour grading also dichotomised. Multivariate model for OS adjusted for ECOG, BMI, FIGO stage, tumour grading and histology.<br>6. Statistical analysis and reporting (a-d): high risk<br>No conceptual framework; unclear on reasons why the particular specific set of variables were selected for univariate screening. Criteria for variable selection into multivariate models unclear. Dichotomisation of continuous variables apparent.<br>**Outcome: progression-free survival**<br>4. Outcome measurement (a-c): low risk<br>Valid and reliable measurement of outcome<br>5. Adjustment for other prognostic factors (a-g): unclear risk<br>Age and BMI dichotomised. Tumour grading also dichotomised. Multivariate model for PFS adjusted for ECOG, BMI, FIGO stage, tumour grading and histology.<br>6. Statistical analysis and reporting (a-d): high risk<br>No conceptual framework; unclear on reasons why the particular specific set of variables were selected for univariate screening. Criteria for variable selection into multivariate models unclear. Dichotomisation of continuous variables apparent. |
| Notes | Follow-up times:<br>• AGO-OVAR 3 trial: women were followed for nearly 50 months in the trial<br>• AGO-OVAR 5 trial: median follow-up time for surviving women in both groups was 42 months (range 0 to 61 months)<br>• AGO-OVAR 7 trial: median KM follow-up time was 54 months for both groups<br>• AGO-OVAR 9 trial: median follow-up time was 49 months in both groups |

**Langstraat 2011**

| ***Study characteristics*** | |
|---|---|
| Methods | Retrospective review of medical records |
| Participants | Women with stage IIIC to IV primary ovarian cancer and managed with the intention of complete tumour cytoreduction (NMRD) followed by treatment with Taxol and platinum-based chemotherapy<br><br>Women had to be 65 years of age and older<br><br>Exclusion: women who received neoadjuvant chemotherapy, underwent initial surgical debulking at another facility or had borderline tumour histology or non-epithelial cancer. Women who required emergent/urgent surgical intervention due to a small bowel obstruction were included if the stated primary surgical goal was to achieve complete cytoreduction, otherwise they were excluded.<br><br>N = 280<br><br>Mean age 73.5 years (range: 65 to 89); 33% 80 years or older<br><br>The group of women was divided into 4 age groups: 65 to 69, 70 to 74, 75 to 79, over 80 for statistical analysis<br><br>ASA 1 to 2: 96, 3 to 4: 181<br><br>Stage IIIC: 210, Stage IV: 67<br><br>Histological subtype; serous: 205, mucinous: 6, endometrioid: 17, clear cell: 6, other: 43<br><br>40% albumin > 3.0 g/dL<br><br>Mean creatinine = 1.05<br><br>USA |
| Residual disease details | Type of surgeon not reported<br><br>Postoperative residual disease was defined as:<br><br>• NMRD 61 (21.8%)<br>• SVRD (0 cm to 1 cm) 120 (42.8%)<br>• LVRD (> 1 cm) 95 (35.5%)<br><br>The surgical complexity score (SCS) was assigned based on the extent of surgical effort and is calculated based on the number and type of procedures the women underwent. High complexity is defined if the score is over 7, and low complexity if the score is 3 or less. |
| Outcomes | OS<br><br>HR (LVRD (> 1 cm) vs NMRD) 4.51 (95% CI 2.92 to 7.17)<br><br>HR (SVRD vs NMRD) 2.24 (95% CI 1.48 to 3.49)<br><br>HRs adjusted for creatinine, surgical complexity score, FIGO stage and age group |
| Risk of bias (QUIPS) | 1. Study participation (a-f): low risk<br><br>Adequate number of participants and description of target population. Baseline characteristics, eligibility criteria, sampling frame and period/place study took place presented clearly.<br><br>2. Study attrition (a-e): unclear risk |

**Langstraat 2011** *(Continued)*

| | |
|---|---|
| | Unclear if patients with incomplete follow-up were excluded before arriving at the stated sample size. Insufficient information to permit judgement.<br><br>3. Prognostic factor measurement (a-f): low risk<br><br>Valid and reliable measurement of RD<br><br>**Outcome level assessment:**<br><br>**Outcome: overall survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Definition of OS not provided but it usually has a standard definition.<br><br>5. Adjustment for other prognostic factors (a-g): unclear risk<br><br>Ascites was dichotomised with arbitrary cutoff of 1000 mL. Age as defined as a continuous and categorical variable in univariate analysis. CA-125 dichotomised with arbitrary cutoff of 750 U/mL. Creatinine dichotomised arbitrarily. Multivariate model predicting OS adjusted for creatinine, surgical complexity score, FIGO stage and age.<br><br>6. Statistical analysis and reporting (a-d): high risk<br><br>No conceptual framework; unclear of variable selection strategy into multivariate model<br><br>**Outcome: progression-free survival**<br><br>Not reported |
| Notes | Mean follow-up was of 3.2 years (range 0 to 15.8 years)<br><br>30-day mortality was observed in 12 of 280 (4.3%) women<br><br>Older women who underwent surgery had a poorer performance score, higher mean creatinine, lower mean albumin and were more likely to have stage III disease. Only 15% of women who underwent surgery in the oldest age group had stage IV disease, compared to 26% of the rest of the cohort.<br><br>Survival benefit was most apparent with complete cytoreduction but this benefit decreased with increasing age (median survival 5 years versus age group 65 to 69 at 5.9 years.<br><br>Despite the trend towards lower surgical complexity in the older women over age 80 years (45%), there was a significant increase in surgical morbidity, mortality and the inability to receive chemotherapy. Similar trend was seen in women aged > 75 years. |

**Lecointre 2020**

| ***Study characteristics*** | |
|---|---|
| Methods | Retrospective, multicentre cohort study in 9 referral centres of France, constituting the FRANCOGYN study group |
| Participants | 501 women with histologically confirmed advanced epithelial ovarian cancer of stages III or IV according to the FIGO classification, diagnosed between January 2000 and June 2017. Participants were split into those with ≤ 4 NACT cycles and > 4 NACT cycles.<br><br>Median age: ≤ 4 NACT cycles: 60.7 years; > 4 NACT cycles: 62.6 years<br>BMI: < 25: 406 (81%); 25 to 30: 2 (1%); > 30: 93 (18%)<br>White ethnicity: 246/284 (87%)<br>Personal or familiar history of gynaecological cancer: 171 (34%)<br>FIGO III: 409 (82%); FIGO IV: 92 (18%) |

**Lecointre 2020** *(Continued)*

| | |
|---|---|
| | Serous histology: 274/478 (57%)<br>Pre-operative CA-125, U/mL: > 500: 302 (60%); ≤ 500: 199 (40%)<br>Charlson index ≥ 1: 103/298 (35%)<br>Tumour grade 1 to 2: 65 (13%); tumour grade 3: 248 (87%) |
| Residual disease details | The type of surgery performed was classified as complete (R0) when all visible tumours were removed (NMRD (referred to RD0 in study)) at the end of the intervention, R1 when it was ≤ 2.5 mm, R2 when it was more than > 2.5 mm but less than 2.5 cm<br><br>NMRD: 346/471 (73%); RD > 0 cm to 2.5cm: 125/471 (27%)<br><br>30 participants had missing RD data |
| Outcomes | Median OS: 54.2 months<br>5-year survival<br>≤ 4 cycles: 45.6%; > 4 cycles: 27.6%<br>10-year survival<br>≤ 4 cycles: 26 %; > 4 cycles: 11%<br><br>In multivariate Cox model controlling for number of NACT cycles (≤ 4, > 4); age (cat); Charlson index; FIGO; lymph node status (N+ vs N0); response to NACT; residual disease (RD > 0 cm to 2.5 cm vs NMRD) (adjusted HR 2.04 (95% CI 1.53 to 2.72))<br><br>Median PFS: 22.9 months<br>5-year survival<br>≤ 4 cycles: 19.7%; > 4 cycles: 11.7%<br><br>In multivariate Cox model controlling for number of NACT cycles (≤ 4, >4); age (cat); response to NACT; residual disease (RD > 0 cm to 2.5 cm vs NMRD) (adjusted HR 1.36 (95% CI 1.05 to 1.76)) |
| Risk of bias (QUIPS) | 1. Study participation (a-f): low risk<br><br>Adequate number of participants and description of target population. Baseline characteristics, eligibility criteria, sampling frame and period/place study took place presented clearly.<br><br>2. Study attrition (a-e): unclear risk<br><br>Unclear if patients with incomplete follow-up were excluded before arriving at the stated sample size. Insufficient information to permit judgement.<br><br>3. Prognostic factor measurement (a-f): unclear risk<br><br>Valid and reliable measurement of RD. 471 (94%) have RD data. Multicentre design may introduce heterogeneity in measurement of RD.<br><br>**Outcome level assessment:**<br><br>**Outcome: overall survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Valid and reliable measurement of outcome<br><br>5. Adjustment for other prognostic factors (a-g): high risk<br><br>Multivariate Cox model for OS adjusted for number of NACT cycles (≤ 4, > 4); age (cat); Charlson index; FIGO; lymph node status (N+ vs N0); response to NACT; residual disease (RD > 0 cm to 2.5 cm vs RD 0 cm)<br><br>Large missing data rate for Charlson index (40%) and response to NACT (24%) - no methods discussed to handle missing data therefore assumed complete case analysis.<br><br>6. Statistical analysis and reporting (a-d): unclear risk |

**Lecointre 2020** *(Continued)*

No conceptual framework; data driven based on P values of univariate associations. Unclear on reasons why the particular specific set of variables were selected for univariate screening. Although multivariate estimates for RD were presented in the text of results, they did not appear in the corresponding tables.

**Outcome: progression-free survival**

4. Outcome measurement (a-c): low risk

Valid and reliable measurement of outcome

5. Adjustment for other prognostic factors (a-g): high risk

Multivariate Cox model for PFS adjusted for number of NACT cycles (≤ 4, > 4); age (cat); response to NACT; residual disease (RD > 0 cm to 2.5 cm vs RD 0 cm)

Large missing data rate for response to NACT (24%) - no methods discussed to handle missing data therefore assumed complete case analysis.

6. Statistical analysis and reporting (a-d): unclear risk

No conceptual framework; data driven based on P values of univariate associations. Unclear on reasons why the particular specific set of variables were selected for univariate screening. Although multivariate estimates for RD were presented in the text of results, they did not appear in the corresponding tables.

| | |
|---|---|
| Notes | Study reports n = 471 with RD data, but due to missing data from other variables in the multivariate model, the HR estimates for OS and PFS may not be based on complete case analysis and could be based on less, unless imputation was used (e.g. multiple imputation by chained equations).<br><br>Median NACT cycles<br><br>≤4 cycles: median 4 (range 3 to 4); > 4 cycles: median 6 (range 5 to 8)<br><br>NACT regime<br><br>Platinum and taxane: 464 (93%); other platinum-based: 37 (7%)<br><br>Response to NACT:<br><br>Complete response: 73/380 (19%); partial: 307/380 (81%)<br><br>Time from diagnosis to IDS, months<br><br>≤ 4 NACT cycles: 3.8 (range 3.1 to 4.7); > 4 cycles: 5.9 (range 5.1 to 7.7)<br><br>Operating duration, minutes<br><br>≤ 4 cycles: 328 (range 300 to 375); > 4 cycles: 360 (range 293 to 450)<br><br>Blood transfusion:<br><br>Yes: 44/77 (57%); no: 33/77 (43%)<br><br>Intraoperative complications:<br><br>Yes: 57/387 (15%); no: 330/387 (85%) |

**Lecuru 2019**

***Study characteristics***

**Lecuru 2019** *(Continued)*

| | |
|---|---|
| Methods | Secondary analysis of the CHIVA double-blind randomised phase II GINECO study. The CHIVA trial explored the role of nintedanib in combination with NACT vs placebo in combination with NACT. |
| Participants | N = 163 participants treated with NACT with FIGO stage IIIC to IV AOC considered as unresectable after laparoscopic (lap) evaluation<br><br>188 participants were originally enrolled into the trial. The decision to exclude 25 participants was not stated. |
| Residual disease details | Women were treated with 3 to 4 cycles of platinum-taxane NACT + oral nintedanib before interval debulking surgery (IDS). CT (up to 6 cycles in total) and nintedanib were pursued postoperatively.<br><br>No definition of optimal cytoreduction provided. Complete surgical resection response (referred to in study as CC0) included as variable but no explicit definition. |
| Outcomes | Multivariable Cox PH model adjusted for ECOG, ascites, neutrophil/lymphocyte ratio, PCI at baseline, RECIST ORR, CC0 at IDS, PCR and treatment arm (nintedanib vs placebo):<br><br>• Complete surgical response (CC0) was predictive of both PFS and OS in multivariable Cox PH models ($P < 0.01$) |
| Risk of bias (QUIPS) | 1. Study participation (a-f): high risk<br><br>Abstract only therefore insufficient information on study participation<br><br>2. Study attrition (a-e): unclear risk<br><br>Unclear if patients with incomplete follow-up were excluded before arriving at the stated sample size. Insufficient information to permit judgement.<br><br>3. Prognostic factor measurement (a-f): low risk<br><br>Valid and reliable measurement of RD<br><br>**Outcome level assessment:**<br><br>**Outcome: overall survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Definition of OS not provided but it usually has a standard definition.<br><br>5. Adjustment for other prognostic factors (a-g): high risk<br><br>Not explicitly stated but implied that model predicting OS adjusted for ECOG, ascites, neutrophil/lymphocyte ratio, Peritoneal Cancer Index at baseline, response rate at end of NACT according to RECIST (RECIST ORR), pathological complete or near complete response rate and treatment arm<br><br>6. Statistical analysis and reporting (a-d): high risk<br><br>No conceptual framework; variable selection criteria undefined and magnitude of effect not reported, only P value<br><br>**Outcome: progression-free survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Definition of PFS not provided but it usually has a standard definition.<br><br>5. Adjustment for other prognostic factors (a-g): high risk<br><br>Not explicitly stated but implied that model predicting OS adjusted for ECOG, ascites, neutrophil/lymphocyte ratio, Peritoneal Cancer Index at baseline, response rate at end of NACT according to RECIST (RECIST ORR), pathological complete or near complete response rate and treatment arm |

**Lecuru 2019** *(Continued)*

| | |
|---|---|
| | 6. Statistical analysis and reporting (a-d): high risk<br><br>No conceptual framework; variable selection criteria undefined and magnitude of effect not reported, only P value |
| Notes | Abstract only<br><br>Refer to Ferron 2019 for trial results for all n = 188 participants<br><br>From Ferron 2019:<br><br>Women with FIGO stage IIIC to IV chemotherapy-naive AEOC considered as unresectable after laparoscopic evaluation were randomised (2:1) to be treated with 3 to 4 cycles (cy) of carboplatin (AUC 5 mg/mL/min) and paclitaxel (175 mg/m$^2$) (CP) before interval debulking surgery (IDS) followed by 2 to 3 cycles of CP for a total of 6 cycles, plus either 200 mg of nintedanib (arm A) or placebo (arm B) twice daily on days 2 to 21 q3 week at cycles 1 and 2, 5 and 6 and maintenance therapy for up to 2 years. |

**Liu 2020**

| ***Study characteristics*** | |
|---|---|
| Methods | Retrospective analysis of past medical data from First Affiliated Hospital of Third Military Medical University from January 2009 to December 2017<br><br>China |
| Participants | 114 women with stage III to IV epithelial ovarian cancer diagnosed by biopsy or cytologic examination based on histological proofs who received neoadjuvant chemotherapy followed by laparoscopic conservative interval debulking surgery (NACT + LIDS)<br><br>Mean age: 51.6 (SD 9.3)<br>Mean BMI: 23.2 (SD 3.3)<br>FIGO III: 94 (82%); FIGO IV: 10 (18%)<br>Serous histology: 97 (85%)<br>Tumour grade<br>High: 92 (81%); medium: 4 (3%); low 3 (3%); unknown: 15 (13%)<br>Lymph node status<br>Positive: 56 (49%); negative: 58 (51%) |
| Residual disease details | NMRD (referred to in study as R0) disease was defined as all diseases that were cytoreduced by electronic devices. If these diseases were not resected using an en bloc approach, leaving SVRD (≤1 cm), authors considered it as optimal (R1).<br><br>NMRD: 66 (58%)<br><br>SVRD (< 1 cm): 48 (42%) |
| Outcomes | Median OS: 56 months<br><br>Univariate association between RD and OS was reported ≥ SVRD (< 1 cm) vs NMRD: HR 9.589 (95% CI 3.911 to 23.507)<br><br>No variable other than RD was included in the "multivariate" model. Therefore, this was not included in the analysis and this is noted in the interpretation of the results.<br><br>Median DFS: 14 months<br><br>After controlling for age (continuous), residual disease (SVRD vs NMRD): adjusted HR 6.022 (95% CI 3.632 to 9.986) |

**Liu 2020** *(Continued)*

| | |
|---|---|
| Risk of bias (QUIPS) | 1. Study participation (a-f): low risk<br><br>Adequate number of participants and description of target population. Baseline characteristics, eligibility criteria, sampling frame and period/place study took place presented clearly.<br><br>2. Study attrition (a-e): unclear risk<br><br>Unclear if patients with incomplete follow-up were excluded before arriving at the stated sample size. Insufficient information to permit judgement.<br><br>3. Prognostic factor measurement (a-f): low risk<br><br>Valid and reliable measurement of RD<br><br>**Outcome level assessment:**<br><br>**Outcome: overall survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Valid and reliable measurement of outcomes<br><br>5. Adjustment for other prognostic factors (a-g): high risk<br><br>No variable other than RD was included in the "multivariate" model. Therefore, this was not included in the analysis and this is noted in the interpretation of the results.<br><br>6. Statistical analysis and reporting (a-d): high risk<br><br>No conceptual framework; unclear how variables were selected into multivariate model and why the absence of key variables.<br>Selection strategy led to multivariate Cox model for OS with RD as the only predictor.<br><br>**Outcome: disease-free survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Valid and reliable measurement of outcomes<br><br>5. Adjustment for other prognostic factors (a-g): high risk<br><br>Only adjustment for age in multivariate model for DFS<br><br>6. Statistical analysis and reporting (a-d): high risk<br><br>No conceptual framework; unclear how variables were selected into multivariate model and why the absence of key variables. |
| Notes | Patients received IV paclitaxel and carboplatin/ cisplatin or IV docetaxel and cisplatin every 3 weeks<br><br>Number of NACT cycles<br><br>2: 67 (59%); 3: 37 (32%); 10: (9%)<br><br>Number of adjuvant chemotherapy cycles<br><br>3 to 4: 30 (26%); 5: 42 (37%); ≥ 6: 42 (37%) |

**Lorusso 2016**

***Study characteristics***

**Lorusso 2016** *(Continued)*

| | |
|---|---|
| Methods | Multicentre, retrospective review of consecutive women who underwent NACT-IDS in 5 Italian centres |
| Participants | N = 193 participants with advanced-stage ovarian cancer |
| Residual disease details | 3 NACT cycles: 77 (44%)<br>4 NACT cycles: 74 (38%)<br>5 NACT cycles or more: 43 (22%)<br>Text suggests residual disease was treated as NMRD vs any macroscopic RD (> 0 cm) |
| Outcomes | 5-year overall survival (OS) was 46% and 31% for women having 3 and 4+ cycles of NACT<br>10-year OS was 26% and 18% for women having 3 and 4+ cycles of NACT<br>"A trend towards worse OS was observed for women with residual disease at IDS": HR 1.29 (95% CI 0.98 to 1.70), P = 0.06<br>Unknown number of covariates in model except for ECOG performance status. Residual disease variable presumed to be RD > 0 cm vs NMRD. |
| Risk of bias (QUIPS) | 1. Study participation (a-f): high risk<br>Abstract only therefore insufficient information on study participation<br>2. Study attrition (a-e): unclear risk<br>Unclear if patients with incomplete follow-up were excluded before arriving at the stated sample size. Insufficient information to permit judgement.<br>3. Prognostic factor measurement (a-f): low risk<br>Valid and reliable measurement of RD<br>**Outcome level assessment:**<br>**Outcome: overall survival**<br>4. Outcome measurement (a-c): low risk<br>Definition of OS not provided but it usually has a standard definition.<br>5. Adjustment for other prognostic factors (a-g): high risk<br>Unclear on which variables were adjusted for but we know there is at least ECOG and number of NACT cycles<br>6. Statistical analysis and reporting (a-d): high risk<br>No conceptual framework; unclear on reasons why the particular specific set of variables were selected for multivariate model<br>**Outcome: progression-free survival**<br>Not reported |
| Notes | Abstract only |

**Luger 2020**

| ***Study characteristics*** | |
|---|---|
| Methods | Retrospectively review of patients diagnosed between 2000 and 2016<br>Austria |
| Participants | 178 stage III and IV ovarian cancer patients<br>Median age at diagnoses was 64.6 years (interquartile range (IQR) 50.8 to 72.7)<br>Only patients without surgically removed enlarged cardiophrenic lymph nodes (CPLN) were eligible for this study<br>FIGO III: 91 (51%); FIGO IV: 87 (49%)<br>Histology<br>Serous: 157 (88%); mucinous: 3 (2%); endometrioid: 13 (7%); clear cell: 5 (3%)<br>Tumour grade: 1: 17 (10%); 2: 82 (46%); 3: 79 (44%)<br>Median follow-up duration: 49.6 months (IQR 32.9 to 66.3) |
| Residual disease details | All patients received primary upfront primary debulking surgery (PDS) by dedicated teams including at least one certified gynaecologic oncologist, and all received adjuvant platinum-based chemotherapy.<br>The authors defined “No residual disease” as complete macroscopic tumour resection at the end of debulking surgery<br>Residual disease groups:<br>NMRD: 133 (75%)<br>RD > 0 cm: 45 (25%) |
| Outcomes | Overall and progression-free survival |
| Risk of bias (QUIPS) | 1. Study participation (a-f): low risk<br>Adequate number of participants and description of target population. Baseline characteristics, eligibility criteria, sampling frame and period/place study took place presented clearly.<br>2. Study attrition (a-e): unclear risk<br>Unclear if patients with incomplete follow-up were excluded before arriving at the stated sample size. Insufficient information to permit judgement.<br>3. Prognostic factor measurement (a-f): low risk<br>Valid and reliable measurement of RD<br>**Outcome level assessment:**<br>**Outcome: overall survival**<br>4. Outcome measurement (a-c): low risk<br>Definition of OS not provided but it usually has a standard definition<br>5. Adjustment for other prognostic factors (a-g): low risk<br>HR for OS was adjusted for age (> 64.6 years), CA-125, paraaortic nodes (positive), stage, residual disease, and CPLN dimension in multivariable Cox model<br>6. Statistical analysis and reporting (a-d): high risk |

**Luger 2020** *(Continued)*

| | |
|---|---|
| | No conceptual framework; unclear of variable selection criteria in multivariate model<br><br>**Outcome: progression-free survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Definition of PFS not provided but it usually has a standard definition<br><br>5. Adjustment for other prognostic factors (a-g): low risk<br><br>HR for PFS was adjusted for age (> 64.6 years), CA-125, paraaortic nodes (positive), stage, residual disease, and CPLN dimension in multivariable Cox model<br><br>6. Statistical analysis and reporting (a-d): high risk<br><br>No conceptual framework; unclear of variable selection criteria in multivariate model |
| Notes | Residual disease in multivariate model for: PFS: HR 2.44 (95% CI 1.23 to 4.84), P = 0.011; OS: HR 2.17 (95% CI 1.11 to 4.69), P = 0.028. The upper 95% CI for OS was entered into forest plots as 4.26 so slight margin of error in the reported statistic). Multivariate model was adjusted for age, CA-125, histologically positive paraaortic lymph nodes, FIGO stage (IIIA to IIIC vs FIGO IVA and IVB), cardiophrenic lymph node (CPLN) and residual disease.<br><br>Recurrence was observed in 66.9% (n = 119) of patients and the median progression-free survival was 12.0 months (IQR 5.5 to 30.5). 80 patients (44.9%) died during a median time of follow-up of 49.6 months (IQR 32.89 to 66.26).<br><br>Adjuvant chemotherapy:<br>Carboplatin + paclitaxel: 150 (84%); carboplatin: 24 (14%); carboplatin + endoxan: 4 (2%)<br><br>Platinum response:<br>Refractory + resistant: 35 (20%); sensitive: 143 (80%)<br><br>A systematic pelvic and paraaortic lymphadenectomy (removal of ≥ 20 retroperitoneal lymph nodes was performed in 84.2% of patients<br>Systematic retroperitoneal lymphadenectomy (removal of ≥ 20 nodes): 150 (84.2%)<br>Sampling retroperitoneal lymphadenectomy (removal of < 20 nodes): 8 (4%)<br>Median number of removed nodes: 26 (IQR 7 to 37)<br>88 (68%) had exhibited histologically proven retroperitoneal lymph node metastases<br><br>Intraperitoneal carcinomatosis radiologically evident in 151 (85%)<br>Radiological diagnosis of upper abdominal spread in 72 (41%) |

**McGuire 1995**

| ***Study characteristics*** | |
|---|---|
| Methods | Retrospective analysis of a prospective randomised controlled trial comparing different chemotherapy dosing schedules. It aimed to determine the importance of chemotherapy dose intensity on survival, progression-free survival (PFS) and response. This was not a trial of surgery but the report allows a comparison of survival outcomes for subgroups women with stage III ovarian cancer who have had < 2 cm or ≥ 2 cm of residual disease following surgery and therefore is relevant to this review. |
| Participants | 458 women with FIGO stage III and IV epithelial ovarian cancer were recruited. These were women who had more than 1 cm residual disease following initial surgery.<br><br>27 women were ineligible: incorrect stage (n = 5), incorrect primary tumour (n = 9), incorrect cell type (n = 7), history of prior malignancy (n = 3), prior chemotherapy (n = 1) and other (n = 2)<br><br>Women with borderline ovarian tumours (low malignant potential) were excluded |

**McGuire 1995** *(Continued)*

| | |
|---|---|
| | Recruitment was from December 1986 to April 1990 and all women had undergone a surgical procedure<br><br>The median age at study entry was 60 years (range: 20 to 83)<br><br>305 (67%) and 153 (33%) women had FIGO stage III and IV disease, respectively<br><br>Tumour cell type: serous 312 (68.1%), endometrioid: 64 (14%), mucinous; 12 (2.6%), clear cell: 12 (2.6%), other: 58 (12.7%)<br><br>Tumour grade: 1: 26 (9%), 2: 114 (39%), 3: 152 (52%), not specified 2 (1%)<br><br>GOG score: 0: 150 (32.8), 2: 213 (46.5%), 3: 95 (20.7%) |
| Residual disease details | Residual disease was noted as follows:<br><br>1. LVRD between 1 cm and 2 cm for women with stage III disease: 31 (6.8%)<br>2. LVRD greater than 2 cm for women with stage III disease: 274 (58.9%)<br>3. LVRD between 1 cm and 2 cm for women with stage IV disease: 54 (11.8%)<br>4. LVRD greater than 2 cm for women with stage IV disease: 99 (21.6%)<br><br>**Definition of optimal surgery:**<br><br>All women were 'suboptimally' cytoreduced with > 1 cm of residual disease<br><br>**Chemotherapy:**<br><br>2 trial arms with women receiving either standard chemotherapy: cyclophosphamide 500 mg/m$^2$ and cisplatin 50 mg/m$^2$ intravenously every 3 weeks for 8 courses OR intense chemotherapy: cyclophosphamide 1000 mg/m$^2$ and cisplatin 100 mg/m$^2$ intravenously every 3 weeks for 4 courses. Dose modification was rigidly controlled to maintain intensity. |
| Outcomes | Overall survival and progression-free survival: HR adjusted for age, GOG performance status, histological sub-type, stage/residual disease and measurable disease using Cox model:<br><br>III, ≥ 2 cm vs III, 1 to 2 cm: HR 1.91<br><br>IV, 1 cm to 2 cm vs III, 1 to 2 cm: HR 1.89<br><br>IV, ≥ 2 cm vs III, 1 to 2 cm: HR 2.29<br><br>Overall and progression-free survival (PFS) were measured from the date of randomisation. All eligible women were included in the analysis of outcomes. All causes of death were used to calculate survival, and the estimates were based on Kaplan-Meier procedures. |
| Risk of bias (QUIPS) | 1. Study participation (a-f): low risk<br><br>Adequate number of participants and description of target population. Baseline characteristics, eligibility criteria, sampling frame and period/place study took place presented clearly.<br><br>2. Study attrition (a-e): unclear risk<br><br>Unclear if patients with incomplete follow-up were excluded before arriving at the stated sample size. Insufficient information to permit judgement.<br><br>3. Prognostic factor measurement (a-f): low risk<br><br>Valid and reliable measurement of RD<br><br>**Outcome level assessment:**<br><br>**Outcome: overall survival**<br><br>4. Outcome measurement (a-c): low risk |

**McGuire 1995** *(Continued)*

| | |
|---|---|
| | Valid and reliable measurement of outcome. OS was measured from the date of randomisation.<br><br>5. Adjustment for other prognostic factors (a-g): unclear risk<br><br>Multivariate model for OS adjusted for age, GOG performance status, histological subtype and measurable disease<br><br>6. Statistical analysis and reporting (a-d): high risk<br><br>No conceptual framework; unclear of variable selection criteria in multivariate model. Magnitude of effect not reported with confidence interval and only P value was available.<br><br>**Outcome: progression-free survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Valid and reliable measurement of outcome. PFS was measured from the date of randomisation.<br><br>5. Adjustment for other prognostic factors (a-g): unclear risk<br><br>Multivariate model for PFS adjusted for age, GOG performance status, histological subtype and measurable disease<br><br>6. Statistical analysis and reporting (a-d): high risk<br><br>No conceptual framework; unclear of variable selection criteria in multivariate model. Magnitude of effect not reported with confidence interval and only P value was available. |
| Notes | Mean and median length of follow-up were not reported. Since this trial was a trial of chemotherapeutic regimens, the randomisation did not aim to compare the effect of different degrees of surgical debulking. The findings borne out on multivariate analysis are similar to those in retrospective and cohort studies. The prospective nature of this study has, however, facilitated the collection of a fairly complete data set and gives this work some authority.<br><br>Other variables in Cox model:<br><br>Age (years): reference group: women aged less than 55 years (P = 0.47): 55 to 65: HR 1.08; > 65: HR 1.38<br><br>GOG performance status: reference group: GOG 0 (P = 0.009) 1: HR 1.26, 2: HR 1.56<br><br>Histological subtype: reference group: serous adenocarcinoma (P < 0.001):<br><br>Endometrioid: HR 0.951, mucinous: HR 8.31, clear cell: HR 1.79, other: HR 0.84<br><br>Measurable disease: reference group:<br>No: (P = 0.01)<br><br>Yes: HR 1.43<br><br>From the study both advancing age and worsening performance status were associated with poorer survival. In addition, mucinous histology is associated with an 8.3 times greater death rate than serous histology (P < 0.001).<br><br>The study shows residual disease after surgery impacts on survival. Even in 'suboptimal' cytoreduction (residual disease greater than 1 cm), women with stage III disease and residual disease diameter less than 2 cm exhibited lower death rates than either those with stage III diease and residual disease diameter of ≥ 2 cm, or those with stage IV disease. |

**Melamed 2017a**

***Study characteristics***

**Melamed 2017a** *(Continued)*

| | |
|---|---|
| Methods | Retrospective cohort study |
| Participants | 307 women with stage IIIC to IV epithelial clear cell carcinoma were included in the analysis<br><br>Age group:<br><br>• < 40: 10 (3.3%)<br>• 40 to 49: 59 (19.2%)<br>• 50 to 59: 131 (42.7%)<br>• 60 to 69: 82 (26.7%)<br>• 70 to 79: 23 (7.5%)<br>• 80+: 2 (0.7%)<br><br>Median age was 56 years<br><br>Race/ethnicity:<br><br>• Asian: 25 (8.1%)<br>• Black: 18 (5.9%)<br>• Hispanic: 24 (7.8%)<br>• White: 240 (78.2%)<br><br>Stage:<br><br>• IIIC: 241 (78.5%)<br>• IV: 66 (21.5%)<br><br>USA |
| Residual disease details | Speciality of surgeon not reported<br><br>All women underwent primary cytoreductive surgery and adjuvant chemotherapy<br><br>Residual disease status was classified as follows:<br><br>• NMRD: 141 (45.9%)<br>• SVRD (1 cm or less): 77 (25.1%)<br>• LVRD measuring > 1 cm: 23 (7.5%)<br>• Unknown: 66 (21.5%) |
| Outcomes | The primary outcome for OS was time from diagnosis to death from any cause, or to last contact, as recorded by the cancer registrar<br><br>NMRD: (AHR 0.34, 95% CI 0.18 to 0.64)<br><br>SVRD (≤ 1 cm): (AHR 0.94, 95% CI 0.50 to 1.75)<br><br>LVRD (> 1 cm): (AHR referent) |
| Risk of bias (QUIPS) | 1. Study participation (a-f): low risk<br><br>Adequate number of participants and description of target population. Baseline characteristics, eligibility criteria, sampling frame and period/place study took place presented clearly.<br><br>2. Study attrition (a-e): unclear risk<br><br>Unclear if patients with incomplete follow-up were excluded before arriving at the stated sample size. Insufficient information to permit judgement.<br><br>3. Prognostic factor measurement (a-f): low risk |

**Melamed 2017a** *(Continued)*

| | |
|---|---|
| | Valid and reliable measurement of RD<br><br>**Outcome level assessment:**<br><br>**Outcome: overall survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Valid and reliable measurement of outcome. OS was time from diagnosis to death from any cause, or to last contact, as recorded by the cancer registrar.<br><br>5. Adjustment for other prognostic factors (a-g): high risk<br><br>Age arbitrarily categorised. Multivariate model predicting OS adjusted for age, race/ethnicity, stage, region, insurance status, treating facility type, hospital annual ovarian cancer volume and presence of comorbidities<br><br>6. Statistical analysis and reporting (a-d): unclear risk<br><br>Authors reported that covariates were selected a priori but difficult to verify<br><br>**Outcome: progression-free survival**<br><br>Not reported |
| Notes | Analysis is a subgroup of women who were analysed from a study that identified 6013 women with stage IIIC and IV high-grade serous, 307 with clear cell and 140 with mucinous histology<br><br>The median follow-up was 34.1 months |

**Melamed 2017b**

| ***Study characteristics*** | |
|---|---|
| Methods | Retrospective cohort study |
| Participants | 6013 women with stage IIIC to IV epithelial high-grade serous ovarian cancer were included in the analysis<br><br>Age group, n (%):<br><br>• < 40: 117 (1.8%)<br>• 40 to 49: 859 (13.3%)<br>• 50 to 59: 1827 (28.3%)<br>• 60 to 69: 2047 (31.7%)<br>• 70 to 79: 1297 (20.1%)<br>• 80+: 314.8%)<br><br>Median age was 63 years<br><br>Race/ethnicity, n (%):<br><br>• Asian: 236 (3.7%)<br>• Black: 467 (7.2%)<br>• Hispanic: 377 (5.8%)<br>• White: 5318 (82.3%)<br>• Other/unknown: 62 (1.0%)<br><br>Stage, n (%): |

**Melamed 2017b** *(Continued)*

| | |
|---|---|
| | • IIIC: 4954 (76.7%)<br>• IV: 1506 (23.3%)<br><br>USA |
| Residual disease details | Speciality of surgeon not reported<br><br>All women underwent primary cytoreductive surgery and adjuvant chemotherapy<br><br>Residual disease status was classified as follows:<br><br>• NMRD: 2048 (34.1%)<br>• SVRD measuring 1 cm or less: 1848 (30.7%)<br>• LVRD measuring > 1 cm: 546 (9.1%)<br>• Unknown: 1571 (26.1%) |
| Outcomes | The primary outcome for OS was time from diagnosis to death from any cause, or to last contact, as recorded by the cancer registrar<br><br>NMRD: (AHR 0.58, 95% CI 0.49 to 0.69)<br><br>SVRD (≤ 1 cm): (AHR 0.85, 95% CI 0.72 to 1.01)<br><br>LVRD (> 1 cm): (AHR referent) |
| Risk of bias (QUIPS) | 1. Study participation (a-f): low risk<br><br>Adequate number of participants and description of target population. Baseline characteristics, eligibility criteria, sampling frame and period/place study took place presented clearly.<br><br>2. Study attrition (a-e): unclear risk<br><br>Unclear if patients with incomplete follow-up were excluded before arriving at the stated sample size. Insufficient information to permit judgement.<br><br>3. Prognostic factor measurement (a-f): low risk<br><br>Valid and reliable measurement of RD<br><br>**Outcome level assessment:**<br><br>**Outcome: overall survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Valid and reliable measurement of outcome. OS was time from diagnosis to death from any cause, or to last contact, as recorded by the cancer registrar.<br><br>5. Adjustment for other prognostic factors (a-g): high risk<br><br>Age arbitrarily categorised. Multivariate model predicting OS adjusted for age, race/ethnicity, stage, region, insurance status, treating facility type, hospital annual ovarian cancer volume and presence of comorbidities<br><br>6. Statistical analysis and reporting (a-d): unclear risk<br><br>Authors reported that covariates were selected a priori but difficult to verify<br><br>**Outcome: progression-free survival**<br><br>Not reported |
| Notes | Analysis is a subgroup of women who were analysed from a study that identified 6013 women with stage IIIC and IV high-grade serous, 307 with clear cell and 140 with mucinous histology |

**Melamed 2017b** *(Continued)*

| | |
|---|---|
| | The median follow-up was 34.1 months |

**Paik 2018**

| ***Study characteristics*** | |
|---|---|
| Methods | Retrospective analysis of data obtained from electronic medical records |
| Participants | 419 EOC women of stages IIIB, IIIC or IV with high-grade serous type histology were investigated<br><br>48 (11.5%) with a normal-sized ovary (less than 4 cm in the longest diameter, with a tumour size greater than 5 × 5 mm within the ovarian substance)<br><br>Mean age of women was 54.5 ± 10.3 years<br><br>Women with enlarged-ovarian tumour were younger (54.0 ± 10.3 vs 58.4 ± 9.2 years) than those in the normal-sized ovary group<br><br>The mean size of ovary was 7.5 ± 3.9 cm for the whole group:<br><br>• With enlarged-ovarian tumour (n = 371): 8.1 ± 3.8 cm<br>• With normal-sized ovary (n = 48); 3.2 ± 1.1 cm<br><br>FIGO stage IIIB: 15 (3.6%); stage IIIC: 335 (84.7%); stage IV: 49 (11.7%)<br><br>Initial CA-125 (U/mL): 1922.4 ± 2968.9<br><br>ASA physical status:<br><br>• I: 191 (45.6<br>• II: 178 (42.5)<br>• III: 18 (4.3)<br>• Unknown: 32 (7.6)<br><br>Korea |
| Residual disease details | Speciality of surgeon not reported<br><br>Women were treated with primary debulking surgery (PDS) with adjuvant chemotherapy for primary treatment<br><br>Residual disease status after PDS (cm) was classified as follows, n(%):<br><br>• NMRD: 107 (25.5%)<br>• SVRD (< 1 cm): 147 (35.1%)<br>• LVRD (≥ 1 cm): 165 (39.4%)<br><br>For adjuvant chemotherapy, the first cycle of combination chemotherapy consisting of taxane/platinum was initiated routinely within 2 weeks of surgery<br><br>Subsequent chemotherapy cycles were performed every 3 weeks for 6 cycles, but there could have been variation in the number of cycles depending on women situation<br><br>Overall survival (OS) was defined as the time between initial diagnosis and women death or loss to follow-up<br><br>Progression-free survival (PFS) was designated as the time between diagnosis and women recurrence/progression or loss to follow-up |

**Paik 2018** *(Continued)*

| | |
|---|---|
| Outcomes | Multivariate Cox proportional hazards analysis of PFS and OS to adjust for risk-associated prognostic clinical features<br><br>Residual disease status after PDS (cm):<br><br>• NMRD: PFS and OS (HR 1)<br>• SVRD (< 1 cm): PFS (HR 1.591, 95% CI 1.153 to 2.193); OS (HR 2.291, 95% CI 1.398 to 3.752)<br>• LVRD (≥ 1 cm): PFS (HR 1.698, 95% CI 1.239 to 2.326); OS (HR 2.549, 95% CI 1.564 to 4.152) |
| Risk of bias (QUIPS) | 1. Study participation (a-f): low risk<br><br>Adequate number of participants and description of target population. Baseline characteristics, eligibility criteria, sampling frame and period/place study took place presented clearly.<br><br>2. Study attrition (a-e): unclear risk<br><br>Unclear if patients with incomplete follow-up were excluded before arriving at the stated sample size. Insufficient information to permit judgement.<br><br>3. Prognostic factor measurement (a-f): low risk<br><br>Valid and reliable measurement of RD<br><br>**Outcome level assessment:**<br><br>**Outcome: overall survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Valid and reliable measurement of outcome. OS was defined as the time between initial diagnosis and women death or loss to follow-up.<br><br>5. Adjustment for other prognostic factors (a-g): unclear risk<br><br>CA-125 arbitrarily dichotomised at cutoff of 35 mL. Multivariate model for OS adjusted for age, CA-125, FIGO stage and normal sized ovary<br><br>6. Statistical analysis and reporting (a-d): high risk<br><br>No conceptual framework; unclear variable selection criteria into multivariate model<br><br>**Outcome: progression-free survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Valid and reliable measurement of outcome. PFS was designated as the time between diagnosis and women recurrence/progression or loss to follow-up.<br><br>5. Adjustment for other prognostic factors (a-g): unclear risk<br><br>CA-125 arbitrarily dichotomised at cutoff of 35 mL. Multivariate model for PFS adjusted for age, CA-125, FIGO stage and normal sized ovary.<br><br>6. Statistical analysis and reporting (a-d): high risk<br><br>No conceptual framework; unclear variable selection criteria into multivariate model |
| Notes | In total cohort with a median follow-up period of 43 months (range, 3 to 164 months),<br><br>Inferior overall survival (OS) was shown in the normal-sized ovary group (median OS, 71.2 vs 41.4 months<br><br>At the time of analysis, of the 419 enrolled women, 298 (71.1%) experienced a relapse, and 192 (45.8%) died after a median observation time of 43 months (range, 3 to 164 months) |

**Paik 2018** *(Continued)*

Other variables in cox model:

Age (continuous): PFS (HR 0.966, 95% CI 0.985 to 1.007); OS (HR 1.003, 95% CI 0.989 to 1.017)

CA-125 level (U/mL):

- < 35: PFS and OS (HR 1)
- ≥ 35: PFS (HR 2.167, 95% CI 1.020 to 4.601); OS (HR 4.437, 95% CI 1.077 to 17.549)

FIGO stage:

- IIIB: PFS and OS (HR 1)
- IIIC: PFS (HR 1.130, 95% CI 0.529 to 2.414); OS (HR 0.638, 95% CI 0.280 to 1.453)
- IV: PFS (HR 1.178, 95% CI 0.520 to 2.671); OS (HR 0.621, 95 % CI 0.249 to 1.550)

Normal-sized ovary:

- No: PFS and OS (HR 1)
- Yes: PFS (HR 1.180, 95% CI 0.839 to 1.660); OS (HR 1.593, 95% CI 1.097 to 2.314)

For primary surgical treatment, bilateral salpingo-oophorectomy, hysterectomy, peritoneal washing, retroperitoneal lymphadenectomy, omentectomy and tumourectomy of any metastatic lesions were performed routinely

**Peiretti 2010**

| ***Study characteristics*** | |
|---|---|
| Methods | Retrospective study |
| Participants | 259 with advanced epithelial ovarian and fallopian tube cancer met the inclusion criteria<br>Median age was 58 years (range: 22 to 77 years)<br>Primary site disease: ovary 256 (98%); fallopian tube 3 (2%)<br>FIGO stages: IIIC: 199 (76%); IV: 60 (24%)<br>Tumour grades: grade 1 to 2: 53 (21%); grade 3: 198 (76%); grade N/A: 8 (3%)<br>Histological type:<br>• Serous: 184 (71%)<br>• Endometrioid: 39 (15%)<br>• Clear cell: 8 (3%)<br>• Mixed: 26 (10%)<br>• Others: 2 (1%)<br>Peritoneal carcinomatosis: yes: 188 (72%); no: 71 (28%)<br>Location of largest mass:<br>• Pelvis: 130 (50%)<br>• Omentum: 110 (42%)<br>• Upper abdomen: 14 (5%)<br>• Retroperitoneal node: 1 (0.4%)<br>• Other: 4 (1.6%)<br>Intraoperative units blood transfused, n (%): |

**Peiretti 2010** *(Continued)*

| | |
|---|---|
| | • None: 147 (56%)<br>• 1 to 2: 67 (26%)<br>• 3 to 4: 31 (12%)<br>• > 5: 14 (5%)<br><br>Postoperative units blood transfused, n (%):<br><br>• None: 122 (47%)<br>• 1 to 2: 113 (43%)<br>• 3 to 4: 23 (8%)<br>• > 5: 4 (2%)<br><br>Size of largest mass (cm): ≤ 10: 98 (38%); > 10: 161 (62%)<br><br>Median CA-125 (range): 913 U/mL (17 to 52,817)<br><br>Median ascites (range): 1500 cc (100 to 15,000)<br><br>Spain and Italy |
| Residual disease details | All these women underwent an attempt of maximal surgical cytoreduction unless there was unresectable disease as determined by the attending surgeon. Speciality of surgeon not reported.<br><br>Postoperative platinum-based chemotherapy was administered in all women<br><br>Residual tumour classed as:<br><br>• NMRD: 115 (44%)<br>• 1 mm to 5 mm: 50 (19%)<br>• 6 mm to 10 mm: 33 13%)<br>• 11 mm to 20 mm: 18 (7%)<br>• > 20 mm: 43 (17%)<br><br>Progression-free survival (PFS) was defined as the time interval from date of surgery to the date of the documented first recurrence of disease |
| Outcomes | At multivariate analysis, age greater than 60 years (P = 0.025), stage IV vs IIIC (P = 0.037) and any residual disease (P = 0.032) were shown to have an independent association with worse PFS<br><br>Median estimated blood loss (range): 700 cc (50 to 6000)<br><br>The median length of hospital stay was 9 days<br><br>Median length of surgery (range): 270 minutes (70 to 480) |
| Risk of bias (QUIPS) | 1. Study participation (a-f): low risk<br><br>Adequate number of participants and description of target population. Baseline characteristics, eligibility criteria, sampling frame and period/place study took place presented clearly.<br><br>2. Study attrition (a-e): unclear risk<br><br>Unclear if patients with incomplete follow-up were excluded before arriving at the stated sample size. Insufficient information to permit judgement.<br><br>3. Prognostic factor measurement (a-f): low risk<br><br>Valid and reliable measurement of RD<br><br>**Outcome level assessment:**<br><br>**Outcome: overall survival** |

**Peiretti 2010** *(Continued)*

| | |
|---|---|
| | 4. Outcome measurement (a-c): low risk<br><br>Valid and reliable measurement of outcome. OS was defined as the time interval from date of surgery to the date of death or last follow-up<br><br>5. Adjustment for other prognostic factors (a-g): high risk<br><br>Not reported in multivariate analyses. Only univariate results.<br><br>6. Statistical analysis and reporting (a-d): high risk<br><br>Not reported in multivariate analyses<br><br>**Outcome: progression-free survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Valid and reliable measurement of outcome. PFS was defined as the time interval from date of surgery to the date of the documented first recurrence of disease.<br><br>5. Adjustment for other prognostic factors (a-g): high risk<br><br>Age categorised. Multivariate model predicting PFS adjusted for age and FIGO stage. Unclear if ascites was included in multivariate model or not.<br><br>6. Statistical analysis and reporting (a-d): high risk<br><br>No conceptual framework; variable selection criteria for multivariate analyses unstated. Multivariate results (hazard ratios) PFS not displayed, only P values. |
| Notes | After a median follow-up of 29.8 months, PFS and overall median survival (OS) were 19.9 and 57.6 months respectively<br><br>92% of the women completed 5 or more cycles of platinum-based systematic chemotherapy<br>At univariate analysis, factors significantly associated with decreased PFS included: age greater than median (N60 years), stage IV, presence of ascites N1000 cc, presence of diffuse peritoneal carcinomatosis and macroscopic residual disease |

**Peiretti 2012**

| | |
|---|---|
| ***Study characteristics*** | |
| Methods | Retrospective medical chart review |
| Participants | 238 consecutive women who underwent rectosigmoid colectomy as part of cytoreductive surgery for ovarian cancer during the study interval were included<br><br>Median age was 59.7 years (range: 22 to 85 years)<br><br>FIGO stage IIC: 3 (1%); IIIA: 1(0.4%); IIIB: 2 (0.8%); IIIC: 174 (73%); IV: 58 (24%)<br><br>Primary site disease:<br><br>• Ovary: 230 (96%)<br>• Fallopian tube: 4 (2%)<br>• Peritoneal cancer: 4(2%)<br><br>Tumour grade:<br><br>• 1 to 2: 51 (22%)<br>• 3: 184 (77%) |

**Peiretti 2012** *(Continued)*

| | |
|---|---|
| | • N/A: 3 (1%)<br><br>Histological subtype:<br><br>• Serous: 200 (84%)<br>• Endometrioid: 15 (6%)<br>• Clear cell: 5 (3%)<br>• Mixed: 18 (7%)<br><br>Median ascites (range): 1500 $cm^3$ 100 to 11,000)<br><br>Italy (157) and USA (81) |
| Residual disease details | All operations were performed by gynaecologic oncologists<br><br>Postoperative platinum-based chemotherapy was administered in all women<br><br>• 62% underwent carbo-platinum and Taxol regimen<br>• Doxorubicin liposomal, gemcitabine and topotecan were the other chemotherapeutic drugs used in association with platinum<br><br>Complete cytoreduction was defined as no visible residual tumour at the completion of the primary operation.<br><br>Reported categories for residual disease (mm) where as follows - no. of women (%):<br><br>• NMRD: 99 (41%)<br>• SVRD (1 mm to 10 mm): 106 (44%)<br>• LVRD (> 10 mm): 32 (15%) |
| Outcomes | The risk factor significantly associated with decreased overall survival (OS) was the presence of any macroscopic residual disease at the end of surgery (P = 0.003)<br><br>The median overall survival time from the time of surgery for all women was 55 months<br><br>A statistically significant difference (P = 0.002) was observed in OS between the group with no macroscopic residual disease (median of 72 months) and the other women with any other gross residual disease (median of 42 months)<br><br>Median estimated blood loss (range): 1000 $cm^3$ (200 to 8500)<br><br>Intraoperative blood transfusion: 152 (64%)<br><br>Postoperative blood transfusion: 150 (63%)<br><br>Median length of hospitalisation (days): 10 (range: 4 to 24 days) |
| Risk of bias (QUIPS) | 1. Study participation (a-f): low risk<br><br>Adequate target population. Baseline characteristics, eligibility criteria, sampling frame and period/place study took place presented clearly. Sample consists of small subset (n = 3, 1%) of stage IIC participants.<br><br>2. Study attrition (a-e): unclear risk<br><br>Unclear if patients with incomplete follow-up were excluded before arriving at the stated sample size. Insufficient information to permit judgement.<br><br>3. Prognostic factor measurement (a-f): unclear risk<br><br>Valid and reliable measurement of RD. Multicentre design may introduce heterogeneity in measurement of RD |

**Peiretti 2012** *(Continued)*

| | |
|---|---|
| | **Outcome level assessment:**<br>**Outcome: overall survival**<br>4. Outcome measurement (a-c): low risk<br>Definition of OS not provided but it usually has a standard definition<br>5. Adjustment for other prognostic factors (a-g): low risk<br>Multivariate model predicting OS adjusted for age, stage, histology, grade and presence of ascites<br>6. Statistical analysis and reporting (a-d): high risk<br>No conceptual framework; variable selection criteria for multivariate analysis unstated<br>**Outcome: progression-free survival**<br>Not reported |
| Notes | Mean or median length of follow-up were not reported<br>Among all groups of women 85% were able to complete at least 5 cycles of (platinum-based) systematic chemotherapy<br>50% of women recurred during the study period. Among them, 74% had a recurrence in the upper abdomen. 8% of the women presented with abdominal recurrence associated to pelvic disease.<br>Only 5% of the women showed a relapse in the pelvis<br>14% of the women presented with distant metastases at the time of recurrence<br>Both univariate and multivariate analyses including the following variables were performed: age, stage, histology, grade, presence of ascites and residual tumour at end of surgery, however no HR are presented in the study |

**Petrillo 2014**

| ***Study characteristics*** | |
|---|---|
| Methods | Single-centre retrospective of medical data (January 1995 to December 2010) retrieved from the electronic database of the Gynecologic Oncology Unit of the Catholic University of Rome and Campobasso |
| Participants | N = 322 women were admitted to the Gynecologic Oncology Unit of the Catholic University of Rome and Campobasso, with a diagnosis of advanced ovarian, tubal or peritoneal cancer. All these women were judged as having unresectable advanced disease after initial surgical exploration and submitted to NACT followed by IDS.<br>≤ 65 years: 226 (70.2%)<br>> 65 years: 96 (29.8%)<br>FIGO: IIIC – 251 (77.7%); IV – 72 (22.3%)<br>Histology: serous – 264 (82%); other – 58 (18%)<br>Tumour grade: G1 – 9 (2.7%); grade 2/3 – 313 (97.3%)<br>Ascites: 247 (76.7%)<br>Median CA-125 at diagnosis: 548 (range: 9 to 9999) |

**Petrillo 2014** *(Continued)*

| | |
|---|---|
| | Carcinomatosis at diagnosis: 285 (88.5%)<br><br>Within FIGO IV (n = 72)<br><br>Presence of pleural effusion: 37<br><br>Metastasis in liver, spleen or lung: 34 |
| Residual disease details | 3 to 4 NACT cycles: 216 (82.3%)<br><br>6 NACT cycles: 57 (17.7%)<br><br>NACT regimen: carboplatin alone – 51 (15.8%); carboplatin/paclitaxel or pegylated-liposomal doxorubicin (PLD) – 271 (84.2%)<br><br>Pathological response to NACT:<br><br>• Complete (cPR in cases with no macroscopic residual neoplastic cells in all the surgical specimens, including the adnexa): 21 (6.5%)<br>• Microscopic response (without macroscopic lesions but with microscopic foci (maximum diameter ≤3 mm)): 104 (32.3%)<br>• Macroscopic response (persistent macroscopic site of disease after NACT were classified as a macroscopic response): 197 (61.2%)<br><br>Study did not provide a definition of optimal cytoreduction<br><br>• NMRD: 236 (73.3%)<br>• SVRD (0 cm to 1 cm): 36 (11.2%)<br>• LVRD (RD > 1 cm): 50 (15.5%) |
| Outcomes | Overall survival defined as time elapsed between diagnosis and death or date of last follow-up (second half of 2012 in all women)<br><br>Death from disease: 239 (74.2%)<br><br>Median OS in those who had complete response (NMRD) from NACT: 72 months<br><br>Median OS in those who had optimal response: 38 months<br><br>Median OS in those who had suboptimal response: 29 months<br><br>Multivariable Cox PH for OS adjusted for pathological response to NACT:<br><br>• Residual tumour at IDS (RT = 0 vs RT ≤ 1 vs explorative laparotomy): X2 = 24.951, P = 0.001<br><br>Progression-free survival (PFS) calculated from the date of diagnosis to the date of first relapse or the date of the last follow-up (second half of 2012 in all women)<br><br>Recurrences: 285 (88.2%)<br><br>Median PFS in those who had complete response (NMRD) from NACT: 36 months<br><br>Median PFS in those who had optimal response: 16 months<br><br>Median PFS in those who had suboptimal response: 13 months<br><br>Multivariable Cox PH for PFS adjusted for age, carcinomatosis at diagnosis, CA-125, pathological response to NACT:<br><br>• Residual tumour at IDS (RT = 0 vs RT ≤ 1 vs explorative laparotomy): X2 = 39.716, P = 0.001<br><br>* No adjusted HR estimates provided for OS or PFS |
| Risk of bias (QUIPS) | 1. Study participation (a-f): low risk |

**Petrillo 2014** *(Continued)*

| | |
|---|---|
| | Adequate number of participants and description of target population. Baseline characteristics, eligibility criteria, sampling frame and period/place study took place presented clearly.<br><br>2. Study attrition (a-e): unclear risk<br><br>Unclear if patients with incomplete follow-up were excluded before arriving at the stated sample size. Insufficient information to permit judgement.<br><br>3. Prognostic factor measurement (a-f): low risk<br><br>Valid and reliable measurement of RD<br><br>**Outcome level assessment:**<br><br>**Outcome: overall survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Valid and reliable measurement of outcome; OS defined as time elapsed between diagnosis and death or date of last follow-up<br><br>5. Adjustment for other prognostic factors (a-g): high risk<br><br>Unstated why explorative laparotomy is a category within the RD variable. Model predicting OS only adjusted for pathological response to NACT.<br><br>6. Statistical analysis and reporting (a-d): high risk<br><br>No conceptual framework; data driven based on P values of univariate associations. Results for multivariate analysis of OS not reported using hazard ratios.<br><br>**Outcome: progression-free survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Valid and reliable measurement of outcome; PFS calculated from the date of diagnosis to the date of first relapse or the date of the last follow-up<br><br>5. Adjustment for other prognostic factors (a-g): high risk<br><br>Unstated why explorative laparotomy is a category within the RD variable. Model predicting PFS adjusted for Age, carcinomatosis at diagnosis, CA-125 and pathological response to NACT.<br><br>6. Statistical analysis and reporting (a-d): high risk<br><br>No conceptual framework; data driven based on P values of univariate associations. Results for multivariate analysis of PFS not reported using hazard ratios. |
| Notes | Median follow-up: 47 months (range: 3 to 181) |

**Phillips 2018**

| ***Study characteristics*** | |
|---|---|
| Methods | Single-centre retrospective study |
| Participants | N = 398 women undergoing interval debulking surgery (IDS) for stage 3 or 4 epithelial ovarian, tubal or peritoneal cancer (advanced ovarian cancer, AOC). All women were managed by subspecialty trained gynaecological oncologists at the Pan-Birmingham Gynaecological Cancer Centre (PBGCC), Birmingham, United Kingdom<br><br>Mean age: 63.9 (95% CI 42.2 to 85.6) |

**Phillips 2018** *(Continued)*

| | |
|---|---|
| | FIGO: III – 273 (68.6%); IV – 123 (31.4%)<br><br>Histology: serous – 370 (93%); undifferentiated – 1 (0.3%); endometrioid – 1 (0.3%); carcinosarcoma – 12 (3%); mixed – 8 (2%); clear cell – 2 (0.5%); unknown – 4 (1%)<br><br>Tumour grade: G1 – 13 (3.3%); G2 – 2 (0.5%); G3 – 374 (94%); unknown – 9 (2.3%)<br><br>Disease site: ovary – 252 (63.3%); fallopian – 90 (22.6%); primary peritoneal: 56 (14.1%) |
| Residual disease details | ≤ 4 NACT cycles: 231 (58%)<br><br>• Group 1 (≤ 4 cycles) with 111 (48.1%) receiving standard treatment with 3 cycles of NACT and the remaining 120 (51.9%) receiving an additional cycle to facilitate timing of IDS<br><br>≥ 5 NACT cycles: 167 (42%)<br><br>NACT regimen:<br><br>• Carboplatin: 94 (23.6%)<br>• Paclitaxel and carboplatin: 304 (76.4%)<br>• Additional bevacizumab: 25 (6%<br><br>Surgical complexity score:<br><br>• Low (0 to 3): 263 (66.1%)<br>• Inter (4 to 7): 89 (22.4%)<br>• High (8+): 46 (11.6%)<br><br>Median adjuvant CT after IDS: 3 cycles<br><br>**'Optimal' cytoreduction** defined as SVRD or NMRD (RD 0 cm to 1 cm) (n = 310, 77.9%):<br><br>• NMRD: 255 (64.1%)<br>• RD greater than 0 cm but less than 1 cm (RD < 1): 55 (13.8%)<br>• RD of 1 cm and above (RD ≥ 1): 88 (22.1%) |
| Outcomes | Median OS: 40.1 months<br><br>Median OS in NMRD: 51.8 months<br><br>Median OS in SVRD < 1 cm: 29.5<br><br>Median OS in LVRD ≥ 1 cm: 28.9<br><br>Multivariable Cox PH for OS adjusted for FIGO stage, chemotherapy regime (carbo/Taxol vs carboplatin):<br><br>Within group 1 (≤ 4 cycles NACT; n = 231)<br><br>SVRD < 1 cm (vs NMRD): HR 1.5723 (95% CI 0.928 to 2.664), P > 0.05; RD ≥ 1 (vs NMRD): HR 1.7709 (95% CI 1.069 to 2.933), P = 0.0264; SVRD < 1 cm (vs LVRD ≥ 1 cm): HR 0.8879 (95% CI 0.460 to 1.715), P > 0.05<br><br>Within group 2 (> 4 cycles NACT; n = 167)<br><br>SVRD < 1 cm (vs NMRD): HR 2.781 (95% CI 1.663 to 4.650), P = 0.0001; LVRD ≥ 1 (vs NMRD): HR 2.6729 (95% CI 1.759 to 4.062), P < 0.00001; SVRD < 1 cm (vs LVRD ≥ 1 cm): HR 1.04 (95% CI 0.613 - 1.765), P > 0.05 |
| Risk of bias (QUIPS) | 1. Study participation (a-f): low risk<br><br>Adequate number of participants and description of target population. Baseline characteristics, eligibility criteria, sampling frame and period/place study took place presented clearly.<br><br>2. Study attrition (a-e): unclear risk |

**Phillips 2018** *(Continued)*

| | |
|---|---|
| | Unclear if patients with incomplete follow-up were excluded before arriving at the stated sample size. Insufficient information to permit judgement.<br><br>3. Prognostic factor measurement (a-f): low risk<br><br>Valid and reliable measurement of RD<br><br>**Outcome level assessment:**<br><br>**Outcome: overall survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Definition of OS not provided but it usually has a standard definition.<br><br>5. Adjustment for other prognostic factors (a-g): high risk<br><br>Model predicting OS adjusted for FIGO stage, and chemotherapy regime (carbo/Taxol vs carboplatin)<br><br>6. Statistical analysis and reporting (a-d): high risk<br><br>No conceptual framework; unclear how FIGO stage and chemotherapy regime were chosen to be in multivariate model<br><br>**Outcome: progression-free survival**<br><br>Not reported |
| Notes | Median BMI: 25 |

**Polterauer 2012**

| ***Study characteristics*** | |
|---|---|
| Methods | Prospective, multicentre study (5 specialised European centres for gynaecologic oncology)<br><br>Women enrolment between February 2005 and December 2008 |
| Participants | 226 women with epithelial ovarian cancer FIGO Stages IIA to IV in whom radical cytoreductive surgery was performed and standard chemotherapy with paclitaxel and carboplatin was applied. Women having received neoadjuvant chemotherapy followed by interval debulking were excluded<br><br>Mean age 57.5 year (SD 11.9)<br><br>FIGO stages II, III and IV: 15 (6.6%), 174 (76.9%) and 37 (16.4%); FIGO stages IIIC and IV: 198 women (87.6%)<br><br>Histological type serous/other: 194/32<br><br>NMRD: 69.4%<br><br>SVRD (≤ 1 cm): 87.2% (NB: this category also includes NMRD)<br><br>Austria |
| Residual disease details | Residual disease was defined as:<br><br>Any RD (SVRD (≤ 1 cm) or LVRD (> 1 cm)<br><br>Complete debulking (NMRD) |
| Outcomes | 3-year OS (unadjusted) with NMRD: 72.4%; minimal RD: 65.8%; gross RD: 45.2% |

**Polterauer 2012** *(Continued)*

| | |
|---|---|
| | Subgroup analysis of stages IIIC and IV: 3-year OS (unadjusted) with NMRD 69.7% (SE 5.3%); any RD 53.6% (SE 8.3%) (P = 0.003)<br><br>HR (apparently for 'Any RD' vs 'No RD', adjusted for FIGO-stage, histological grade, histological type and age) 1.4 (95% CI 1.0 to 2.1)<br><br>"Multivariable survival analysis revealed residual tumour size (p=0.04) and older women age (p =0.02) as independent prognosticators for impaired overall survival. Complete cytoreduction was predictive for a higher rate of treatment response (p=0.001) and was associated with prolonged progression-free and overall survival (p<0.001 and p=0.001)."<br><br>HR for PFS (apparently for 'Any RD' vs 'NMRD', adjusted for FIGO stage, histological grade, histological type and age) 1.6 (95% CI 1.3 to 2.1)<br><br>Univariate survival analysis of categorical variables by the log-rank test. Multiple forward stepwise Cox regression analysis. |
| Risk of bias (QUIPS) | 1. Study participation (a-f): low risk<br><br>Adequate number of participants and description of target population. Baseline characteristics, eligibility criteria, sampling frame and period/place study took place presented clearly.<br><br>2. Study attrition (a-e): unclear risk<br><br>Unclear if patients with incomplete follow-up were excluded before arriving at the stated sample size. Insufficient information to permit judgement.<br><br>3. Prognostic factor measurement (a-f): unclear risk<br><br>Valid and reliable measurement of RD. Multicentre design may introduce heterogeneity in measurement of RD<br><br>**Outcome level assessment:**<br><br>**Outcome: overall survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Valid and reliable measurement of outcome<br><br>5. Adjustment for other prognostic factors (a-g): low risk<br><br>Cohort was recruited with objective to identify and verify clinical and molecular prognostic/predictive factors in ovarian cancer. Possible confounding prognostic factors would also have been included in study. Multivariate model for OS adjusted for FIGO stage, histological grade, histology subtype and age<br><br>6. Statistical analysis and reporting (a-d): unclear risk<br><br>No conceptual framework; variable selection criteria for multivariate analysis unstated.<br><br>**Outcome: progression-free survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Valid and reliable measurement of outcome<br><br>5. Adjustment for other prognostic factors (a-g): low risk<br><br>Cohort was recruited with objective to identify and verify clinical and molecular prognostic/predictive factors in ovarian cancer. Possible confounding prognostic factors would also have been included in study. Multivariate models for PFS adjusted for FIGO stage, histological grade, histology subtype and age<br><br>6. Statistical analysis and reporting (a-d): unclear risk |

**Polterauer 2012** *(Continued)*

| | |
|---|---|
| | No conceptual framework; variable selection criteria for multivariate analysis unstated |
| Notes | Source of funding: the European commission (FP6 Specific Targeted Research or Innovation Project)<br><br>Declaration of interest: none declared<br><br>Median follow-up: 25.0 months (range: 1 to 49)<br><br>Retrospective non-randomised study. Blinding not reported (but not applicable). Adjusted HRs are derived from a prognostic model. No details on how modelling was performed, but this seems to have been done based on significance testing (and not on including putative confounders in the analysis, irrespective of statistical significance).<br><br>Women and disease characteristics not reported according to debulking status. NB: possible overlap with Hofstetter 2013. |

**Shibutani 2020**

| ***Study characteristics*** | |
|---|---|
| Methods | The purpose of this study was to determine the optimal regimen of neoadjuvant chemotherapy (NAC) for advanced epithelial ovarian, fallopian tube and peritoneal cancers<br><br>Retrospective study of data from the Hyogo Cancer Center between January 2006 and December 2015.<br><br>Japan |
| Participants | 171 patients with advanced epithelial ovarian, fallopian tube or primary peritoneal cancer who underwent dose-dense tri-weekly administration of paclitaxel and carboplatin (TC) or TC as NAC followed by IDS<br><br>The median age of patients was 61 (range 35 to 79) years<br><br>Performance status of patients: 0 for 47 patients (27%); 1 for 79 patients (46%); 2 for 38 patients (22%); and 3 for 7 patients (4%) |
| Residual disease details | Patients who underwent NAC followed by interval debulking surgery<br><br>The median number of NAC cycles was 4 (range 2 to 10). The total number of cycles during the first treatment was 7 (range 4 to 16).<br><br>Dose-dense paclitaxel and carboplatin (TC) was administered in 101 patients (59%); tri-weekly TC was administered 70 patients (41%)<br><br>**Residual disease groups:**<br><br>SVRD < 1 cm: 150 (88%)<br><br>LVRD > 1 cm: 21 (12%) |
| Outcomes | Overall survival and progression-free survival |
| Risk of bias (QUIPS) | 1. Study participation (a-f): low risk<br><br>Adequate number of participants and description of target population. Baseline characteristics, eligibility criteria, sampling frame and period/place study took place presented clearly.<br><br>2. Study attrition (a-e): unclear risk<br><br>Unclear if patients with incomplete follow-up were excluded before arriving at the stated sample size. Insufficient information to permit judgement. |

**Shibutani 2020** *(Continued)*

| | |
|---|---|
| | 3. Prognostic factor measurement (a-f): low risk<br><br>Valid and reliable measurement of RD<br><br>**Outcome level assessment:**<br><br>**Outcome: overall survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Valid and reliable measurement of outcome. The overall survival was calculated from the date of the first chemotherapy to the date of death or last contact.<br><br>5. Adjustment for other prognostic factors (a-g): high risk<br><br>Only univariate analysis of OS. Not included in multivariate analyses.<br><br>6. Statistical analysis and reporting (a-d): high risk<br><br>No multivariate model predicting OS despite there being one for PFS<br><br>**Outcome: progression-free survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Valid and reliable measurement of outcome. Progression-free survival was calculated from the date of the first chemotherapy to the date of death or disease progression.<br><br>5. Adjustment for other prognostic factors (a-g): low risk<br><br>HR for PFS was adjusted for age (< 61 vs ≥ 61), PS (0 to 1 vs 2 to 3), stage (III vs IV), disease (ovary vs others), histology, residual disease, NAC cycles and NAC regimens in multivariable Cox model<br><br>6. Statistical analysis and reporting (a-d): high risk<br><br>No conceptual framework; variable selection criteria for multivariate analysis unstated<br><br>No multivariate model predicting OS despite there being one for PFS |
| Notes | The median observation period was 41 (range 4 to 138) months<br><br>Median progression-free survival was 21 (95% CI 18 to 23) months and 15 (95% CI 13 to 17) months in the dose-dense TC and conventional TC group, respectively (HR 0.69, 95% CI 0.46 to 0.96; P = 0.02)<br><br>The median overall survival was 59 (95% CI 46 to 72) and 40 (95% CI 32 to 57) months in the dose-dense TC group and conventional TC group (HR 0.72, 95% CI 0.48 to 1.06; P = 0.09)<br><br>Multivariate analysis for progression-free survival demonstrated that dose-dense TC represented an independent prognostic factor (HR 0.70, 95% CI 0.50 to 0.99; P = 0.04).<br><br>PFS multivariate prognostic factors were as follows: FIGO stage (HR 0.68, 95% CI 0.48 to 0.96 (table says 0.90); P = 0.03) and residual disease at IDS (HR 0.55, 95% CI 0.34 to 0.96 (table says 0.90 and this appears to be the correct estimate when log estimates are entered; P = 0.02). Also when reference is changed this estimate is: HR 1.82 (95% CI 1.12 to 2.97). |

**Shim 2016**

| ***Study characteristics*** | |
|---|---|
| Methods | Retrospective study |
| Participants | 276 women with FIGO stage III or IV ovarian cancer consecutively treated |

**Shim 2016** *(Continued)*

| | |
|---|---|
| | Median age at diagnosis was 54 years (range: 20 to 80 years)<br><br>258 (93.5%) women received postoperative platinum-based chemotherapy<br><br>South Korea |
| Residual disease details | Speciality of surgeon not reported<br><br>Surgery followed by platinum-taxane chemotherapy<br><br>The 25%, 50% and 75% quartiles of intervals from surgery to start of chemotherapy were 18, 22 and 28 days, respectively |
| Outcomes | Time to chemotherapy (TTC) was analysed and correlated with outcome<br><br>The following were significant prognostic factors for progression-free survival in multivariate analysis:<br><br>• TTC (≤ 28 vs > 28 days; HR 1.578, 95% CI 1.057 to 2.355)<br>• Complete debulking with NMRD (HR 0.419, 95% CI 0.274 to 0.640)<br>• Preoperative albumin level (HR 0.549, 95% CI 0.382 to 0.791) |
| Risk of bias (QUIPS) | 1. Study participation (a-f): high risk<br><br>Abstract only therefore insufficient information on study participation<br><br>2. Study attrition (a-e): unclear risk<br><br>Unclear if patients with incomplete follow-up were excluded before arriving at the stated sample size. Insufficient information to permit judgement.<br><br>3. Prognostic factor measurement (a-f): low risk<br><br>Valid and reliable measurement of RD<br><br>**Outcome level assessment:**<br><br>**Outcome: overall survival**<br><br>Not reported<br><br>**Outcome: progression-free survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Definition of PFS not provided but it usually has a standard definition<br><br>5. Adjustment for other prognostic factors (a-g): high risk<br><br>Time to chemotherapy arbitrarily categorised. Model predicting PFS adjusted for time to chemotherapy and preoperative albumin level.<br><br>6. Statistical analysis and reporting (a-d): high risk<br><br>No conceptual framework; unclear on reasons why the particular specific set of variables were selected for multivariate model. PFS used as outcome but no overall survival. |
| Notes | Findings are from an abstract<br><br>OS not reported<br><br>Mean and median length of follow-up were not reported |

**Shim 2016** *(Continued)*

| | |
|---|---|
| | Although delayed TTC (> 28 days) did not possess prognostic significance in women without postoperative residual disease (n = 94), it significantly correlated with progression-free survival in women with postoperative RD (n = 164, HR 1.893, 95% CI 1.209 to 2.962) |

**Stoeckle 2014**

| ***Study characteristics*** | |
|---|---|
| Methods | Single-centre retrospective study |
| Participants | N = 118 women diagnosed with primary ovarian carcinoma, epithelial cell type (stages IIIC with carcinomatosis and IV) who were treated by NACT + late IDS (after 6 cycles) in the taxane/platinum period (1998 to 2010)<br><br>Median age: 64 (range: 37 to 88)<br><br>FIGO: IIIC – 82 (69%); IV – 36 (31%)<br><br>Histology: serous – 111 (94%); non-serous – 7 (6%)<br><br>Had lymph node assessment: 105 (89%)<br><br>Median node count: 32 (range: 4 to 81)<br><br>Lymph node involvement<br><br>• Positive: 56 (47%)<br>• Negative: 49 (42%)<br>• N/A: 13 (11%) |
| Residual disease details | All women had sampling biopsy.<br><br>• Laparoscopy: 77 (65.3%)<br>• Diagnostic laparotomy: 17 (14.4%)<br><br>Median NACT cycles: 6 (range: 5 to 13)<br><br>NACT regimen<br><br>• Carboplatin – 4 (3.4%);<br>• Paclitaxel and carboplatin – 114 (96.6%)<br><br>All IDS performed by 2 surgeons (co-authors on paper) with experience in ovarian cancer surgery<br><br>Resection categories (other than peritoneal stripping)<br><br>• Salpingo-oophorectomy: 109 (92%)<br>• Total abdominal hysterectomy: 109 (92%)<br>• Omentectomy: 113 (96%)<br>• Appendectomy: 102 (86%)<br>• Pelvic lymph node dissection: 104 (88%)<br>• Aortic lymph node dissection: 93 (79%)<br>• Bowel surgery: 32 (27%)<br>• Other organ resection (spleen, liver, small bowel etc.): 17 (14%)<br><br>Number of resection categories<br><br>Median: 6 |

**Stoeckle 2014** *(Continued)*

| | |
|---|---|
| | Range: 0 to 8<br><br>Standard surgery: 54 (46%)<br><br>Extended surgery: 64 (54%)<br><br>**'Optimal' cytoreduction** defined as RD < 1 cm (n = 111, 94%)<br><br>• NMRD (referred to in study as RD0): 80 (68%)<br>• SVRD (RD 0.1 cm to 1 cm): 31 (26%)<br>• LVRD (RD ≥ 1 cm): 7 (6%)<br><br>* In multivariable analysis, it is NMRD vs RD > 0 cm |
| Outcomes | Overall survival defined as time from date of initial diagnosis to date of death of any cause<br><br>Median OS: 42 months<br><br>Median OS in no macroscopic RD group (RD 0 cm): 50 months<br><br>Median OS in RD > 0: 38 months<br><br>Multivariable Cox PH for OS adjusted for tumour grade, WHO performance status, ASA, bowel surgery (yes/no), FIGO stage:<br><br>RD > 0 cm vs NMRD: HR 2.2 (95% CI 1.2 to 4.0), P = 0.01<br><br>Progression-free survival (PFS) was calculated from the date of initial diagnosis to date of progression. Progression was defined as locoregional or metastatic recurrences after complete remission or progression of disease in women without complete remission.<br><br>Median PFS: 17.2 months<br><br>No multivariable analysis for PFS<br><br>Death within 30 days of surgery: 2 (1.7%) |
| Risk of bias (QUIPS) | 1. Study participation (a-f): low risk<br><br>Adequate number of participants and description of target population. Baseline characteristics, eligibility criteria, sampling frame and period/place study took place presented clearly.<br><br>2. Study attrition (a-e): unclear risk<br><br>Unclear if patients with incomplete follow-up were excluded before arriving at the stated sample size. Insufficient information to permit judgement.<br><br>3. Prognostic factor measurement (a-f): low risk<br><br>Valid and reliable measurement of RD<br><br>**Outcome level assessment:**<br><br>**Outcome: overall survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Valid and reliable measurement of outcome. Overall survival defined as time from date of initial diagnosis to date of death of any cause.<br><br>5. Adjustment for other prognostic factors (a-g): unclear risk<br><br>Model predicting OS adjusted for tumour grade, WHO performance status, ASA, bowel surgery (yes/no), FIGO stage<br><br>6. Statistical analysis and reporting (a-d): unclear risk |

**Stoeckle 2014** *(Continued)*

| | |
|---|---|
| | No conceptual framework; data driven based on P values of univariate associations<br><br>**Outcome: progression-free survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Valid and reliable measurement of outcome. Progression was defined as locoregional or metastatic recurrences after complete remission or progression of disease in women without complete remission.<br><br>5. Adjustment for other prognostic factors (a-g): high risk<br><br>Model predicting PFS was not adjusted for any other prognostic factor<br><br>6. Statistical analysis and reporting (a-d): high risk<br><br>Model predicting PFS was not adjusted for any other prognostic factor |
| Notes | Median follow-up: 37 months<br><br>ASA score:<br><br>• 1: 35 (30%)<br>• 2 to 3: 83 (70%)<br><br>WHO performance status<br><br>• 0 to 1: 80 (68%)<br>• 2 to 3: 38 (32%)<br><br>At IDS, 96 (81%) presented with visible tumour. Median tumour size was 2 mm.<br><br>Median length of hospital stay (all women): 10 (2 to 44)<br><br>Median length of stay (women with complications): 16 (range: 7 to 44)<br><br>Major morbidity was defined as a complication requiring a prolonged hospital stay (more than 10 days), re-hospitalisation or reoperation (by surgery or interventional imaging) needing correction by major medication (e.g. prolonged IV antibiotics or blood transfusion (5 packed red blood cells), or causing death during the first postoperative month<br><br>21 women (18%) had major complications, for a total of 24 major complications<br><br>• Infection: 11<br>• Blood loss needing transfusion > 5 PRBC: 7<br>• Thromboembolic event: 2<br>• Cerebrovascular accident: 1<br>• Myocardial infarction: 1<br>• Bowel obstruction: 1<br>• Chylous ascites: 1<br><br>Rehospitalisation: 10 women<br><br>Reoperation by surgery or imaging techniques: 8 women |

**Tewari 2016**

| ***Study characteristics*** | |
|---|---|
| Methods | Retrospective analysis |

**Tewari 2016** *(Continued)*

| | |
|---|---|
| Participants | 1718 women with newly diagnosed International Federation of Gynecology and Obstetrics stage III and IV ovarian, peritoneal or fallopian tube carcinoma were included in the analysis<br>Median age (years): microscopic (58.5); optimal (60.1); suboptimal (60.2)<br><br>Performance status - frequency (%):<br><br>• Normal, asymptomatic: 848 (49.3%)<br>• Symptomatic, ambulatory: 745 (43.4%)<br>• Symptomatic, in bed < 50%: 125 (7.3%)<br><br>Top-level FIGO stage: III: 1241 (72.2%); IV: 477 (27.8%)<br><br>Histology: serous: 1477 (86%); mixed epithelial: 76 (4.4%); endometrioid: 56 (3.3%); clear-cell/mucinous: 60 (3.5%); other: 24 (1.4%)<br><br>Ascites: no: 346 (20.1%); yes: 1372 (79.9%)<br><br>Progression-free survival status: censored: 268 (15.6%); progression or death: 1450 (84.4%)<br><br>Overall survival status: censored: 840 (48.9%); death: 878 (51.1%)<br><br>USA |
| Residual disease details | Speciality of surgeon was not reported<br><br>Primary cytoreductive surgery followed by platinum based chemotherapy<br><br>Treatment arms: frequency (%)<br><br>• I (standard chemotherapy): 580 (33.8%)<br>• II (concurrent bevacizumab): 570 (33.2%)<br>• III (extended bevacizumab): 568 (33%)<br><br>Residual disease, n (%)<br><br>• NMRD: 85 (4.9%)<br>• SVRD (≤ 1 cm): 701 (40.8%)<br>• LVRD (> 1 cm): 932 (54.2%) |
| Outcomes | Overall survival: HR adjusted for:<br><br>TSIC = 15 days: ≤ 1 cm (AHR 1.41, 95% CI 0.77 to 2.58); > 1 cm (AHR 1.87, 95% CI 1.05 to 3.31)<br><br>Residual = micro, 40 days:<br><br>• Race/ethnicity = White (AHR 1.27, 95% CI 1.15 to 1.40)<br>• Race/ethnicity = Asian (AHR 1.51, 95% CI 1.27 to 1.80)<br>• Race/ethnicity = Black (AHR 1.18, 95% CI 1.00 to 1.40)<br>• Race/ethnicity = Hispanic (AHR 1.18, 95% CI 0.97 to 1.43)<br>• Race/ethnicity = other (AHR 1.41, 95% CI 1.15 to 1.74)<br><br>Residual ≤ 1 cm, 40 days:<br><br>• Race/ethnicity = Asian (AHR 1.17, 95% CI 1.01 to 1.35)<br><br>Residual > 1 cm, 40 days<br><br>• Race/ethnicity = Asian (AHR 1.24, 95% CI 1.07 to 1.44)<br><br>Histology<br><br>• Serous: (AHR 1 - referent)<br>• Mixed epithelial: (AHR 1.33, 95% CI 0.97 to 1.84) |

**Tewari 2016** *(Continued)*

- Endometrioid: (AHR 0.70, 95% CI 0.44 to 1.11)
- Clear-cell/mucinous: (AHR 4.97, 95% CI 2.46 to 10.05)
- Other: (AHR 1.14, 95% CI 0.73 to 1.78)

| | |
|---|---|
| Risk of bias (QUIPS) | 1. Study participation (a-f): low risk<br><br>Adequate number of participants and description of target population. Baseline characteristics, eligibility criteria, sampling frame and period/place study took place presented clearly.<br><br>2. Study attrition (a-e): unclear risk<br><br>Unclear if patients with incomplete follow-up were excluded before arriving at the stated sample size. Insufficient information to permit judgement.<br><br>3. Prognostic factor measurement (a-f): low risk<br><br>Valid and reliable measurement of RD<br><br>**Outcome level assessment:**<br><br>**Outcome: overall survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Definition of OS not provided but it usually has a standard definition<br><br>5. Adjustment for other prognostic factors (a-g): low risk<br><br>Arbitrary dichotomisation of time from surgery to chemotherapy. Multivariate model predicting OS adjusted for age, race, performance status, tumour grade, FIGO stage, histology, ascites, CA-125, time from surgery to chemotherapy and interaction terms<br><br>6. Statistical analysis and reporting (a-d): high risk<br><br>No conceptual framework; unclear of variable selection criteria for multivariate analysis<br><br>**Outcome: progression-free survival**<br><br>Not reported |
| Notes | At 15 days, time to initiation of chemotherapy does not increase the risk of death for any women, whereas at 40 days most women have an increased risk of death. This represents a change-point in increasing time at which some women start to become affected negatively. |

**Tseng 2018**

| ***Study characteristics*** | |
|---|---|
| Methods | Retrospective cohort study |
| Participants | 978 women with stage IIIB to IV ovarian, fallopian tube or primary peritoneal carcinoma<br><br>Median age was 61 years (range: 19 to 95 years)<br><br>FIGO stage - n (%):<br><br>- IIIB: 33 (3%)<br>- IIIC: 761 (78%)<br>- IV: 184 (19%)<br><br>Histology - n (%): |

**Tseng 2018** *(Continued)*

| | |
|---|---|
| | • Serous: 869 (89%)<br>• Other: 109 (11%)<br><br>Estimated blood loss: 700 mL (range: 5 mL to 8000 mL)<br><br>Median hospital length of stay was 8 days (range 1 to 22 days)<br><br>USA |
| Residual disease details | Speciality of surgeon not reported<br><br>All women underwent primary debulking surgery followed by intraperitoneal (IP) chemotherapy in (n = 949, 99%)<br><br>Residual disease was classed as follows:<br><br>• NMRD (defined as complete gross resection (CGR) in study) - 0 mm: 408 (42%)<br>• SVRD (1 to 10 mm): 378 (39%)<br>• LVRD (> 10 mm): 192 (20%) |
| Outcomes | Multivariable analysis of factors associated with PFS adjusted for PDS-year group<br><br>Residual disease:<br><br>• NMRD: (AHR: reference)<br>• SVRD: (AHR 1.393, 95% CI 1.174 to 1.654)<br>• LVRD: (AHR 1.921, 95% CI 1.547 to 2.386)<br><br>Multivariable analysis of factors associated with OS adjusted for PDS-year group<br><br>Residual disease:<br><br>• NMRD: (AHR: reference)<br>• SVRD: (AHR 1.36, 95% CI 1.118 to 1.653)<br>• LVRD: (AHR 1.751, 95% CI 1.378 to 2.224)<br><br>Median operative time was 280 minutes (range, 36 to 893 minutes)<br><br>Median length of hospital stay (LOS) was 8 days (range: 1 to 22 days)<br><br>30-day all-cause mortality was 0.4% (4 deaths) |
| Risk of bias (QUIPS) | 1. Study participation (a-f): low risk<br><br>Adequate number of participants and description of target population. Baseline characteristics, eligibility criteria, sampling frame and period/place study took place presented clearly.<br><br>2. Study attrition (a-e): unclear risk<br><br>Unclear if patients with incomplete follow-up were excluded before arriving at the stated sample size. Insufficient information to permit judgement.<br><br>3. Prognostic factor measurement (a-f): low risk<br><br>Valid and reliable measurement of RD<br><br>**Outcome level assessment:**<br><br>**Outcome: overall survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Valid and reliable measurement of outcome |

**Tseng 2018** *(Continued)*

| | |
|---|---|
| | 5. Adjustment for other prognostic factors (a-g): low risk<br><br>Multivariate models for OS adjusted for age, albumin, FIGO stage, ASA score, histology, BRCA, tumour index, and postoperative intraperitoneal chemotherapy<br><br>6. Statistical analysis and reporting (a-d): high risk<br><br>No conceptual framework; unclear of variable selection criteria for multivariate analysis<br><br>**Outcome: progression-free survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Valid and reliable measurement of outcome<br><br>5. Adjustment for other prognostic factors (a-g): low risk<br><br>Multivariate models for PFS adjusted for age, albumin, FIGO stage, ASA score, histology, BRCA, tumour index and postoperative intraperitoneal chemotherapy<br><br>6. Statistical analysis and reporting (a-d): high risk<br><br>No conceptual framework; unclear of variable selection criteria for multivariate analysis |
| Notes | Median follow-up for the entire cohort was 77.7 months (range: 1.3 to 198 months) |

**Van Geene 1996**

| ***Study characteristics*** | |
|---|---|
| Methods | Prospective cohort study: the 2 groups were defined from data collected prospectively at laparotomy.<br><br>All women with ovarian cancer referred to the departments of gynaecological oncology at 2 hospitals between 1981 and 1989 were entered into prospective surgical studies. |
| Participants | During the 8-year period in the study a total of 256 women with previously untreated primary EOC were referred for consideration of surgery and chemotherapy. 37 women with stage II disease were excluded from this analysis leaving 219 women with stage III to IV disease to form the basis of the study.<br><br>Median age at study entry was 57 years (range: 24 to 75 years)<br><br>180 (82%) and 39 (18%) women had FIGO stage III and IV disease respectively<br><br>Histological cell type was as follows: serous: 134 (61%), endometrioid: 34 (15%), mucinous: 32 (15%), clear cell: 7 (3%), undifferentiated: 12 (6%)<br><br>50 (25%) women had tumour grade classified as being well, 68 (34%) had grade as moderate, 75 (37%) had poor grade and in 9 (4%) women the grade was unknown<br><br>101 (46%) women had GOG performance status 0, 94 (43%) had status 1, 23 (10.5%) women had status 2 and for 1 (0.5%) woman their status was unknown<br><br>Mode of spread was as follows: bulky: 100 (46%), spreading: 119 (54%)<br><br>UK |
| Residual disease details | Reported categories for residual disease were as follows:<br><br>1. RD < 2 cm: 92 (42%) of which 15 were deemed to have had NMRD<br>2. LVRD (> 2 cm): 127 (58%) |

**Van Geene 1996** *(Continued)*

| | |
|---|---|
| | All women received cis-platinum containing chemotherapy at the dose of 75 mg/m$^2$ up to a total of 6 courses depending on response and toxicity |
| Outcomes | Overall survival: HR adjusted for performance status and pattern of spread using Cox model:<br><br>> 2 cm vs < 2 cm: HR 1.83, P < 0.0001<br><br>We requested the exact P value and 95% CI from the study authors but the data were no longer available.<br><br>Table 4 is confusing as no macroscopic RD and less than 2cm RD was compared to > 2 cm. This was grouped in table 2. |
| Risk of bias (QUIPS) | 1. Study participation (a-f): unclear risk<br><br>There was insufficient information to permit judgement<br><br>2. Study attrition (a-e): unclear risk<br><br>There was insufficient information to permit judgement<br><br>3. Prognostic factor measurement (a-f): unclear risk<br><br>There was insufficient information to permit judgement<br><br>**Outcome level assessment:**<br><br>**Outcome: overall survival**<br><br>4. Outcome measurement (a-c): unclear risk<br><br>There was insufficient information to permit judgement<br><br>5. Adjustment for other prognostic factors (a-g): unclear risk<br><br>There was insufficient information to permit judgement<br><br>6. Statistical analysis and reporting (a-d): unclear risk<br><br>There was insufficient information to permit judgement<br><br>**Outcome: progression-free survival**<br><br>Not reported |
| Notes | The 2 groups were defined from data collected prospectively at laparotomy. Women with small-volume (≤ 0.5 cm) but widespread disease (> 10 metastatic nodules) were assigned to the seedling group and women with large-volume disease (> 0.5 cm) spread outside the pelvis were assigned to the bulky disease group. Optimal debulking, i.e. residual disease less than 2 cm, was achieved in 92 (42%) of the women with similar rates between the 2 groups (P = 0.09). Complete macroscopic clearance was achieved in only 15 women, all of which were in the bulky spread group.<br><br>Complete macroscopic clearance (NMRD) was achieved in only 15 women, all of which were in the bulky spread group. |

**Wimberger 2010**

| | |
|---|---|
| ***Study characteristics*** | |
| Methods | Retrospective data set review (retrieved from 3 prospective, randomised phase III trials: AGO-OVAR (OVAR-3/-5/-7)) |

**Wimberger 2010** *(Continued)*

| | |
|---|---|
| Participants | Cohort of women from three prospective, randomised phase III trials: AGO-OVAR (OVAR-3/-5/-7) in between 1995 and 2002<br><br>Previously untreated epithelial ovarian cancer FIGO stage IV, at least 18 years of age and required to have adequate haematologic, renal and hepatic function, defined as follows: absolute neutrophil count (ANC) of at least 1.5 × 109 cells/L, platelet count of at least 100 × 109 cells/L, serum creatinine and bilirubin of no more than 1.25 × upper normal limit<br><br>N = 573, all FIGO stage IV disease: malignant pleural effusion = 214 (37.3%), parenchymal hepatic metastases = 146 (25.5%), other sites disease = 213 (37.2%)<br><br>Median age was 59 years (range 19 to 83); age < 50 (17.6%), 50 to 65 (59.5%), > 65 (22.9%)<br><br>ECOG performance status: 0 (28.2%), 1 (54.6%), 2 (17.2%)<br><br>Histological subtypes; serous (68.2%), endometrioid: (6.9%), mucinous (16.0%)<br><br>Peritoneal carcinomatosis: yes (87.8%), no (12.2)<br><br>France and Germany |
| Residual disease details | Residual disease were defined as:<br><br>• NMRD (12.3%)<br>• SVRD (1 to 10 mm) (29.3%)<br>• LVRD (> 10 mm) (58.4%)<br><br>Women were randomly assigned to one of two treatment arms consisting of either carboplatin or cisplatin and paclitaxel, or a combination of carboplatin and paclitaxel versus the same combination with epirubicin or topotecan. All women were scheduled to receive at least 6 courses of platinum-taxane intravenously every 3 weeks. |
| Outcomes | **Women with stage IV**<br><br>**Kaplan-Meier**<br><br>Median OS (unadjusted) of NMRD 54.6 months, SVRD 25.8 months, LVRD > 1 cm 23.9 months<br><br>Median PFS (unadjusted) of NMRD 19.1 months, SVRD 13.6 months, LVRD (> 1 cm) 11.3 months<br><br>**Multivariant analysis for OS:**<br><br>SVRD vs NMRD: HR 1.87 (95% CI 1.21 to 2.89)<br><br>LVRD (> 1 cm) vs NMRD: HR 2.13 (95% CI 1.40 to 3.23)<br><br>**Multivariate analysis for PFS:**<br><br>SVRD vs NMRD: HR 1.51 (95% CI 1.05 to 2.19)<br><br>LVRD (> 1 cm) vs NMRD: HR 1.82 (95% CI 0.28 to 2.59)<br><br>HRs adjusted for age, performance status, histological type, presence of peritoneal carcinomatosis and multiple sites (Y/N) |
| Risk of bias (QUIPS) | 1. Study participation (a-f): low risk<br><br>Adequate number of participants and description of target population. Baseline characteristics, eligibility criteria, sampling frame and period/place study took place presented clearly.<br><br>2. Study attrition (a-e): unclear risk<br><br>Unclear if patients with incomplete follow-up were excluded before arriving at the stated sample size. Insufficient information to permit judgement. |

**Wimberger 2010** *(Continued)*

| | |
|---|---|
| | 3. Prognostic factor measurement (a-f): low risk<br><br>Valid and reliable measurement of RD<br><br>**Outcome level assessment:**<br><br>**Outcome: overall survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Valid and reliable measurement of outcome<br><br>5. Adjustment for other prognostic factors (a-g): unclear risk<br><br>Age arbitrarily categorised. Multivariate model for OS adjusted for age, ECOG, histology, peritoneal carcinomatosis and number of stage IV disease sites<br><br>6. Statistical analysis and reporting (a-d): high risk<br><br>No conceptual framework; unclear of variable selection criteria for multivariate analysis<br><br>**Outcome: progression-free survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Valid and reliable measurement of outcome<br><br>5. Adjustment for other prognostic factors (a-g): unclear risk<br><br>Age arbitrarily categorised. Multivariate model for PFS adjusted for age, ECOG, histology, peritoneal carcinomatosis and number of stage IV disease sites<br><br>6. Statistical analysis and reporting (a-d): high risk<br><br>No conceptual framework; unclear of variable selection criteria for multivariate analysis |
| Notes | All women with stage IV disease in 3 RCTs:<br><br>OVAR-3 trial (1995 to 1997): 69 women received carboplatin-paclitaxel (7 women had complete resection)<br><br>64 women received cisplatin-paclitaxel (6 women had complete resection)<br><br>OVCAR-5 trial (1997 to 1999: 112 carboplatin-paclitaxel (14 complete resection, 61 LVRD > 1 cm)<br><br>106 carboplatin-paclitaxel-epirubicin (12 complete resection, 63 LVRD > 1 cm)<br><br>OVCAR-7 trial (1999 to 2002): 104 carboplatin-paclitaxel (15 complete resection)<br><br>118 carboplatin-paclitaxel-topotecan (15 complete resection)<br><br>The difference in proportion of women with zero residual disease in all 3 trials is not statistically significant (OVAR-3, P = 0.88, OVAR-5 P = 0.79 and OVAR-7, P = 0.71). No significant trend difference in women recruited during the different time period. No relation between residual disease and the number of applied chemotherapy cycles. Therefore, all 3 trials were considered sufficiently similar to be combined for this study and analysis.<br><br>Median OS was statistically reduced in FIGO stage IV 26.1 months compared to stage IIIC |

**Winter 2007**

***Study characteristics***

**Winter 2007** *(Continued)*

| | |
|---|---|
| Methods | The current study was a retrospective review of data from women treated with platinum and paclitaxel combination chemotherapy on one of 6 prospective randomised clinical trials conducted by GOG: protocols 111, 114, 132, 152, 158 and 172<br><br>GOG 111: included LVRD (> 1 cm) stage III/IV EOC (eligible women = 123)<br><br>GOG 114: included SVRD (< 1 cm) stage III EOC (eligible women = 226)<br><br>GOG 132: included LVRD (> 1 cm) stage III/IV EOC (eligible women = 147)<br><br>GOG 152: included LVRD (> 1 cm) stage III EOC (eligible women = 397)<br><br>GOG 158: included LVRD (> 1 cm) stage III EOC (eligible women = 792)<br><br>GOG 172: included SVRD (≤ 1 cm) stage III EOC (eligible women = 210) |
| Participants | Data from 1895 women with stage III invasive EOC who underwent primary surgical cytoreduction followed by paclitaxel/platinum chemotherapy, while participating in one of six GOG clinical trials, was analysed for the present study<br><br>The median age was 57 years (range: 16 to 86 years)<br><br>All 1895 women had FIGO stage III<br><br>Histological cell type was as follows: serous: 1392 (73.5%), endometrioid: 166 (8.8%), mucinous: 34 (1.8%), mixed epithelial: 142 (7.5%), adenocarcinoma unspecified: 49 (2.6%), clear cell: 62 (3.3%), undifferentiated: 26 (1.4%), other: 24 (1.3%)<br><br>179 (9.5%) women had tumour grade 1, 719 (37.9%) had grade 2 and 997 (52.6%) women had tumour grade 3<br><br>Tumour grade details: 1: 179 (9.5%), 2: 719 (37.9%), 3: 997 (52.6%)<br><br>Ethnicity details: White: 1669 (88.1%), African-American: 111 (5.9%), other: 115 (6.1%) |
| Residual disease details | Reported categories for residual disease were as follows:<br><br>1. NMRD: 437 (23.1%)<br>2. SVRD (0.1 cm to 1 cm): 791 (41.7%)<br>3. LVRD (> 1 cm): 667 (35.2%)<br><br>Optimal was not defined, yet women were divided into 3 groups for analysis, based on RD status (as above). The following chemotherapy schedules were given in the 6 trials:<br><br>• GOG 111: IV paclitaxel 135 mg/m$^2$, cisplatin 75 mg/m$^2$, 6 cycles<br>• GOG 114: IV paclitaxel 135 mg/m$^2$, cisplatin 75 mg/m$^2$, 6 cycles<br>• GOG 132: IV paclitaxel 135 mg/m$^2$, cisplatin 75 mg/m$^2$, 6 cycles<br>• GOG 152: IV paclitaxel 135 mg/m$^2$, cisplatin 75 mg/m$^2$, 6 cycles ± interval debulking<br>• GOG 158: IV paclitaxel 135 mg/m$^2$ (24 hours), cisplatin 75 mg/m$^2$, 6 cycles or IV paclitaxel 175 mg/m$^2$ (3 hours), carboplatin AUC 7.5, 6 cycles<br>• GOG 172: IV paclitaxel 135 mg/m$^2$, cisplatin 75 mg/m$^2$, 6 cycles |
| Outcomes | Overall survival and progression-free survival: HR adjusted for age (discrete), race, GOG performance status, histology and tumour grade using Cox model:<br><br>SVRD vs NMRD: HR 2.11 (95% CI 1.78 to 2.49), P < 0.001 and HR 1.96 (95% CI 1.70 to 2.26), P < 0.001 for OS and PFS respectively<br><br>LVRD (> 1 cm) vs NMRD: HR 2.47 (95% CI 2.09 to 2.92), P < 0.001 and HR 2.36 (95% CI 2.04 to 2.73), P < 0.001 for OS and PFS respectively |

**Winter 2007** *(Continued)*

| | |
|---|---|
| Risk of bias (QUIPS) | 1. Study participation (a-f): low risk<br><br>Adequate number of participants and description of target population. Baseline characteristics, eligibility criteria, sampling frame and period/place study took place presented clearly.<br><br>2. Study attrition (a-e): unclear risk<br><br>Unclear if patients with incomplete follow-up were excluded before arriving at the stated sample size. Insufficient information to permit judgement.<br><br>3. Prognostic factor measurement (a-f): low risk<br><br>Valid and reliable measurement of RD<br><br>**Outcome level assessment:**<br><br>**Outcome: overall survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Valid and reliable measurement of outcome<br><br>5. Adjustment for other prognostic factors (a-g): unclear risk<br><br>Age arbitrarily categorised. Model predicting OS adjusted for age, race, GOG performance status, histology, and tumour grade<br><br>6. Statistical analysis and reporting (a-d): unclear risk<br><br>In methods, authors reported that all variables considered as potential prognostic factors were included in multivariate analyses, suggesting some conceptual framework<br><br>**Outcome: progression-free survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Valid and reliable measurement of outcome<br><br>5. Adjustment for other prognostic factors (a-g): unclear risk<br><br>Age arbitrarily categorised. Model predicting PFS adjusted for age, race, GOG performance status, histology, and tumour grade.<br><br>6. Statistical analysis and reporting (a-d): unclear risk<br><br>In methods, authors reported that all variables considered as potential prognostic factors were included in multivariate analyses, suggesting some conceptual framework. |
| Notes | 1505 recurrences and 1323 deaths were identified during a median follow-up period of 43 months:<br><br>The median PFS was 17.1 months (95% CI 16.4 to 17.8 months)<br><br>The median OS was 45.3 months (95% CI 43.0 to 47.7 months)<br><br>**PFS for disease residual:**<br>NMRD: N = 437, PFS was 33.0 months, 0.1 cm to 1.0 cm: N = 791, PFS) was 16.8 months, LVRD (> 1 cm): N = 667, PFS was 14.1 months, P < 0.001<br><br>**OS for disease residual:**<br>NMRD: N = 437, OS was 71.9 months, SVRD: N = 791, OS was 42.4 months, LVRD (> 1.0 cm): N = 667, OS was 35.0 months, P < 0.001<br><br>Increasing age was associated with decreased PFS and OS. Median PFS and OS were shorter for women with a performance status (PS) of 1 or 2 when compared with those with a PS of 0. No difference in median PFS was evident between PS 1 and PS 2 women, whereas the difference in median OS between the |

**Winter 2007** *(Continued)*

| | |
|---|---|
| | same groups was observed. Based on tumour histology, women with endometrioid histology had improved clinical outcomes compared with those with serous tumours. Women with mucinous or clear-cell tumours had decreased PFS and OS. Women with mucinous cell type had a median OS of only 15 months compared with 24, 45 and 56 months for clear-cell, serous and endometrioid cell types, respectively.<br><br>Women with NMRD had the longest PFS and OS 33 and 72 months, respectively compared with women with any gross residual disease. The differences in median PFS and OS between the SVRD and LVRD (> 1 cm) groups were also evident, albeit small (3 months in median PFS and 7 months in median OS). Women with grade 2 or 3 tumours were associated with decreased PFS and OS. Race was not significantly associated with PFS or OS. |

**Winter 2008**

| ***Study characteristics*** | |
|---|---|
| Methods | Retrospective review of 4 RCTs. The current study was a retrospective review of data from women with stage IV EOC treated with platinum and paclitaxel combination chemotherapy on one of four prospective randomised clinical trials conducted by the GOG: protocols 111, 132, 152 and 162 |
| Participants | 360 women with stage IV invasive EOC who underwent primary surgical cytoreduction followed by paclitaxel/platinum chemotherapy while participating in one of four GOG clinical trials.<br><br>The median age of women was 59 years (range: 24 to 86 years)<br><br>317 (88%) women were white, 28 (8%) were black and 15 (4%) were of other ethnic origin<br><br>97 (27%) had GOG performance status 0, 203 (56%) had status 1 and 60 (17%) had status 2<br><br>24 (7%) women had tumour grade 1, 112 (31%) grade 2 and 224 (62%) had grade 3 disease<br><br>Histology was as follows: serous 268 (74.5%), endometrioid 28 (8%), mucinous 7 (2%), clear cell 12 (3%), adenocarcinoma unspecified 9 (2.5%), mixed epithelial 22 (6%), undifferentiated 9 (2.5%), other 5 (1.5%).<br><br>The median residual tumour size was 3 cm (range 0.0 to 40.0)<br><br>Stage IV disease site was as follows: distant: 45 (12.5%), parenchymal liver: 64 (17.75%), pleural effusion: 172 (47.75%), subcutaneous: 32 (9%), others: 3 (1%), multiple sites: 44 (12%) |
| Residual disease details | The maximum diameter of residual tumour that was used to define optimal cytoreduction: 1 cm (in original RCTs). All 4 RCTs included suboptimal disease (> 1 cm).<br><br>Residual disease was noted as follows:<br><br>• NMRD: 29 (8%)<br>• SVRD of 0.1 cm to 1 cm: 78 women (22%)<br>• LVRD of 1.1 cm to 2 cm: 50 women (14%)<br>• LVRD of 2.1 cm to 3 cm: 40 women (11%)<br>• LVRD of 3.1 cm to 4 cm: 30 women (8.25%)<br>• LVRD of 4.1 cm to 5 cm: 44 women (12%)<br>• LVRD of 5.1 cm to 6 cm: 30 women (8.25%)<br>• LVRD larger than 6 cm: 59 women (16.5%)<br><br>'Optimal' cytoreduction was defined as RD < 1 cm and a sensitivity analysis was performed defining RD as < 2 cm |

**Winter 2008** *(Continued)*

| | |
|---|---|
| | All women were treated with primary surgical cytoreduction and 6 cycles of a 24-hour infusion of intravenous paclitaxel 135 mg/m$^2$, followed by intravenous cisplatin 75 mg/m$^2$ |
| Outcomes | • Overall survival: HR adjusted for several prognostic categories<br>• Optimal: NMRD:<br>  ◦ SVRD (< 1 cm) vs NMRD: HR 1.93 (95% CI 1.17 to 3.20)<br>  ◦ 1 cm to 5cm vs NMRD: HR 1.83 (95% CI 1.14 to 2.94)<br>  ◦ > 5 cm vs NMRD: HR 2.72 (95% CI 1.65 to 4.47)<br>• Optimal: SVRD (≤ 1.0 cm):<br>  ◦ LVRD (> 1 cm) HR 1.30 (95% CI 1.00 to 1.59)<br>• Optimal: ≤ 2 cm RD:<br>  ◦ LVRD (> 2 cm) HR 1.17 (95% CI 0.92 to 1.49)<br>• Progression-free survival: HR adjusted for several prognostic categories<br>• Optimal: NMRD:<br>  ◦ SVRD (< 1 cm) vs NMRD: HR 1.99 (95% CI 1.24 to 3.18)<br>  ◦ 1 cm to 5cm vs NMRD: HR 2.15 (95% CI 1.38 to 3.34)<br>  ◦ > 5 cm vs NMRD: HR 2.96 (95% CI 1.86 to 4.71)<br>• Optimal: SVRD (≤ 1 cm) RD:<br>  ◦ LVRD (> 1 cm) HR: 1.49 (95% CI 1.16 to 1.92)<br>• Optimal: ≤ 2.0 cm RD:<br>  ◦ LVRD (> 2 cm) HR: 1.27 (95% CI 1.01 to 1.61) |
| Risk of bias (QUIPS) | 1. Study participation (a-f): low risk<br><br>Adequate number of participants and description of target population. Baseline characteristics, eligibility criteria, sampling frame and period/place study took place presented clearly.<br><br>2. Study attrition (a-e): unclear risk<br><br>Unclear if patients with incomplete follow-up were excluded before arriving at the stated sample size. Insufficient information to permit judgement.<br><br>3. Prognostic factor measurement (a-f): low risk<br><br>Valid and reliable measurement of RD<br><br>**Outcome level assessment:**<br><br>**Outcome: overall survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Valid and reliable measurement of outcome<br><br>5. Adjustment for other prognostic factors (a-g): unclear risk<br><br>Multivariate model for OS adjusted for histology and stage IV disease site<br><br>6. Statistical analysis and reporting (a-d): unclear risk<br><br>In methods, authors reported that all variables considered as potential prognostic factors were included in multivariate analyses, suggesting some conceptual framework. However, age, race, GOG PS and tumour grade were excluded secondary at univariate analysis due to their P values falling above significance threshold<br><br>**Outcome: progression-free survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Valid and reliable measurement of outcome<br><br>5. Adjustment for other prognostic factors (a-g): unclear risk |

**Winter 2008** *(Continued)*

| | |
|---|---|
| | Multivariate model for PFS adjusted for histology and stage IV disease site<br><br>6. Statistical analysis and reporting (a-d): unclear risk<br><br>In methods, authors reported that all variables considered as potential prognostic factors were included in multivariate analyses, suggesting some conceptual framework. However, age, race, GOG PS and tumour grade were excluded secondary at univariate analysis due to their P values falling above significance threshold |
| Notes | The median length of follow-up was 28 months<br><br>When evaluating the association of clinicopathologic factors with residual disease status, there was no difference between the RD groups and demographic, clinical and pathologic factors<br><br>Stage IV site did not seem to have significant association with RD group distributions |

**Zhang 2018**

| ***Study characteristics*** | |
|---|---|
| Methods | Single-centre, retrospective study undertaken on women treated between January 2003 and December 2013, at the Department of Gynecology, Weifang Yidu Central Hospital, China |
| Participants | N = 200 women diagnosed with stage IIIC to IV invasive ovarian, fallopian tube or peritoneal high-grade serous carcinoma, who were treated with platinum-based NAC followed by IDS and adjuvant chemotherapy.<br><br>Median age: 61 (range: 38 to 80)<br><br>FIGO: IIIC – 169 (84.5%); IV – 31 (15.5%)<br><br>Pre-operative ascites<br><br>• < 500 mL: 116 (58%)<br>• ≥ 500 mL: 84 (42%)<br><br>Median CA-125 at diagnosis: 952 U/mL (range: 75 to 23,400)<br><br>Median pre-operative CA-125: 572 (range: 43 to 986)<br><br>Median CA-125 decreasing kinetics (ratio of the initial serum CA-125 level to the preoperative serum CA-125 level): 2.3 (range: 0.8 to 30.2)<br><br>≤ 3 tumour sites: 50 (25%)<br><br>> 3 tumour sites: 150 (75%) |
| Residual disease details | Median NACT cycles: 3 (range: 1 to 8)<br><br>NAC was administrated intraperitoneally for 90 (45%) women and intravenously for 110 (55%) women<br><br>Median adjuvant CT cycles: 5 (range: 3 to 7)<br><br>**'Optimal' cytoreduction** defined as RD < 1 cm (n = 156, 78%):<br><br>• NMRD (referred to in study as RD0): 59 (29.5%)<br>• SVRD (RD < 1 cm): 97 (48.5%)<br>• LVRD between 1 cm to 2 cm inclusive: 8 (4%)<br>• LVRD (> 2 cm): 30 (15%)<br>• Unknown: 6 (3%) |

**Zhang 2018** *(Continued)*

| | |
|---|---|
| Outcomes | Overall survival defined as interval between treatment initiation and death<br><br>Median OS in participants with ascites regression: 32.1<br><br>Median OS in participants without ascites regression: 25.2<br><br>Multivariable Cox PH for OS adjusted for pre-operative ascites, number of tumour sites, CA-125 at diagnosis, CA-125 decreasing kinetics:<br><br>• LVRD (> 1 cm (vs SVRD < 1 cm): HR 2.58, 95% CI 1.71 to 4.24), P < 0.01<br><br>Progression-free survival defined as interval between the beginning of treatment and documented disease progression or death from any cause in women with no evidence of progression<br><br>Median PFS in participants with ascites regression: 22.3<br><br>Median PFS in participants without ascites regression: 18<br><br>Multivariable Cox PH for PFS adjusted for pre-operative ascites, number of tumour sites, number of NAC cycles, CA-125 at diagnosis, CA-125 decreasing kinetics:<br><br>• LVRD (> 1 cm (vs SVRD): HR 2.43, 95% CI 1.44 to 4.08), P < 0.01 |
| Risk of bias (QUIPS) | 1. Study participation (a-f): low risk<br><br>Adequate number of participants and description of target population. Baseline characteristics, eligibility criteria, sampling frame and period/place study took place presented clearly.<br><br>2. Study attrition (a-e): unclear risk<br><br>Unclear if patients with incomplete follow-up were excluded before arriving at the stated sample size. Insufficient information to permit judgement.<br><br>3. Prognostic factor measurement (a-f): low risk<br><br>Valid and reliable measurement of RD<br><br>**Outcome level assessment:**<br><br>**Outcome: overall survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Valid and reliable measurement of outcome. OS defined as interval between treatment initiation and death.<br><br>5. Adjustment for other prognostic factors (a-g): unclear risk<br><br>Baseline CA-125 and preoperative CA-125 are likely to introduce multicollinearity. Model predicting OS adjusted for pre-operative ascites, number of tumour sites, CA-125 at diagnosis, CA-125 decreasing kinetics<br><br>6. Statistical analysis and reporting (a-d): unclear risk<br><br>No conceptual framework; data driven based on P values of univariate associations<br><br>**Outcome: progression-free survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Valid and reliable measurement of outcome. PFS defined as interval between the beginning of treatment and documented disease progression or death from any cause in women with no evidence of progression.<br><br>5. Adjustment for other prognostic factors (a-g): unclear risk |

**Zhang 2018** *(Continued)*

| | |
|---|---|
| | Baseline CA-125 and preoperative CA-125 are likely to introduce multicollinearity. Model predicting PFS adjusted for age, preoperative ascites, FIGO stage, tumour sites, baseline CA-125, preoperative CA-125, number of NACT cycles and CA-125 decreasing kinetics<br><br>6. Statistical analysis and reporting (a-d): unclear risk<br><br>No conceptual framework; data driven based on P values of univariate associations |
| Notes | Median follow-up: 43.5 months<br><br>Ascites regression defined as an ascites volume of less than 500 mL<br><br>**Inclusion criteria**<br><br>(i) Women histologically diagnosed as stage IIIc or IV invasive ovarian, fallopian tube or peritoneal high-grade serous carcinoma; (ii) women treated with platinum-based NAC followed by IDS and adjuvant chemotherapy; and (iii) women with an ascites volume of greater than or equal to 500 mL before NAC treatment as assessed by ultrasound examination<br><br>**Exclusion criteria**<br><br>(i) Fragile women who received slow-release evacuation procedure before NAC due to intolerable abdominal distension; (ii) women with extra-abdominal metastatic malignancy; and (iii) women whose preoperative serum cancer antigen 125 (CA-125) levels were less than or equal to 35 U/ mL<br><br>**Treatment protocol**<br><br>A NAC regimen consisting of carbo-platinum (area under the curves 5 to 6) and paclitaxel (135 to 175 mg/ $m^2$) was administered every 3 weeks. IDS was performed approximately 2 to 4 weeks after the NAC regimen. The adjuvant chemotherapy (at least 3 to 4 cycles) was the same as NAC.<br><br>The standard IDS included bilateral/unilateral salpingo-oophorectomy, hysterectomy, appendectomy, pelvic/para-aortic lymphadenectomy and omentectomy. Extensive upper abdominal surgery was defined as splenectomy, diaphragm stripping and/or resection, distal pancreatectomy, cholecystectomy, partial liver resection and partial gastrectomy. Other surgery procedures, such as large/small bowel resection and peritoneal resection, were performed as necessary. |

**Zhu 2016**

| ***Study characteristics*** | |
|---|---|
| Methods | Multicentre, retrospective study |
| Participants | N = 672 women newly diagnosed with epithelial ovarian cancer between June 2008 and December 2015 at the Sun Yat-Sen University Cancer Center and Nan Fang Hospital of Southern Medical University, who were treated with NACT followed by IDS<br><br>Median age: 55 (range: 30 to 70)<br><br>FIGO: III – 564 (83.9%); IV – 108 (16.1%)<br><br>Histology: serous – 484 (72%); non-serous – 188 (28%)<br><br>Tumour grade:<br><br>• G1 – 384 (57.1%)<br>• G2/3 – 288 (42.9%)<br><br>CA-125 at diagnosis, U/mL:<br><br>• ≤ 35: 226 (33.6%) |

**Zhu 2016** *(Continued)*

| | |
|---|---|
| | • > 35: 446 (66.4%)<br><br>Comorbidity:<br><br>• Chronic hepatitis B: 64 (9.5%)<br>• Hypertension: 35 (5.2%)<br>• Diabetes: 27 (4%)<br>• Cardiovascular disease: 4 (0.6%)<br><br>Chemosensitivity (RECIST complete/partial response): 444 (66.1%)<br><br>Chemoresistance: 228 (33.9%) |
| Residual disease details | All participants given 3 cycles of NACT before IDS<br><br>NACT regimen<br><br>• Cisplatin plus paclitaxel: 298 (44.3 %)<br>• Carboplatin plus paclitaxel: 250 (37.2 %)<br>• Carboplatin plus docetaxel: 124 (18.5 %)<br><br>Complete response to NACT (NMRD) in 61 (9.1%)<br><br>**'Optimal' cytoreduction** was defined as RD ≤ 1 cm (n = 486; 72.3%) |
| Outcomes | Overall survival defined as interval between the date of diagnosis and the date of death from any cause or last follow-up<br><br>5-year OS: 36.7%<br><br>Multivariable Cox PH for OS adjusted for FIGO stage, chemosensitivity, Glasgow prognostic score:<br><br>• LVRD (> 1 cm) (vs SVRD): HR 1.332 (95% CI 1.057 to 1.679), P = 0.015<br><br>Progression-free survival defined as time from the date of diagnosis to the date of first relapse, progression, death from any cause or last follow-up<br><br>5-year PFS: 19.3%<br><br>Multivariable Cox PH for PFS adjusted for FIGO stage, chemosensitivity, Glasgow prognostic score:<br><br>• LVRD (> 1 cm) (vs SVRD): HR 1.268 (95% CI 1.051 to 1.589), P = 0.044 |
| Risk of bias (QUIPS) | 1. Study participation (a-f): low risk<br><br>Adequate number of participants and description of target population. Baseline characteristics, eligibility criteria, sampling frame and period/place study took place presented clearly.<br><br>2. Study attrition (a-e): unclear risk<br><br>Unclear if patients with incomplete follow-up were excluded before arriving at the stated sample size. Insufficient information to permit judgement.<br><br>3. Prognostic factor measurement (a-f): unclear risk<br><br>Valid and reliable measurement of RD. Multicentre design may introduce heterogeneity in measurement of RD<br><br>**Outcome level assessment:**<br><br>**Outcome: overall survival**<br><br>4. Outcome measurement (a-c): low risk |

**Zhu 2016** *(Continued)*

| | |
|---|---|
| | Valid and reliable measurement of outcome. OS defined as interval between the date of diagnosis and the date of death from any cause or last follow-up.<br><br>5. Adjustment for other prognostic factors (a-g): high risk<br><br>Age arbitrarily dichotomised. Multivariate models for OS adjusted for FIGO stage, chemosensitivity and Glasgow Prognostic Score<br><br>6. Statistical analysis and reporting (a-d): high risk<br><br>No conceptual framework; unclear on how variables were brought forward to multivariate model<br><br>**Outcome: progression-free survival**<br><br>4. Outcome measurement (a-c): low risk<br><br>Valid and reliable measurement of outcome. PFS defined as time from the date of diagnosis to the date of first relapse, progression, death from any cause or last follow-up<br><br>5. Adjustment for other prognostic factors (a-g): high risk<br><br>Age arbitrarily dichotomised. Multivariate models for PFS adjusted for FIGO stage, chemosensitivity and Glasgow Prognostic Score.<br><br>6. Statistical analysis and reporting (a-d): high risk<br><br>No conceptual framework; unclear on how variables were brought forward to multivariate model |
| Notes | Median follow-up: 38 months (range: 5 to 103)<br><br>ECOG PS (Eastern Cooperative Oncology Group performance status)<br><br>≤ 1: 494 (73.5%); > 2: 178 (26.5%)<br><br>Definition of Glasgow prognostic score: women in whom an elevated CRP level (> 10 mg/L) and hypoalbuminaemia (< 35 g/L) were both present were allocated a score of 2. Women with only one of these two biochemical abnormalities were given a score of 1. Women with neither of these abnormalities received a score of 0. |

1. Study participation
   a. Adequate participation in the study by eligible persons
   b. Description of the target population or population of interest
   c. Description of the baseline study sample
   d. Adequate description of the sampling frame and recruitment
   e. Adequate description of the period and place of recruitment
   f. Adequate description of inclusion and exclusion criteria
2. Study attrition
   a. Adequate response rate for study participants
   b. Description of attempts to collect information on participants who dropped out
   c. Reasons for loss to follow-up are provided
   d. Adequate description of participants lost to follow-up
   e. There are no important differences between participants who completed the study and those who did not
3. Prognostic factor measurement
   a. A clear definition or description of the PF is provided
   b. Method of PF measurement is adequately valid and reliable
   c. Continuous variables are reported or appropriate cutpoints are used
   d. The method and setting of measurement of PF is the same for all study participants
   e. Adequate proportion of the study sample has complete data for the PF
   f. Appropriate methods of imputation are used for missing PF data

4. Outcome measurement
   a. A clear definition of the outcome is provided
   b. Method of outcome measurement used is adequately valid and reliable
   c. The method and setting of outcome measurement is the same for all study participants
5. Adjustment for other prognostic factors
   a. All other important PFs are measured
   b. Clear definitions of the important PFs measured are provided
   c. Measurement of all important PFs is adequately valid and reliable
   d. The method and setting of PF measurement are the same for all study participants
   e. Appropriate methods are used to deal with missing values of PFs, such as multiple imputation
   f. Important PFs are accounted for in the study design
   g. Important PFs are accounted for in the analysis
6. Statistical analysis and reporting
   a. Sufficient presentation of data to assess the adequacy of the analytic strategy
   b. Strategy for model building is appropriate and is based on a conceptual framework or model
   c. The selected statistical model is adequate for the design of the study
   d. There is no selective reporting of result

**Overall risk of bias judgements were made per outcome for each included study**

**Abbreviations:**

ACCI: age-adjusted Charlson Comorbidity Index; AHR: adjusted hazard ratio; AOC: advanced ovarian cancer; ASA: American Society of Anaesthesiologists; BMI: body mass index; CDC: Clavien-Dindo classification; CI: confidence interval; CPLN: cardiophrenic lymph nodes; CRP: c-reactive protein; CRS: cytoreductive surgery; DSS: disease-specific survival; ECOG: Eastern Cooperative Oncology Group; EOC: epithelial ovarian cancer; FIGO: International Federation of Gynecology and Obstetrics; GOG: Gynaecologic Oncology Group; HR: hazard ratio; ICU: intensive care unit; IDS: interval debulking surgery; IP: intraperitoneal; IQR: interquartile range; IV: intravenous; KM: Kaplan–Meier; LVRD: large-volume residual disease; NACT/ACT/CT: neoadjuvant chemotherapy/adjuvant chemotherapy/chemotherapy; NMRD: no macroscopic residual disease; NOS: not otherwise specified; OR: odds ratio; OS: overall survival; PCI: Peritoneal Cancer Index; PDS: primary debulking surgery; PH: proportional hazards; OS: overall survival; PF: prognostic factor; PFS: progression-free survival; RD: residual disease; RR: risk ratio; RT: residual tumour; SD: standard deviation; SE: standard error; SVRD: small-volume residual disease; TSIC: time from surgery to initiation of chemotherapy; UTI: urinary tract infection; WHO: World Health Organization

## Characteristics of excluded studies *[ordered by study ID]*

| Study | Reason for exclusion |
|---|---|
| Alberts 1993 | No survival analysis by RD as all patients had suboptimal surgery (defined as more than 2 cm) |
| Alberts 1996 | No multivariate analysis data |
| Alphs 2006 | Included only 78 patients; 8 patients were early-stage and 9 patients received NAC |
| Altman 2012 | No multivariate analysis data |
| Andersen Soegaard 2005 | This study included only 83 patients, of which 66 received platinum-based chemotherapy. No multivariate analysis was performed. |
| Anuradha 2016 | Scope of study focused on time interval between surgery and chemotherapy |
| Bailey 2006 | Chemotherapy data are absent |
| Baker 1994 | 95% CI or SE (HR) are not reported and the HR point estimate for OS is 1.66 across all categories; it is not clear if the < 1 cm category was used as the reference group when compared to both 1 cm to 2 cm and > 2 cm residual disease |
| Barda 2004 | 27.3% of ovarian cancer received non-platinum chemotherapy |
| Benedetti-Panici 1996 | Included only 66 patients and stage IIb. No survival data per RD. Also included NAC/IDS. |

| Study | Reason for exclusion |
|---|---|
| Bertelsen 1990 | Study does not include a multivariate analysis |
| Bertelsen 1993 | No survival data per residual disease |
| Bian 2016 | No multivariate analysis data |
| Bonnefoi 1999 | 38 patients had NAC and 27 patients had non-platinum chemotherapy |
| Brinkhuis 1996a | No direct comparison by size of residual disease and there is no multivariate analysis |
| Brinkhuis 1996b | 1 group of patients did not receive platinum chemotherapy except at progression. Survival data per RD is reported for all patients collectively. |
| Bristow 1999 | Included only 84 patients |
| Cai 2007 | Included 95 patients. We suspect that IDS cases were included. |
| Ceresoli 2018 | Included only 56 patients at analysis, of which 28 treated with cytoreductive surgery + HIPEC and 28 treated with cytoreductive surgery alone. |
| Chekman 2015 | Did not report outcomes for extent of residual disease by type of initial primary surgery |
| Clamp 2018 | No multivariate analysis data |
| Colozza 1997 | Included only 39 patients |
| Conte 1991 | No survival data per residual disease |
| Conte 1996 | There is no optimal group. No survival data per residual disease. |
| Crawford 2005 | 18% of the cases were stage IC and II |
| Creasman 1990 | All cases were sub-optimal, defined as RD greater than 1 cm; no analysis by RD |
| Cummins 2019 | Full text unavailable |
| Dao 2016 | Included patients who had neoadjuvant chemotherapy |
| Del Campo 1994 | Included only 91 patients |
| de Oliviera 1990 | 1 arm did not receive platinum-based chemotherapy |
| di Re 1996 | 14 patients had borderline tumours. Also included stage II cases. Before 1979, patients received non-platinum chemotherapy. |
| Elgamal 2019 | Full text unavailable |
| Fagotti 2020 | Did not report outcomes for extent of residual disease by type of initial primary surgery |
| Gao 2001 | Only 31 cases |
| Gasimli 2016 | Included selective group of women with cytoreduction of tumour to microscopic optimal disease (0 cm) |
| Geisler 2004 | 24 patients were stage I and II |

| Study | Reason for exclusion |
|---|---|
| Gershenson 1989 | Included only 50 patients |
| Gershenson 1992 | All patients were optimal, defined as RD less than 2 cm. No further analysis of survival by RD. |
| Gershenson 1995 | Included only 51 patients |
| Greggi 2016 | RD thresholds were not part of scope as the study focused on comparison of oncology specialist centres versus non-specialist centres |
| Grem 1991 | Included only 43 patients |
| Hainsworth 1990 | Included only 25 patients |
| Hakes 1992 | Included only 78 patients |
| Hamid 2002 | Only included 62 patients |
| Hardy 1991 | Included only 30 stage IV patients |
| Heitz 2016 | No multivariate analyses were reported |
| Hoskins 1992 | All patients are optimal, i.e. less than 1 cm. Survival data is per preoperative disease volume rather than RD. |
| Hoskins 1996 | Included only 29 patients |
| Hoskins 1997 | No survival by residual disease |
| Itamochi 2002 | Optimal surgery, i.e. size of RD, is not properly defined |
| Kaern 2005 | Included only 31 stage III patients with no control group having RD more than 1 cm |
| Kehoe 2015 | Comparisons of residual disease were based on type of intervention |
| Kessous 2017 | No multivariate analysis data |
| Keyver-Paik 2016 | No multivariate analyses were reported |
| Kirmani 1994 | Included only 29 patients |
| Kristensen 1995 | Included only 27 patients |
| Le 1997 | Data for stage IIIC and IV subgroup was not reported and authors no longer had access to these data |
| Lee 2018 | No multivariate analyses were reported and no response from corresponding author after request for adjusted estimates |
| Loizzi 2016 | Included only 78 patients |
| Lorusso 1998 | Included only 34 patients |
| Malik 1998 | Included only 21 patients |
| Marchetti 1993 | Included only 70 patients |

| Study | Reason for exclusion |
|---|---|
| McGuire 1996 | No multivariate analyses were reported |
| Michaan 2018 | Chemotherapy response score not same as optimal cytoreduction |
| Ngan 1989 | Contained 65 patients only and 15 patients were excluded, so only 50 patients |
| Omura 1989 | 95% CIs and P values from Cox model in adjusted estimates are not reported. Cannot use Parmar's methods given the number of deaths and log rank P value as we need the unadjusted estimate. |
| Onda 2020 | Did not report outcomes for extent of residual disease by type of initial primary surgery |
| Palmer 1992 | Included only 70 patients |
| Piver 1991 | 43 patients did not receive platinum-based chemotherapy. No multivariate analysis. |
| Raspagliesi 2018 | No multivariate analysis data |
| Redman 1986 | Included 89 patients, 11 of whom initially did not receive platinum chemotherapy |
| Risum 2012 | Only 17 women went through NACT-IDS |
| Rodriguez 2013 | Comparisons were in terms of surgical procedures performed and could not be analysed by residual disease thresholds |
| Rose 2004 | Reported on outcome after ''secondary'' debulking surgery. However, Winter 2007 included the results of GOG 152 by residual diease after primary cytoreductive surgery. This has been confirmed through personal communication with GOG statistician (Dr Mark Brady). |
| Ruscito 2016 | Study did not distinguish between PDS and IDS |
| Rutten 2014 | 17% of sample made up of FIGO I and II |
| Salani 2007 | Case-control study |
| Sessa 1991 | No multivariate analysis performed |
| Shapiro 1998 | Included only 26 patients |
| Shinozuka 1999 | Some patients received preoperative chemotherapy |
| Sioulas 2017 | Included women who received combination of intravenous/intraperitoneal chemotherapy and RD was not adequately reported in multivariate analyses |
| Skarlos 1996 | Included patients with stage IIC disease |
| Smits 2015 | Scope of study focused on obese and non-obese patients and included proportion of women who received neoadjuvant chemotherapy |
| Solmaz 2015 | Did not report survival by residual disease |
| Son 2017 | Included only 60 patients |
| Stewart 2015 | Full text unavailable |
| Stewart 2016 | No multivariate analysis |

| Study | Reason for exclusion |
|---|---|
| Strauss 1996 | Included 42 patients only |
| Suidan 2015 | Reported in abstract form only and unlikely that residual disease thresholds were assessed in appropriate multivariate analyses |
| Sun 2000 | Patients who did not receive preoperative chemotherapy are only 76. Nature of chemotherapy received not clear. |
| Sutton 1989 | Included only 56 patients |
| Takano 2006 | Most patients had early-stage disease, which cannot be separated from late-staged cases |
| Takano 2007 | Included early-stage disease (stage IC and II), which cannot be separated from late-staged cases |
| Tay 1996 | Included 62 patients only. Did not include survival data per optimal versus suboptimal. |
| Taylor 1994 | Included only 64 patients |
| Tingulstad 2003 | 6 patients did not receive chemotherapy and 6 patients received non-platinum chemotherapy |
| Todo 2003 | Included patients who received NAC and IDS but did not report by extent of disease |
| Trhlík 2013 | Full text unavailable |
| Uyar 2005 | 18 patients were stage I and II. No survival data per RD. |
| Vallejos 1997 | Included only 30 patients |
| Van Der Burg 1996 | Reported results per residual disease after NAC/IDS |
| Van Driel 2017 | Non-platinum based chemotherapy was given to all the women |
| van Vliet 2015 | Included patients with who received IDS |
| Vergote 2010 | Comparisons of residual disease were based on type of intervention |
| Vergote 2018 | Comparisons of residual disease were based on type of intervention |
| Vidal 2016 | No multivariate analyses were reported |
| Wadler 1996 | Survival reported per residual disease in all patients including 118 who received non-platinum chemotherapy |
| Wallace 2017 | No multivariate analyses were reported |
| Warwick 1995 | 31 patients were stage II. No survival data per RD. |
| Willemse 1992 | Included only 76 patients |
| Wils 1990 | Included only 88 patients |
| Wimberger 2007 | Multivariate analyses did not include residual disease and the study also included women with stage IIB and IIC disease. We attempted to contact the authors for further information but at time of submission of the review there had been no correspondence. |

| Study | Reason for exclusion |
|---|---|
| Yamamoto 2007 | Included 67 ''selected'' patients with rare histological subtype |
| Zang 1999 | Included only 71 patients and 31 of them received neoadjuvant chemotherapy |
| Zhang 2015 | < 100 patients with advanced disease in study |

CI: confidence interval; HIPEC: hyperthermic intraperitoneal chemotherapy; HR: hazard ratio; IDS: interval debulking surgery; NAC: neoadjuvant chemotherapy; OS: overall survival; PDS: Primary debulking surgery; RD: residual disease; SE: standard error

## DATA AND ANALYSES

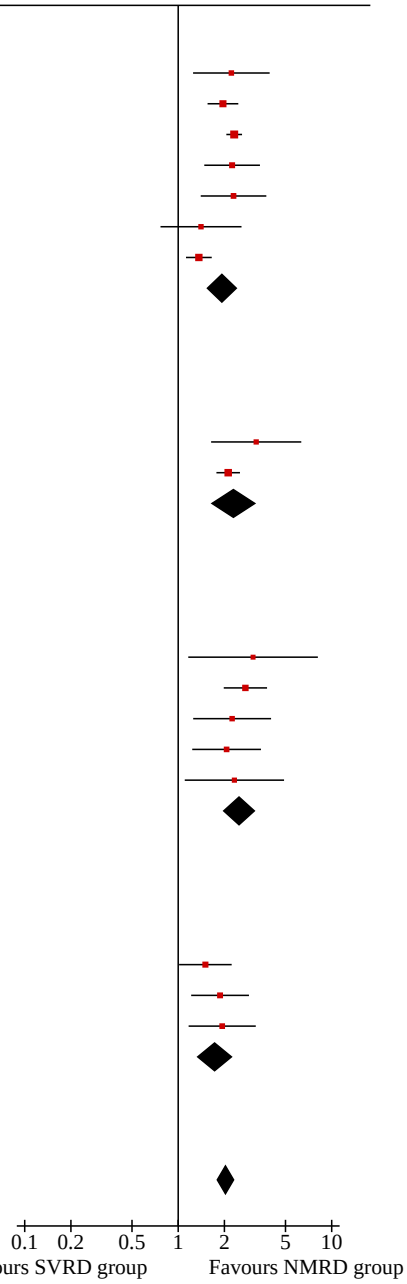
### Comparison 1. PDS: SVRD (< 1 cm) versus NMRD

| Outcome or subgroup title | No. of studies | No. of participants | Statistical method | Effect size |
|---|---|---|---|---|
| 1.1 Overall survival | 17 | | Hazard Ratio (IV, Random, 95% CI) | 2.03 [1.80, 2.29] |
| 1.1.1 Advanced stage (III/IV) | 7 | | Hazard Ratio (IV, Random, 95% CI) | 1.93 [1.55, 2.39] |
| 1.1.2 Stage III | 2 | | Hazard Ratio (IV, Random, 95% CI) | 2.29 [1.66, 3.15] |
| 1.1.3 Stage IIIC | 5 | | Hazard Ratio (IV, Random, 95% CI) | 2.49 [1.98, 3.13] |
| 1.1.4 Stage IV | 3 | | Hazard Ratio (IV, Random, 95% CI) | 1.73 [1.34, 2.22] |
| 1.2 Overall survival - sensitivity analysis using fixed-effect model | 17 | | Hazard Ratio (IV, Fixed, 95% CI) | 2.05 [1.91, 2.20] |
| 1.2.1 Advanced stage (III/IV) | 7 | | Hazard Ratio (IV, Fixed, 95% CI) | 2.01 [1.84, 2.19] |
| 1.2.2 Stage III | 2 | | Hazard Ratio (IV, Fixed, 95% CI) | 2.17 [1.83, 2.58] |
| 1.2.3 Stage IIIC | 5 | | Hazard Ratio (IV, Fixed, 95% CI) | 2.49 [1.98, 3.13] |
| 1.2.4 Stage IV | 3 | | Hazard Ratio (IV, Fixed, 95% CI) | 1.73 [1.34, 2.22] |
| 1.3 Overall survival - sensitivity analysis excluding Klar 2016 | 16 | | Hazard Ratio (IV, Random, 95% CI) | 1.99 [1.75, 2.27] |
| 1.3.1 Advanced stage (III/IV) | 6 | | Hazard Ratio (IV, Random, 95% CI) | 1.81 [1.46, 2.25] |
| 1.3.2 Stage III | 2 | | Hazard Ratio (IV, Random, 95% CI) | 2.29 [1.66, 3.15] |
| 1.3.3 Stage IIIC | 5 | | Hazard Ratio (IV, Random, 95% CI) | 2.49 [1.98, 3.13] |
| 1.3.4 Stage IV | 3 | | Hazard Ratio (IV, Random, 95% CI) | 1.73 [1.34, 2.22] |
| 1.4 Progression-free survival | 10 | | Hazard Ratio (IV, Random, 95% CI) | 1.88 [1.63, 2.16] |

| Outcome or subgroup title | No. of studies | No. of participants | Statistical method | Effect size |
|---|---|---|---|---|
| 1.4.1 Advanced stage (III/IV) | 5 | | Hazard Ratio (IV, Random, 95% CI) | 1.82 [1.43, 2.32] |
| 1.4.2 Stage III | 2 | | Hazard Ratio (IV, Random, 95% CI) | 2.21 [1.54, 3.18] |
| 1.4.3 Stage IIIC | 1 | | Hazard Ratio (IV, Random, 95% CI) | 2.03 [1.25, 3.31] |
| 1.4.4 Stage IV | 2 | | Hazard Ratio (IV, Random, 95% CI) | 1.68 [1.26, 2.24] |
| 1.5 Progression-free survival - sensitivity analysis using fixed-effect model | 10 | | Hazard Ratio (IV, Fixed, 95% CI) | 1.93 [1.80, 2.06] |
| 1.5.1 Advanced stage (III/IV) | 5 | | Hazard Ratio (IV, Fixed, 95% CI) | 1.92 [1.77, 2.08] |
| 1.5.2 Stage III | 2 | | Hazard Ratio (IV, Fixed, 95% CI) | 2.01 [1.76, 2.31] |
| 1.5.3 Stage IIIC | 1 | | Hazard Ratio (IV, Fixed, 95% CI) | 2.03 [1.25, 3.31] |
| 1.5.4 Stage IV | 2 | | Hazard Ratio (IV, Fixed, 95% CI) | 1.68 [1.26, 2.24] |
| 1.6 Progression-free survival - sensitivity analysis excluding Klar 2016 | 9 | | Hazard Ratio (IV, Random, 95% CI) | 1.83 [1.56, 2.13] |
| 1.6.1 Advanced stage (III/IV) | 4 | | Hazard Ratio (IV, Random, 95% CI) | 1.69 [1.33, 2.14] |
| 1.6.2 Stage III | 2 | | Hazard Ratio (IV, Random, 95% CI) | 2.21 [1.54, 3.18] |
| 1.6.3 Stage IIIC | 1 | | Hazard Ratio (IV, Random, 95% CI) | 2.03 [1.25, 3.31] |
| 1.6.4 Stage IV | 2 | | Hazard Ratio (IV, Random, 95% CI) | 1.68 [1.26, 2.24] |

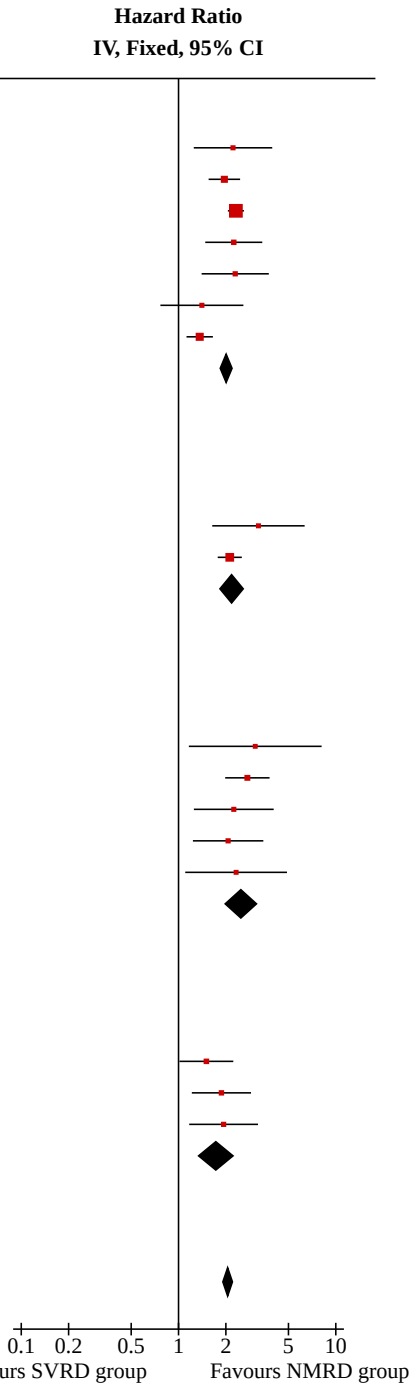
## Analysis 1.1. Comparison 1: PDS: SVRD (< 1 cm) versus NMRD, Outcome 1: Overall survival

| Study or Subgroup | log[Hazard Ratio] | SE | Weight | Hazard Ratio IV, Random, 95% CI |
|---|---|---|---|---|
| **1.1.1 Advanced stage (III/IV)** | | | | |
| Chang 2012a | 0.7975 | 0.293 | 3.5% | 2.22 [1.25 , 3.94] |
| Kahl 2017 | 0.671 | 0.1176 | 10.0% | 1.96 [1.55 , 2.46] |
| Klar 2016 | 0.84 | 0.06 | 13.6% | 2.32 [2.06 , 2.61] |
| Langstraat 2011 | 0.8087 | 0.213 | 5.5% | 2.24 [1.48 , 3.41] |
| Paik 2018 | 0.83 | 0.251 | 4.4% | 2.29 [1.40 , 3.75] |
| Tewari 2016 | 0.342 | 0.31 | 3.2% | 1.41 [0.77 , 2.58] |
| Tseng 2018 | 0.31 | 0.0985 | 11.2% | 1.36 [1.12 , 1.65] |
| **Subtotal (95% CI)** | | | **51.5%** | **1.93 [1.55 , 2.39]** |
| Heterogeneity: Tau² = 0.05; Chi² = 23.15, df = 6 (P = 0.0007); I² = 74% | | | | |
| Test for overall effect: Z = 5.94 (P < 0.00001) | | | | |
| **1.1.2 Stage III** | | | | |
| Cuylan 2018 | 1.17 | 0.345 | 2.7% | 3.22 [1.64 , 6.34] |
| Winter 2007 | 0.75 | 0.09 | 11.8% | 2.12 [1.77 , 2.53] |
| **Subtotal (95% CI)** | | | **14.4%** | **2.29 [1.66 , 3.15]** |
| Heterogeneity: Tau² = 0.02; Chi² = 1.39, df = 1 (P = 0.24); I² = 28% | | | | |
| Test for overall effect: Z = 5.07 (P < 0.00001) | | | | |
| **1.1.3 Stage IIIC** | | | | |
| Aletti 2006 | 1.123 | 0.4959 | 1.4% | 3.07 [1.16 , 8.13] |
| Bristow 2011 | 1.008 | 0.1657 | 7.4% | 2.74 [1.98 , 3.79] |
| Chang 2012b | 0.809 | 0.298 | 3.4% | 2.25 [1.25 , 4.03] |
| Chi 2006 | 0.726 | 0.263 | 4.1% | 2.07 [1.23 , 3.46] |
| Eisenkop 2003 | 0.843 | 0.38 | 2.3% | 2.32 [1.10 , 4.89] |
| **Subtotal (95% CI)** | | | **18.6%** | **2.49 [1.98 , 3.13]** |
| Heterogeneity: Tau² = 0.00; Chi² = 1.17, df = 4 (P = 0.88); I² = 0% | | | | |
| Test for overall effect: Z = 7.80 (P < 0.00001) | | | | |
| **1.1.4 Stage IV** | | | | |
| Ataseven 2016 | 0.408 | 0.201 | 5.9% | 1.50 [1.01 , 2.23] |
| Wimberger 2010 | 0.628 | 0.2211 | 5.3% | 1.87 [1.21 , 2.89] |
| Winter 2008 | 0.66 | 0.257 | 4.3% | 1.93 [1.17 , 3.20] |
| **Subtotal (95% CI)** | | | **15.4%** | **1.73 [1.34 , 2.22]** |
| Heterogeneity: Tau² = 0.00; Chi² = 0.81, df = 2 (P = 0.67); I² = 0% | | | | |
| Test for overall effect: Z = 4.24 (P < 0.0001) | | | | |
| **Total (95% CI)** | | | **100.0%** | **2.03 [1.80 , 2.29]** |
| Heterogeneity: Tau² = 0.02; Chi² = 31.75, df = 16 (P = 0.01); I² = 50% | | | | |
| Test for overall effect: Z = 11.39 (P < 0.00001) | | | | |
| Test for subgroup differences: Chi² = 5.30, df = 3 (P = 0.15), I² = 43.4% | | | | |



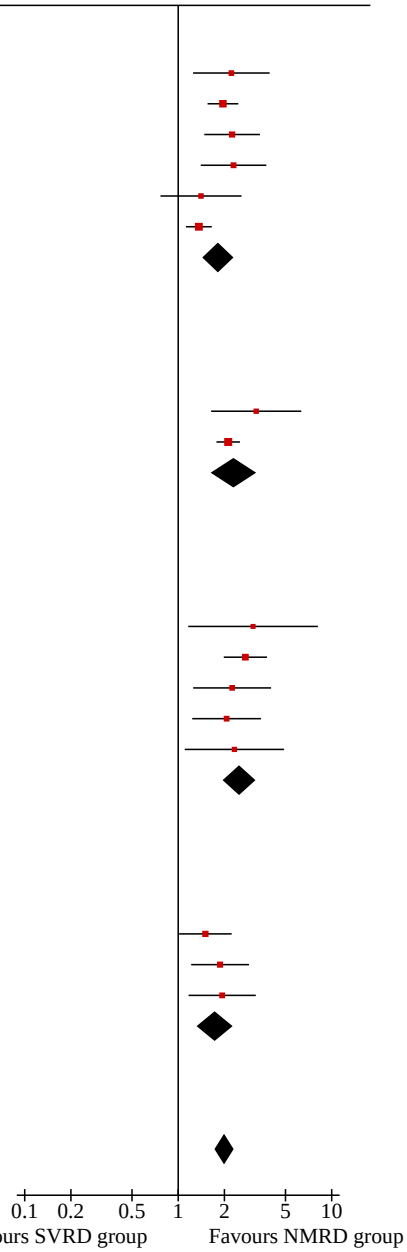
## Analysis 1.2. Comparison 1: PDS: SVRD (< 1 cm) versus NMRD, Outcome 2: Overall survival - sensitivity analysis using fixed-effect model

| Study or Subgroup | log[Hazard Ratio] | SE | Weight | Hazard Ratio IV, Fixed, 95% CI |
|---|---|---|---|---|
| **1.2.1 Advanced stage (III/IV)** | | | | |
| Chang 2012a | 0.7975 | 0.293 | 1.5% | 2.22 [1.25 , 3.94] |
| Kahl 2017 | 0.671 | 0.1176 | 9.3% | 1.96 [1.55 , 2.46] |
| Klar 2016 | 0.84 | 0.06 | 35.7% | 2.32 [2.06 , 2.61] |
| Langstraat 2011 | 0.8087 | 0.213 | 2.8% | 2.24 [1.48 , 3.41] |
| Paik 2018 | 0.83 | 0.251 | 2.0% | 2.29 [1.40 , 3.75] |
| Tewari 2016 | 0.342 | 0.31 | 1.3% | 1.41 [0.77 , 2.58] |
| Tseng 2018 | 0.31 | 0.0985 | 13.2% | 1.36 [1.12 , 1.65] |
| **Subtotal (95% CI)** | | | **65.9%** | **2.01 [1.84 , 2.19]** |
| Heterogeneity: Chi² = 23.15, df = 6 (P = 0.0007); I² = 74% | | | | |
| Test for overall effect: Z = 15.79 (P < 0.00001) | | | | |
| **1.2.2 Stage III** | | | | |
| Cuylan 2018 | 1.17 | 0.345 | 1.1% | 3.22 [1.64 , 6.34] |
| Winter 2007 | 0.75 | 0.09 | 15.9% | 2.12 [1.77 , 2.53] |
| **Subtotal (95% CI)** | | | **16.9%** | **2.17 [1.83 , 2.58]** |
| Heterogeneity: Chi² = 1.39, df = 1 (P = 0.24); I² = 28% | | | | |
| Test for overall effect: Z = 8.92 (P < 0.00001) | | | | |
| **1.2.3 Stage IIIC** | | | | |
| Aletti 2006 | 1.123 | 0.4959 | 0.5% | 3.07 [1.16 , 8.13] |
| Bristow 2011 | 1.008 | 0.1657 | 4.7% | 2.74 [1.98 , 3.79] |
| Chang 2012b | 0.809 | 0.298 | 1.4% | 2.25 [1.25 , 4.03] |
| Chi 2006 | 0.726 | 0.263 | 1.9% | 2.07 [1.23 , 3.46] |
| Eisenkop 2003 | 0.843 | 0.38 | 0.9% | 2.32 [1.10 , 4.89] |
| **Subtotal (95% CI)** | | | **9.4%** | **2.49 [1.98 , 3.13]** |
| Heterogeneity: Chi² = 1.17, df = 4 (P = 0.88); I² = 0% | | | | |
| Test for overall effect: Z = 7.80 (P < 0.00001) | | | | |
| **1.2.4 Stage IV** | | | | |
| Ataseven 2016 | 0.408 | 0.201 | 3.2% | 1.50 [1.01 , 2.23] |
| Wimberger 2010 | 0.628 | 0.2211 | 2.6% | 1.87 [1.21 , 2.89] |
| Winter 2008 | 0.66 | 0.257 | 1.9% | 1.93 [1.17 , 3.20] |
| **Subtotal (95% CI)** | | | **7.8%** | **1.73 [1.34 , 2.22]** |
| Heterogeneity: Chi² = 0.81, df = 2 (P = 0.67); I² = 0% | | | | |
| Test for overall effect: Z = 4.24 (P < 0.0001) | | | | |
| **Total (95% CI)** | | | **100.0%** | **2.05 [1.91 , 2.20]** |
| Heterogeneity: Chi² = 31.75, df = 16 (P = 0.01); I² = 50% | | | | |
| Test for overall effect: Z = 20.06 (P < 0.00001) | | | | |
| Test for subgroup differences: Chi² = 5.23, df = 3 (P = 0.16), I² = 42.7% | | | | |



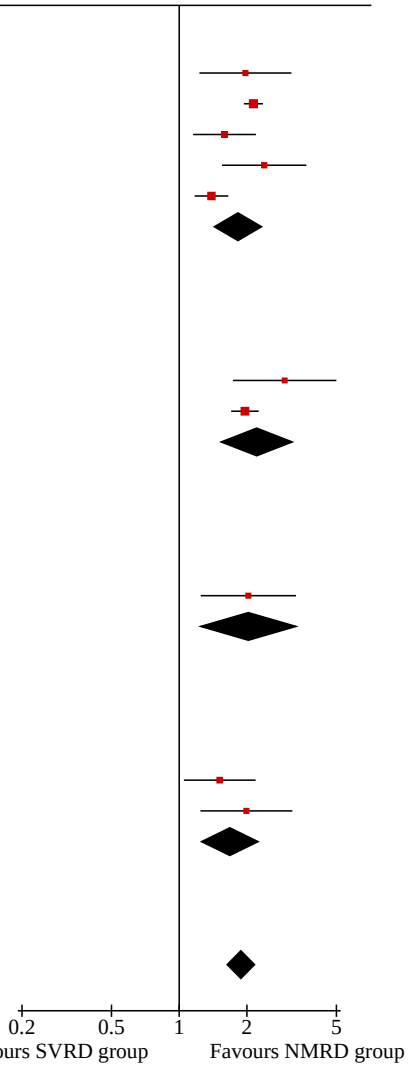
## Analysis 1.3. Comparison 1: PDS: SVRD (< 1 cm) versus NMRD, Outcome 3: Overall survival - sensitivity analysis excluding Klar 2016

| Study or Subgroup | log[Hazard Ratio] | SE | Weight | Hazard Ratio IV, Random, 95% CI |
|---|---|---|---|---|
| **1.3.1 Advanced stage (III/IV)** | | | | |
| Chang 2012a | 0.7975 | 0.293 | 4.0% | 2.22 [1.25 , 3.94] |
| Kahl 2017 | 0.671 | 0.1176 | 11.6% | 1.96 [1.55 , 2.46] |
| Langstraat 2011 | 0.8087 | 0.213 | 6.4% | 2.24 [1.48 , 3.41] |
| Paik 2018 | 0.83 | 0.251 | 5.1% | 2.29 [1.40 , 3.75] |
| Tewari 2016 | 0.342 | 0.31 | 3.7% | 1.41 [0.77 , 2.58] |
| Tseng 2018 | 0.31 | 0.0985 | 13.1% | 1.36 [1.12 , 1.65] |
| **Subtotal (95% CI)** | | | **43.9%** | **1.81 [1.46 , 2.25]** |
| Heterogeneity: Tau² = 0.04; Chi² = 10.77, df = 5 (P = 0.06); I² = 54% | | | | |
| Test for overall effect: Z = 5.36 (P < 0.00001) | | | | |
| **1.3.2 Stage III** | | | | |
| Cuylan 2018 | 1.17 | 0.345 | 3.1% | 3.22 [1.64 , 6.34] |
| Winter 2007 | 0.75 | 0.09 | 13.7% | 2.12 [1.77 , 2.53] |
| **Subtotal (95% CI)** | | | **16.8%** | **2.29 [1.66 , 3.15]** |
| Heterogeneity: Tau² = 0.02; Chi² = 1.39, df = 1 (P = 0.24); I² = 28% | | | | |
| Test for overall effect: Z = 5.07 (P < 0.00001) | | | | |
| **1.3.3 Stage IIIC** | | | | |
| Aletti 2006 | 1.123 | 0.4959 | 1.6% | 3.07 [1.16 , 8.13] |
| Bristow 2011 | 1.008 | 0.1657 | 8.6% | 2.74 [1.98 , 3.79] |
| Chang 2012b | 0.809 | 0.298 | 3.9% | 2.25 [1.25 , 4.03] |
| Chi 2006 | 0.726 | 0.263 | 4.7% | 2.07 [1.23 , 3.46] |
| Eisenkop 2003 | 0.843 | 0.38 | 2.6% | 2.32 [1.10 , 4.89] |
| **Subtotal (95% CI)** | | | **21.5%** | **2.49 [1.98 , 3.13]** |
| Heterogeneity: Tau² = 0.00; Chi² = 1.17, df = 4 (P = 0.88); I² = 0% | | | | |
| Test for overall effect: Z = 7.80 (P < 0.00001) | | | | |
| **1.3.4 Stage IV** | | | | |
| Ataseven 2016 | 0.408 | 0.201 | 6.9% | 1.50 [1.01 , 2.23] |
| Wimberger 2010 | 0.628 | 0.2211 | 6.1% | 1.87 [1.21 , 2.89] |
| Winter 2008 | 0.66 | 0.257 | 4.9% | 1.93 [1.17 , 3.20] |
| **Subtotal (95% CI)** | | | **17.8%** | **1.73 [1.34 , 2.22]** |
| Heterogeneity: Tau² = 0.00; Chi² = 0.81, df = 2 (P = 0.67); I² = 0% | | | | |
| Test for overall effect: Z = 4.24 (P < 0.0001) | | | | |
| **Total (95% CI)** | | | **100.0%** | **1.99 [1.75 , 2.27]** |
| Heterogeneity: Tau² = 0.02; Chi² = 25.43, df = 15 (P = 0.04); I² = 41% | | | | |
| Test for overall effect: Z = 10.33 (P < 0.00001) | | | | |
| Test for subgroup differences: Chi² = 6.22, df = 3 (P = 0.10), I² = 51.8% | | | | |



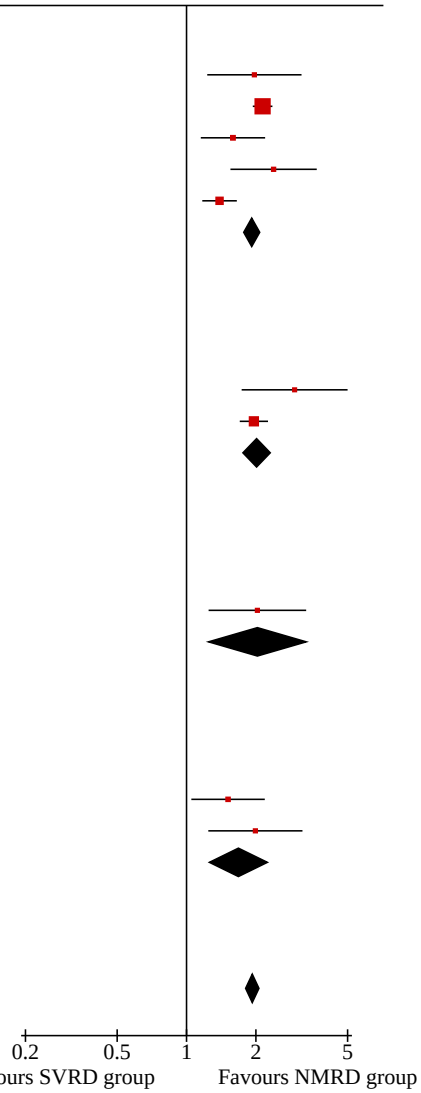
## Analysis 1.4. Comparison 1: PDS: SVRD (< 1 cm) versus NMRD, Outcome 4: Progression-free survival

| Study or Subgroup | log[Hazard Ratio] | SE | Weight | Hazard Ratio IV, Random, 95% CI |
|---|---|---|---|---|
| **1.4.1 Advanced stage (III/IV)** | | | | |
| Chang 2012a | 0.678 | 0.2403 | 6.2% | 1.97 [1.23 , 3.15] |
| Klar 2016 | 0.76 | 0.05 | 18.6% | 2.14 [1.94 , 2.36] |
| Paik 2018 | 0.464 | 0.164 | 9.9% | 1.59 [1.15 , 2.19] |
| Shim 2016 | 0.87 | 0.22 | 7.0% | 2.39 [1.55 , 3.67] |
| Tseng 2018 | 0.33 | 0.088 | 15.6% | 1.39 [1.17 , 1.65] |
| **Subtotal (95% CI)** | | | **57.2%** | **1.82 [1.43 , 2.32]** |
| Heterogeneity: Tau² = 0.05; Chi² = 20.36, df = 4 (P = 0.0004); I² = 80% | | | | |
| Test for overall effect: Z = 4.88 (P < 0.00001) | | | | |
| **1.4.2 Stage III** | | | | |
| Cuylan 2018 | 1.08 | 0.27 | 5.2% | 2.94 [1.73 , 5.00] |
| Winter 2007 | 0.673 | 0.072 | 16.9% | 1.96 [1.70 , 2.26] |
| **Subtotal (95% CI)** | | | **22.2%** | **2.21 [1.54 , 3.18]** |
| Heterogeneity: Tau² = 0.04; Chi² = 2.12, df = 1 (P = 0.15); I² = 53% | | | | |
| Test for overall effect: Z = 4.27 (P < 0.0001) | | | | |
| **1.4.3 Stage IIIC** | | | | |
| Chang 2012b | 0.708 | 0.2487 | 5.9% | 2.03 [1.25 , 3.31] |
| **Subtotal (95% CI)** | | | **5.9%** | **2.03 [1.25 , 3.31]** |
| Heterogeneity: Not applicable | | | | |
| Test for overall effect: Z = 2.85 (P = 0.004) | | | | |
| **1.4.4 Stage IV** | | | | |
| Wimberger 2010 | 0.415 | 0.1872 | 8.5% | 1.51 [1.05 , 2.19] |
| Winter 2008 | 0.688 | 0.24 | 6.2% | 1.99 [1.24 , 3.18] |
| **Subtotal (95% CI)** | | | **14.7%** | **1.68 [1.26 , 2.24]** |
| Heterogeneity: Tau² = 0.00; Chi² = 0.80, df = 1 (P = 0.37); I² = 0% | | | | |
| Test for overall effect: Z = 3.51 (P = 0.0004) | | | | |
| **Total (95% CI)** | | | **100.0%** | **1.88 [1.63 , 2.16]** |
| Heterogeneity: Tau² = 0.03; Chi² = 24.62, df = 9 (P = 0.003); I² = 63% | | | | |
| Test for overall effect: Z = 8.82 (P < 0.00001) | | | | |
| Test for subgroup differences: Chi² = 1.49, df = 3 (P = 0.68), I² = 0% | | | | |

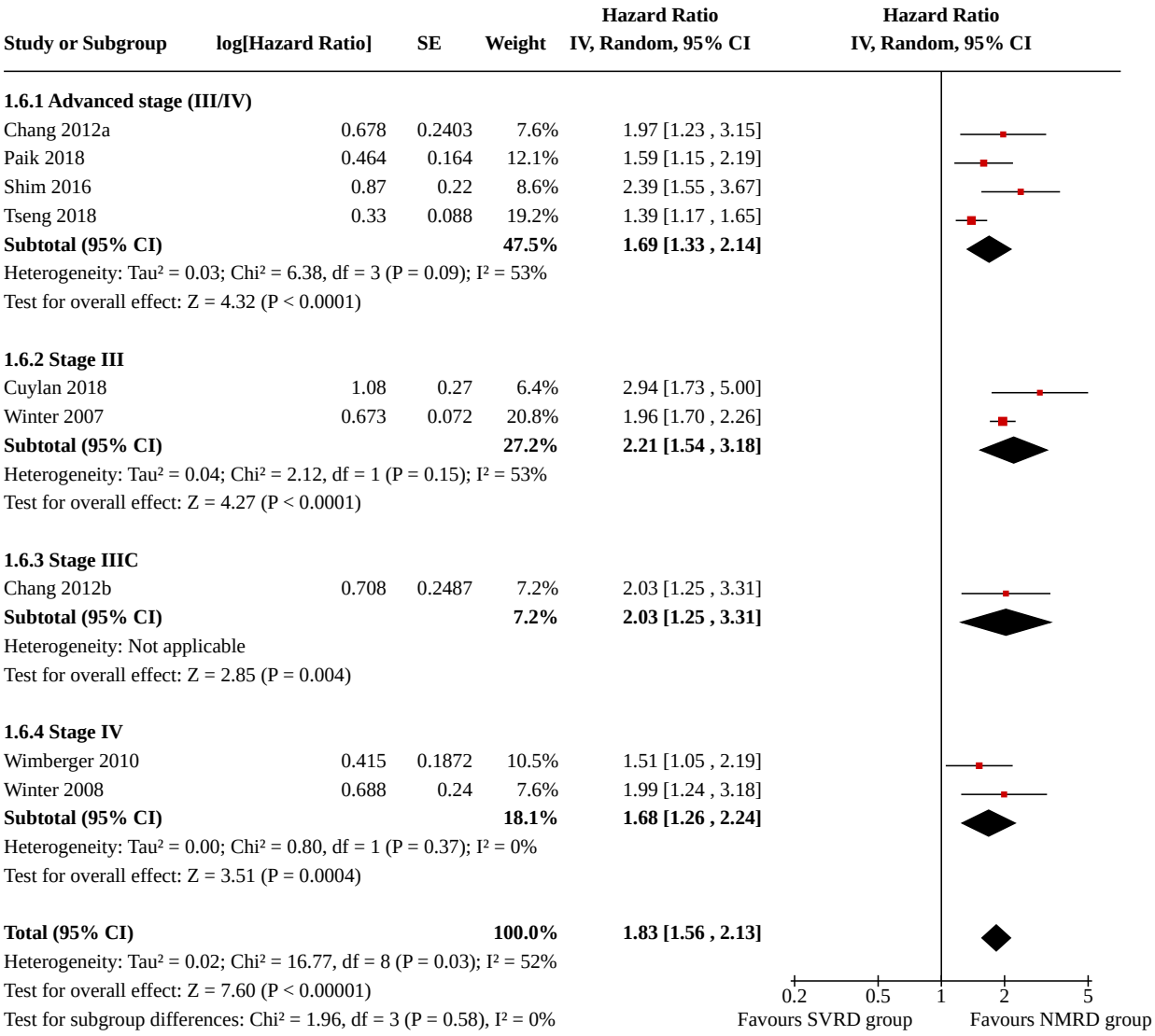


## Analysis 1.5. Comparison 1: PDS: SVRD (< 1 cm) versus NMRD, Outcome 5: Progression-free survival - sensitivity analysis using fixed-effect model

| Study or Subgroup | log[Hazard Ratio] | SE | Weight | Hazard Ratio IV, Fixed, 95% CI |
|---|---|---|---|---|
| **1.5.1 Advanced stage (III/IV)** | | | | |
| Chang 2012a | 0.678 | 0.2403 | 2.0% | 1.97 [1.23 , 3.15] |
| Klar 2016 | 0.76 | 0.05 | 45.8% | 2.14 [1.94 , 2.36] |
| Paik 2018 | 0.464 | 0.164 | 4.3% | 1.59 [1.15 , 2.19] |
| Shim 2016 | 0.87 | 0.22 | 2.4% | 2.39 [1.55 , 3.67] |
| Tseng 2018 | 0.33 | 0.088 | 14.8% | 1.39 [1.17 , 1.65] |
| **Subtotal (95% CI)** | | | **69.2%** | **1.92 [1.77 , 2.08]** |
| Heterogeneity: Chi² = 20.36, df = 4 (P = 0.0004); I² = 80% | | | | |
| Test for overall effect: Z = 16.01 (P < 0.00001) | | | | |
| **1.5.2 Stage III** | | | | |
| Cuylan 2018 | 1.08 | 0.27 | 1.6% | 2.94 [1.73 , 5.00] |
| Winter 2007 | 0.673 | 0.072 | 22.1% | 1.96 [1.70 , 2.26] |
| **Subtotal (95% CI)** | | | **23.7%** | **2.01 [1.76 , 2.31]** |
| Heterogeneity: Chi² = 2.12, df = 1 (P = 0.15); I² = 53% | | | | |
| Test for overall effect: Z = 10.06 (P < 0.00001) | | | | |
| **1.5.3 Stage IIIC** | | | | |
| Chang 2012b | 0.708 | 0.2487 | 1.9% | 2.03 [1.25 , 3.31] |
| **Subtotal (95% CI)** | | | **1.9%** | **2.03 [1.25 , 3.31]** |
| Heterogeneity: Not applicable | | | | |
| Test for overall effect: Z = 2.85 (P = 0.004) | | | | |
| **1.5.4 Stage IV** | | | | |
| Wimberger 2010 | 0.415 | 0.1872 | 3.3% | 1.51 [1.05 , 2.19] |
| Winter 2008 | 0.688 | 0.24 | 2.0% | 1.99 [1.24 , 3.18] |
| **Subtotal (95% CI)** | | | **5.3%** | **1.68 [1.26 , 2.24]** |
| Heterogeneity: Chi² = 0.80, df = 1 (P = 0.37); I² = 0% | | | | |
| Test for overall effect: Z = 3.51 (P = 0.0004) | | | | |
| **Total (95% CI)** | | | **100.0%** | **1.93 [1.80 , 2.06]** |
| Heterogeneity: Chi² = 24.62, df = 9 (P = 0.003); I² = 63% | | | | |
| Test for overall effect: Z = 19.41 (P < 0.00001) | | | | |
| Test for subgroup differences: Chi² = 1.33, df = 3 (P = 0.72), I² = 0% | | | | |



**Analysis 1.6. Comparison 1: PDS: SVRD (< 1 cm) versus NMRD, Outcome 6: Progression-free survival - sensitivity analysis excluding Klar 2016**



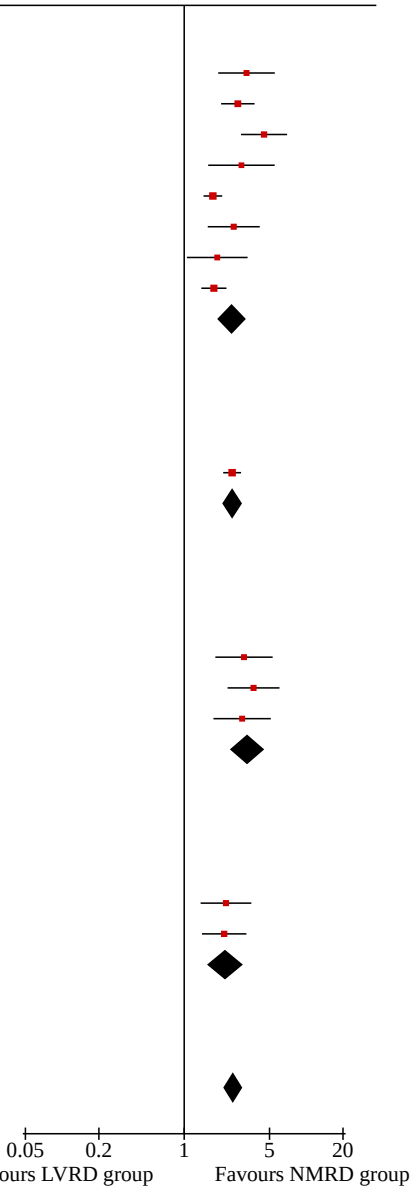

## Comparison 2. PDS: LVRD (> 1 cm) versus NMRD

| Outcome or subgroup title | No. of studies | No. of participants | Statistical method | Effect size |
|---|---|---|---|---|
| 2.1 Overall survival | 14 | | Hazard Ratio (IV, Random, 95% CI) | 2.50 [2.13, 2.94] |
| 2.1.1 Advanced stage (III/IV) | 8 | | Hazard Ratio (IV, Random, 95% CI) | 2.44 [1.90, 3.14] |
| 2.1.2 Stage III | 1 | | Hazard Ratio (IV, Random, 95% CI) | 2.47 [2.09, 2.92] |
| 2.1.3 Stage IIIC | 3 | | Hazard Ratio (IV, Random, 95% CI) | 3.27 [2.42, 4.42] |
| 2.1.4 Stage IV | 2 | | Hazard Ratio (IV, Random, 95% CI) | 2.16 [1.57, 2.96] |

| Outcome or subgroup title | No. of studies | No. of participants | Statistical method | Effect size |
|---|---|---|---|---|
| 2.2 Overall survival - sensitivity analysis using fixed effects model | 14 | | Hazard Ratio (IV, Fixed, 95% CI) | 2.27 [2.09, 2.48] |
| 2.2.1 Advanced stage (III/IV) | 8 | | Hazard Ratio (IV, Fixed, 95% CI) | 2.10 [1.87, 2.35] |
| 2.2.2 Stage III | 1 | | Hazard Ratio (IV, Fixed, 95% CI) | 2.47 [2.09, 2.92] |
| 2.2.3 Stage IIIC | 3 | | Hazard Ratio (IV, Fixed, 95% CI) | 3.27 [2.42, 4.42] |
| 2.2.4 Stage IV | 2 | | Hazard Ratio (IV, Fixed, 95% CI) | 2.16 [1.57, 2.96] |
| 2.3 Overall survival - sensitivity analysis excluding Melamed 2017b and Winter 2007 | 12 | | Hazard Ratio (IV, Random, 95% CI) | 2.65 [2.20, 3.19] |
| 2.3.1 Advanced stage (III/IV) | 7 | | Hazard Ratio (IV, Random, 95% CI) | 2.63 [1.99, 3.47] |
| 2.3.2 Stage IIIC | 3 | | Hazard Ratio (IV, Random, 95% CI) | 3.27 [2.42, 4.42] |
| 2.3.3 Stage IV | 2 | | Hazard Ratio (IV, Random, 95% CI) | 2.16 [1.57, 2.96] |
| 2.4 Progression-free survival | 6 | | Hazard Ratio (IV, Random, 95% CI) | 2.10 [1.84, 2.40] |
| 2.4.1 Advanced stage (III/IV) | 3 | | Hazard Ratio (IV, Random, 95% CI) | 1.92 [1.62, 2.27] |
| 2.4.2 Stage III | 1 | | Hazard Ratio (IV, Random, 95% CI) | 2.36 [2.04, 2.73] |
| 2.4.3 Stage IIIC | 1 | | Hazard Ratio (IV, Random, 95% CI) | 2.56 [1.54, 4.26] |
| 2.4.4 Stage IV | 1 | | Hazard Ratio (IV, Random, 95% CI) | 1.82 [1.28, 2.59] |

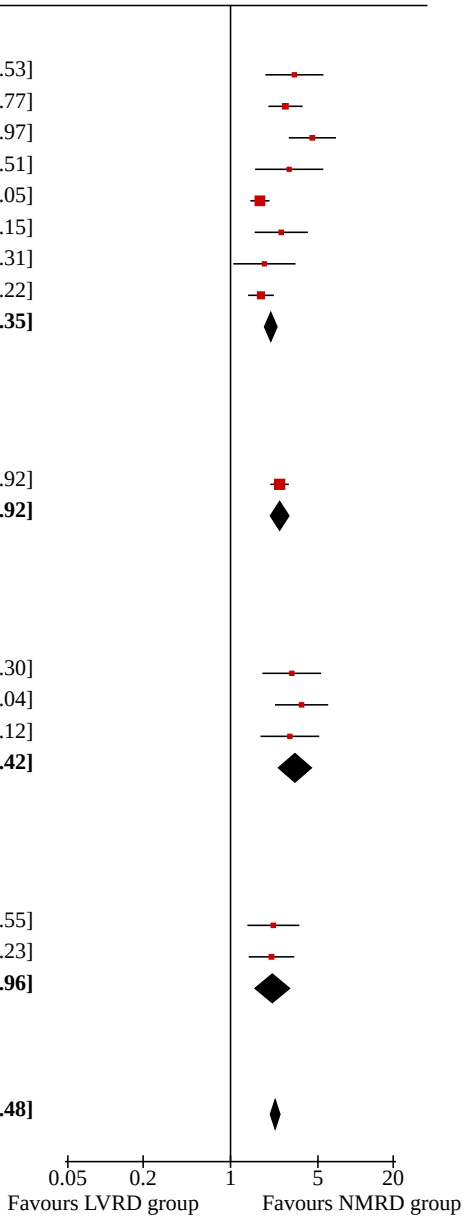
## Analysis 2.1. Comparison 2: PDS: LVRD (> 1 cm) versus NMRD, Outcome 1: Overall survival

| Study or Subgroup | log[Hazard Ratio] | SE | Weight | Hazard Ratio IV, Random, 95% CI |
|---|---|---|---|---|
| **2.1.1 Advanced stage (III/IV)** | | | | |
| Chang 2012a | 1.1756 | 0.2723 | 5.4% | 3.24 [1.90 , 5.53] |
| Kahl 2017 | 1.012 | 0.161 | 8.9% | 2.75 [2.01 , 3.77] |
| Langstraat 2011 | 1.5063 | 0.2218 | 6.8% | 4.51 [2.92 , 6.97] |
| Melamed 2017a | 1.08 | 0.32 | 4.4% | 2.94 [1.57 , 5.51] |
| Melamed 2017b | 0.541 | 0.09 | 11.5% | 1.72 [1.44 , 2.05] |
| Paik 2018 | 0.9343 | 0.25 | 6.0% | 2.55 [1.56 , 4.15] |
| Tewari 2016 | 0.6235 | 0.292 | 5.0% | 1.87 [1.05 , 3.31] |
| Tseng 2018 | 0.56 | 0.121 | 10.4% | 1.75 [1.38 , 2.22] |
| **Subtotal (95% CI)** | | | **58.4%** | **2.44 [1.90 , 3.14]** |
| Heterogeneity: Tau² = 0.09; Chi² = 26.34, df = 7 (P = 0.0004); I² = 73% | | | | |
| Test for overall effect: Z = 7.00 (P < 0.00001) | | | | |
| **2.1.2 Stage III** | | | | |
| Winter 2007 | 0.904 | 0.086 | 11.6% | 2.47 [2.09 , 2.92] |
| **Subtotal (95% CI)** | | | **11.6%** | **2.47 [2.09 , 2.92]** |
| Heterogeneity: Not applicable | | | | |
| Test for overall effect: Z = 10.51 (P < 0.00001) | | | | |
| **2.1.3 Stage IIIC** | | | | |
| Chang 2012b | 1.1282 | 0.2757 | 5.4% | 3.09 [1.80 , 5.30] |
| Chi 2006 | 1.308 | 0.25 | 6.0% | 3.70 [2.27 , 6.04] |
| Eisenkop 2003 | 1.093 | 0.276 | 5.4% | 2.98 [1.74 , 5.12] |
| **Subtotal (95% CI)** | | | **16.7%** | **3.27 [2.42 , 4.42]** |
| Heterogeneity: Tau² = 0.00; Chi² = 0.40, df = 2 (P = 0.82); I² = 0% | | | | |
| Test for overall effect: Z = 7.71 (P < 0.00001) | | | | |
| **2.1.4 Stage IV** | | | | |
| Ataseven 2016 | 0.788 | 0.244 | 6.2% | 2.20 [1.36 , 3.55] |
| Wimberger 2010 | 0.754 | 0.214 | 7.0% | 2.13 [1.40 , 3.23] |
| **Subtotal (95% CI)** | | | **13.2%** | **2.16 [1.57 , 2.96]** |
| Heterogeneity: Tau² = 0.00; Chi² = 0.01, df = 1 (P = 0.92); I² = 0% | | | | |
| Test for overall effect: Z = 4.78 (P < 0.00001) | | | | |
| **Total (95% CI)** | | | **100.0%** | **2.50 [2.13 , 2.94]** |
| Heterogeneity: Tau² = 0.05; Chi² = 35.35, df = 13 (P = 0.0007); I² = 63% | | | | |
| Test for overall effect: Z = 11.09 (P < 0.00001) | | | | |
| Test for subgroup differences: Chi² = 3.93, df = 3 (P = 0.27), I² = 23.7% | | | | |



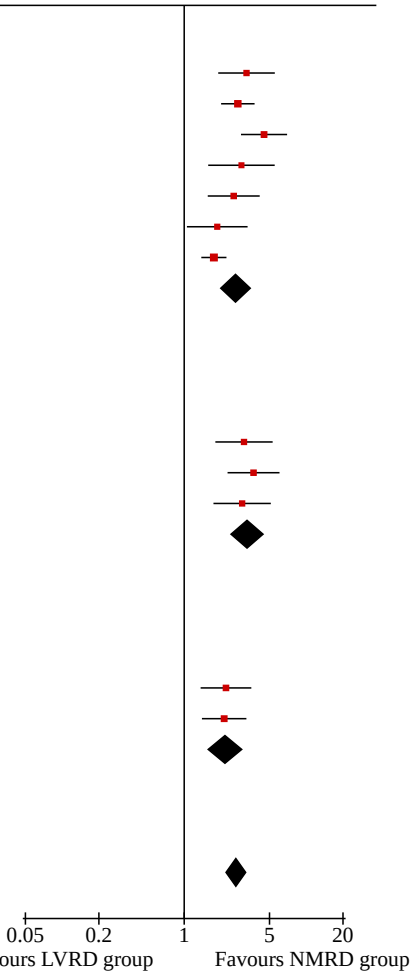
## Analysis 2.2. Comparison 2: PDS: LVRD (> 1 cm) versus NMRD, Outcome 2: Overall survival - sensitivity analysis using fixed effects model

| Study or Subgroup | log[Hazard Ratio] | SE | Weight | Hazard Ratio IV, Fixed, 95% CI |
|---|---|---|---|---|
| **2.2.1 Advanced stage (III/IV)** | | | | |
| Chang 2012a | 1.1756 | 0.2723 | 2.6% | 3.24 [1.90 , 5.53] |
| Kahl 2017 | 1.012 | 0.161 | 7.5% | 2.75 [2.01 , 3.77] |
| Langstraat 2011 | 1.5063 | 0.2218 | 3.9% | 4.51 [2.92 , 6.97] |
| Melamed 2017a | 1.08 | 0.32 | 1.9% | 2.94 [1.57 , 5.51] |
| Melamed 2017b | 0.541 | 0.09 | 23.8% | 1.72 [1.44 , 2.05] |
| Paik 2018 | 0.9343 | 0.25 | 3.1% | 2.55 [1.56 , 4.15] |
| Tewari 2016 | 0.6235 | 0.292 | 2.3% | 1.87 [1.05 , 3.31] |
| Tseng 2018 | 0.56 | 0.121 | 13.2% | 1.75 [1.38 , 2.22] |
| **Subtotal (95% CI)** | | | **58.3%** | **2.10 [1.87 , 2.35]** |
| Heterogeneity: Chi² = 26.34, df = 7 (P = 0.0004); I² = 73% | | | | |
| Test for overall effect: Z = 12.86 (P < 0.00001) | | | | |
| **2.2.2 Stage III** | | | | |
| Winter 2007 | 0.904 | 0.086 | 26.1% | 2.47 [2.09 , 2.92] |
| **Subtotal (95% CI)** | | | **26.1%** | **2.47 [2.09 , 2.92]** |
| Heterogeneity: Not applicable | | | | |
| Test for overall effect: Z = 10.51 (P < 0.00001) | | | | |
| **2.2.3 Stage IIIC** | | | | |
| Chang 2012b | 1.1282 | 0.2757 | 2.5% | 3.09 [1.80 , 5.30] |
| Chi 2006 | 1.308 | 0.25 | 3.1% | 3.70 [2.27 , 6.04] |
| Eisenkop 2003 | 1.093 | 0.276 | 2.5% | 2.98 [1.74 , 5.12] |
| **Subtotal (95% CI)** | | | **8.2%** | **3.27 [2.42 , 4.42]** |
| Heterogeneity: Chi² = 0.40, df = 2 (P = 0.82); I² = 0% | | | | |
| Test for overall effect: Z = 7.71 (P < 0.00001) | | | | |
| **2.2.4 Stage IV** | | | | |
| Ataseven 2016 | 0.788 | 0.244 | 3.2% | 2.20 [1.36 , 3.55] |
| Wimberger 2010 | 0.754 | 0.214 | 4.2% | 2.13 [1.40 , 3.23] |
| **Subtotal (95% CI)** | | | **7.5%** | **2.16 [1.57 , 2.96]** |
| Heterogeneity: Chi² = 0.01, df = 1 (P = 0.92); I² = 0% | | | | |
| Test for overall effect: Z = 4.78 (P < 0.00001) | | | | |
| **Total (95% CI)** | | | **100.0%** | **2.27 [2.09 , 2.48]** |
| Heterogeneity: Chi² = 35.35, df = 13 (P = 0.0007); I² = 63% | | | | |
| Test for overall effect: Z = 18.70 (P < 0.00001) | | | | |
| Test for subgroup differences: Chi² = 8.60, df = 3 (P = 0.04), I² = 65.1% | | | | |

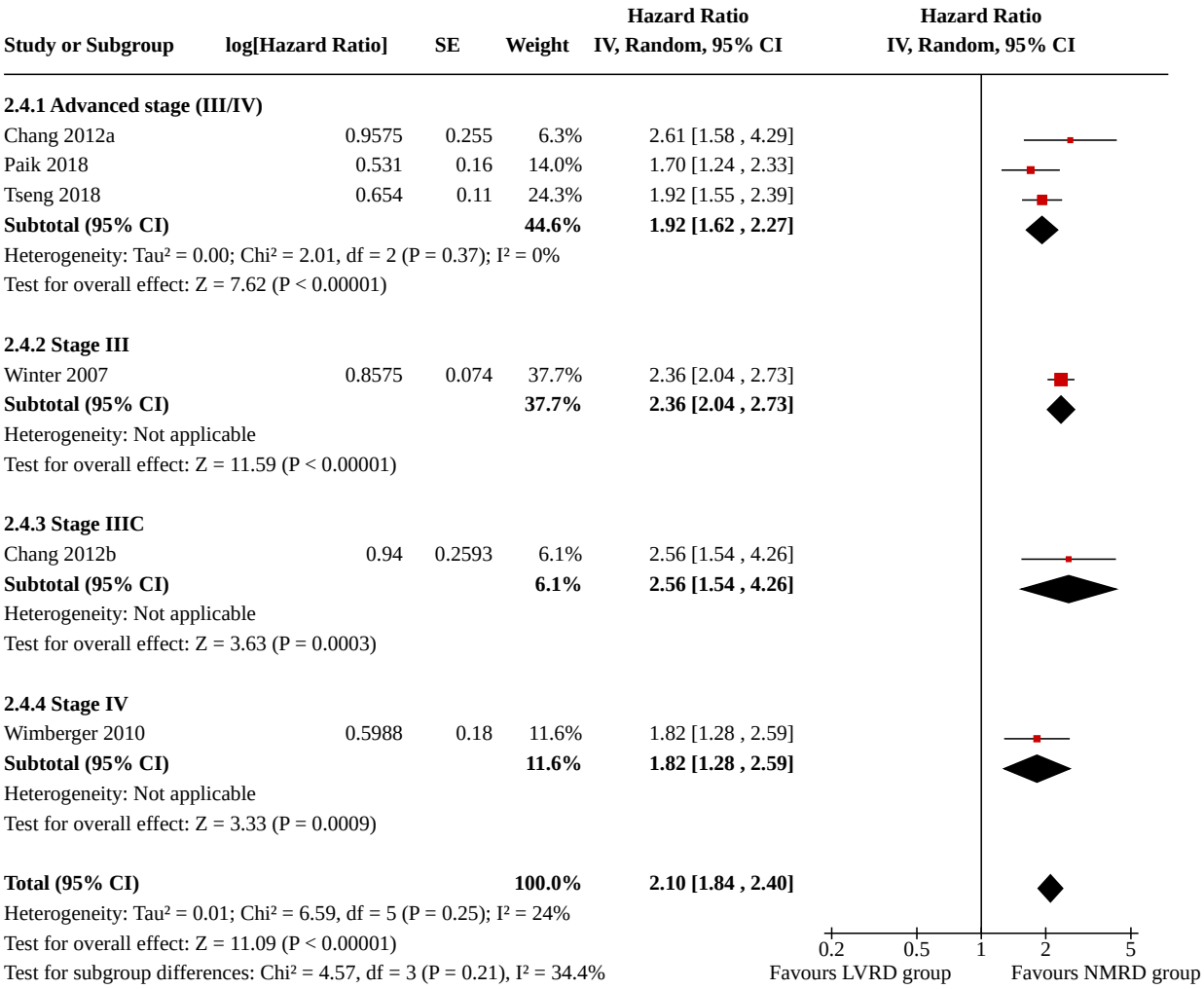


## Analysis 2.3. Comparison 2: PDS: LVRD (> 1 cm) versus NMRD, Outcome 3: Overall survival - sensitivity analysis excluding Melamed 2017b and Winter 2007

| Study or Subgroup | log[Hazard Ratio] | SE | Weight | Hazard Ratio IV, Random, 95% CI |
|---|---|---|---|---|
| **2.3.1 Advanced stage (III/IV)** | | | | |
| Chang 2012a | 1.1756 | 0.2723 | 7.1% | 3.24 [1.90 , 5.53] |
| Kahl 2017 | 1.012 | 0.161 | 11.5% | 2.75 [2.01 , 3.77] |
| Langstraat 2011 | 1.5063 | 0.2218 | 8.8% | 4.51 [2.92 , 6.97] |
| Melamed 2017a | 1.08 | 0.32 | 5.8% | 2.94 [1.57 , 5.51] |
| Paik 2018 | 0.9343 | 0.25 | 7.8% | 2.55 [1.56 , 4.15] |
| Tewari 2016 | 0.6235 | 0.292 | 6.5% | 1.87 [1.05 , 3.31] |
| Tseng 2018 | 0.56 | 0.121 | 13.5% | 1.75 [1.38 , 2.22] |
| **Subtotal (95% CI)** | | | **61.0%** | **2.63 [1.99 , 3.47]** |
| Heterogeneity: Tau² = 0.09; Chi² = 18.03, df = 6 (P = 0.006); I² = 67% | | | | |
| Test for overall effect: Z = 6.83 (P < 0.00001) | | | | |
| **2.3.2 Stage IIIC** | | | | |
| Chang 2012b | 1.1282 | 0.2757 | 7.0% | 3.09 [1.80 , 5.30] |
| Chi 2006 | 1.308 | 0.25 | 7.8% | 3.70 [2.27 , 6.04] |
| Eisenkop 2003 | 1.093 | 0.276 | 7.0% | 2.98 [1.74 , 5.12] |
| **Subtotal (95% CI)** | | | **21.8%** | **3.27 [2.42 , 4.42]** |
| Heterogeneity: Tau² = 0.00; Chi² = 0.40, df = 2 (P = 0.82); I² = 0% | | | | |
| Test for overall effect: Z = 7.71 (P < 0.00001) | | | | |
| **2.3.3 Stage IV** | | | | |
| Ataseven 2016 | 0.788 | 0.244 | 8.0% | 2.20 [1.36 , 3.55] |
| Wimberger 2010 | 0.754 | 0.214 | 9.2% | 2.13 [1.40 , 3.23] |
| **Subtotal (95% CI)** | | | **17.2%** | **2.16 [1.57 , 2.96]** |
| Heterogeneity: Tau² = 0.00; Chi² = 0.01, df = 1 (P = 0.92); I² = 0% | | | | |
| Test for overall effect: Z = 4.78 (P < 0.00001) | | | | |
| **Total (95% CI)** | | | **100.0%** | **2.65 [2.20 , 3.19]** |
| Heterogeneity: Tau² = 0.05; Chi² = 22.58, df = 11 (P = 0.02); I² = 51% | | | | |
| Test for overall effect: Z = 10.33 (P < 0.00001) | | | | |
| Test for subgroup differences: Chi² = 3.51, df = 2 (P = 0.17), I² = 43.1% | | | | |



**Analysis 2.4. Comparison 2: PDS: LVRD (> 1 cm) versus NMRD, Outcome 4: Progression-free survival**



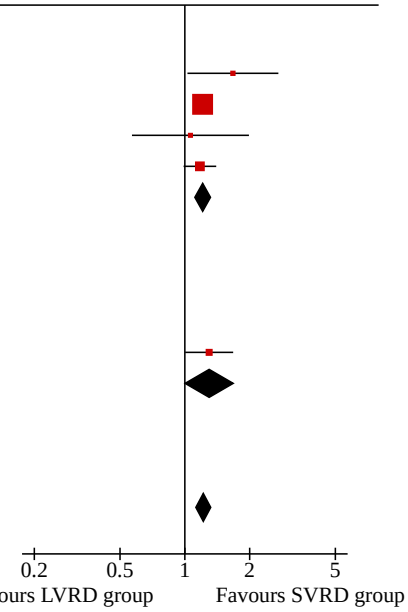

## Comparison 3. PDS: LVRD (> 1 cm) versus SVRD (< 1 cm)

| Outcome or subgroup title | No. of studies | No. of participants | Statistical method | Effect size |
|---|---|---|---|---|
| 3.1 Overall survival | 5 | | Hazard Ratio (IV, Random, 95% CI) | 1.22 [1.13, 1.32] |
| 3.1.1 Advanced stage (III/IV) | 4 | | Hazard Ratio (IV, Random, 95% CI) | 1.21 [1.11, 1.32] |
| 3.1.2 Stage IV | 1 | | Hazard Ratio (IV, Random, 95% CI) | 1.30 [1.00, 1.68] |
| 3.2 Overall survival sensitivity analysis excluding Klar 2016 | 4 | | Hazard Ratio (IV, Random, 95% CI) | 1.23 [1.08, 1.41] |
| 3.2.1 Advanced stage (III/IV) | 3 | | Hazard Ratio (IV, Random, 95% CI) | 1.21 [1.03, 1.42] |

| Outcome or subgroup title | No. of studies | No. of participants | Statistical method | Effect size |
|---|---|---|---|---|
| 3.2.2 Stage IV | 1 | | Hazard Ratio (IV, Random, 95% CI) | 1.30 [1.00, 1.68] |
| 3.3 Overall survival sensitivity analysis excluding 0 cm | 3 | | Hazard Ratio (IV, Random, 95% CI) | 1.20 [1.10, 1.30] |
| 3.3.1 Advanced stage (III/IV) | 3 | | Hazard Ratio (IV, Random, 95% CI) | 1.20 [1.10, 1.30] |
| 3.4 Overall survival sensitivity analysis including studies that included 0 cm | 2 | | Hazard Ratio (IV, Random, 95% CI) | 1.37 [1.09, 1.72] |
| 3.4.1 Advanced stage (III/IV) | 1 | | Hazard Ratio (IV, Random, 95% CI) | 1.67 [1.03, 2.72] |
| 3.4.2 Stage IV | 1 | | Hazard Ratio (IV, Random, 95% CI) | 1.30 [1.00, 1.68] |
| 3.5 Progression-free survival | 2 | | Hazard Ratio (IV, Random, 95% CI) | 1.30 [1.08, 1.56] |
| 3.5.1 Advanced stage (III/IV) | 1 | | Hazard Ratio (IV, Random, 95% CI) | 1.22 [1.12, 1.33] |
| 3.5.2 Stage IV | 1 | | Hazard Ratio (IV, Random, 95% CI) | 1.49 [1.16, 1.92] |

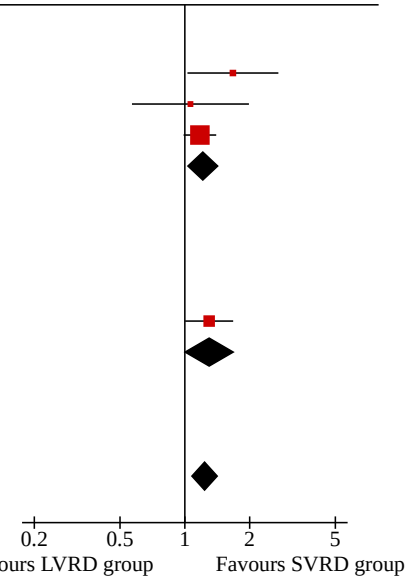
## Analysis 3.1. Comparison 3: PDS: LVRD (> 1 cm) versus SVRD (< 1 cm), Outcome 1: Overall survival

| Study or Subgroup | log[Hazard Ratio] | SE | Weight | Hazard Ratio IV, Random, 95% CI |
|---|---|---|---|---|
| **3.1.1 Advanced stage (III/IV)** | | | | |
| Chan 2003 | 0.514 | 0.248 | 2.7% | 1.67 [1.03 , 2.72] |
| Klar 2016 | 0.19 | 0.05 | 65.5% | 1.21 [1.10 , 1.33] |
| Melamed 2017a | 0.06 | 0.319 | 1.6% | 1.06 [0.57 , 1.98] |
| Melamed 2017b | 0.161 | 0.089 | 20.7% | 1.17 [0.99 , 1.40] |
| **Subtotal (95% CI)** | | | **90.5%** | **1.21 [1.11 , 1.32]** |
| Heterogeneity: Tau² = 0.00; Chi² = 1.98, df = 3 (P = 0.58); I² = 0% | | | | |
| Test for overall effect: Z = 4.48 (P < 0.00001) | | | | |
| **3.1.2 Stage IV** | | | | |
| Winter 2008 | 0.26 | 0.131 | 9.5% | 1.30 [1.00 , 1.68] |
| **Subtotal (95% CI)** | | | **9.5%** | **1.30 [1.00 , 1.68]** |
| Heterogeneity: Not applicable | | | | |
| Test for overall effect: Z = 1.98 (P = 0.05) | | | | |
| **Total (95% CI)** | | | **100.0%** | **1.22 [1.13 , 1.32]** |
| Heterogeneity: Tau² = 0.00; Chi² = 2.23, df = 4 (P = 0.69); I² = 0% | | | | |
| Test for overall effect: Z = 4.87 (P < 0.00001) | | | | |
| Test for subgroup differences: Chi² = 0.25, df = 1 (P = 0.61), I² = 0% | | | | |



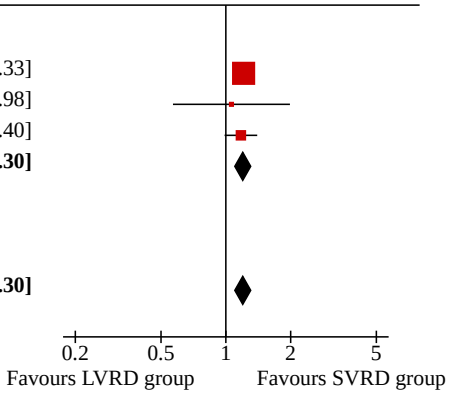

## Analysis 3.2. Comparison 3: PDS: LVRD (> 1 cm) versus SVRD (< 1 cm), Outcome 2: Overall survival sensitivity analysis excluding Klar 2016

| Study or Subgroup | log[Hazard Ratio] | SE | Weight | Hazard Ratio IV, Random, 95% CI |
|---|---|---|---|---|
| **3.2.1 Advanced stage (III/IV)** | | | | |
| Chan 2003 | 0.514 | 0.248 | 7.7% | 1.67 [1.03 , 2.72] |
| Melamed 2017a | 0.06 | 0.319 | 4.7% | 1.06 [0.57 , 1.98] |
| Melamed 2017b | 0.161 | 0.089 | 59.9% | 1.17 [0.99 , 1.40] |
| **Subtotal (95% CI)** | | | **72.3%** | **1.21 [1.03 , 1.42]** |
| Heterogeneity: Tau² = 0.00; Chi² = 1.98, df = 2 (P = 0.37); I² = 0% | | | | |
| Test for overall effect: Z = 2.37 (P = 0.02) | | | | |
| **3.2.2 Stage IV** | | | | |
| Winter 2008 | 0.26 | 0.131 | 27.7% | 1.30 [1.00 , 1.68] |
| **Subtotal (95% CI)** | | | **27.7%** | **1.30 [1.00 , 1.68]** |
| Heterogeneity: Not applicable | | | | |
| Test for overall effect: Z = 1.98 (P = 0.05) | | | | |
| **Total (95% CI)** | | | **100.0%** | **1.23 [1.08 , 1.41]** |
| Heterogeneity: Tau² = 0.00; Chi² = 2.17, df = 3 (P = 0.54); I² = 0% | | | | |
| Test for overall effect: Z = 3.06 (P = 0.002) | | | | |
| Test for subgroup differences: Chi² = 0.19, df = 1 (P = 0.66), I² = 0% | | | | |



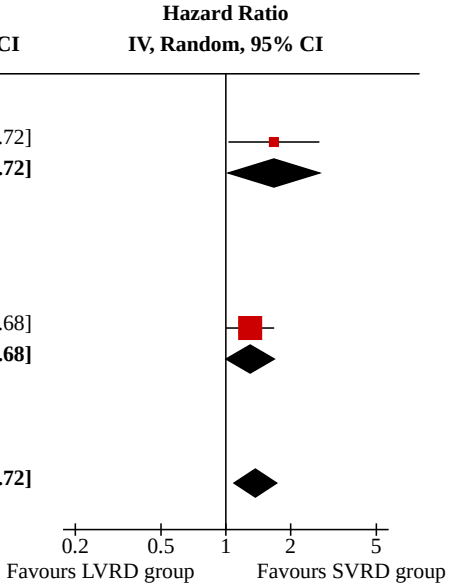
## Analysis 3.3. Comparison 3: PDS: LVRD (> 1 cm) versus SVRD (< 1 cm), Outcome 3: Overall survival sensitivity analysis excluding 0 cm

| Study or Subgroup | log[Hazard Ratio] | SE | Weight | Hazard Ratio IV, Random, 95% CI |
|---|---|---|---|---|
| **3.3.1 Advanced stage (III/IV)** | | | | |
| Klar 2016 | 0.19 | 0.05 | 74.6% | 1.21 [1.10 , 1.33] |
| Melamed 2017a | 0.06 | 0.319 | 1.8% | 1.06 [0.57 , 1.98] |
| Melamed 2017b | 0.161 | 0.089 | 23.6% | 1.17 [0.99 , 1.40] |
| **Subtotal (95% CI)** | | | **100.0%** | **1.20 [1.10 , 1.30]** |
| Heterogeneity: Tau² = 0.00; Chi² = 0.23, df = 2 (P = 0.89); I² = 0% | | | | |
| Test for overall effect: Z = 4.19 (P < 0.0001) | | | | |
| **Total (95% CI)** | | | **100.0%** | **1.20 [1.10 , 1.30]** |
| Heterogeneity: Tau² = 0.00; Chi² = 0.23, df = 2 (P = 0.89); I² = 0% | | | | |
| Test for overall effect: Z = 4.19 (P < 0.0001) | | | | |
| Test for subgroup differences: Not applicable | | | | |

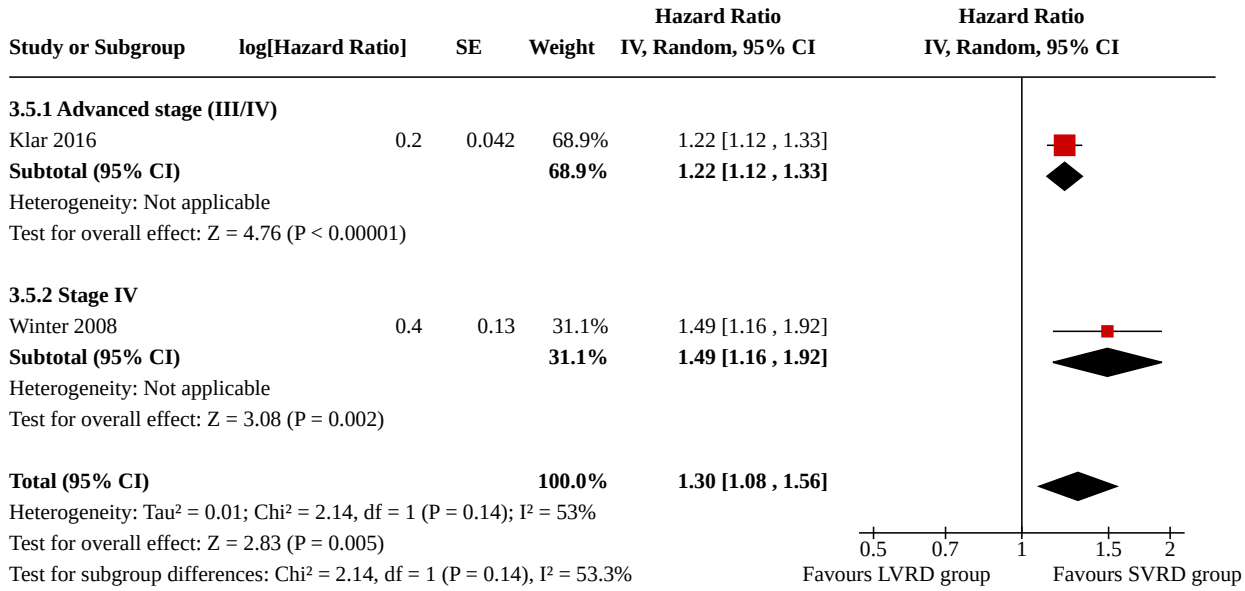



## Analysis 3.4. Comparison 3: PDS: LVRD (> 1 cm) versus SVRD (< 1 cm), Outcome 4: Overall survival sensitivity analysis including studies that included 0 cm

| Study or Subgroup | log[Hazard Ratio] | SE | Weight | Hazard Ratio IV, Random, 95% CI |
|---|---|---|---|---|
| **3.4.1 Advanced stage (III/IV)** | | | | |
| Chan 2003 | 0.514 | 0.248 | 21.8% | 1.67 [1.03 , 2.72] |
| **Subtotal (95% CI)** | | | **21.8%** | **1.67 [1.03 , 2.72]** |
| Heterogeneity: Not applicable | | | | |
| Test for overall effect: Z = 2.07 (P = 0.04) | | | | |
| **3.4.2 Stage IV** | | | | |
| Winter 2008 | 0.26 | 0.131 | 78.2% | 1.30 [1.00 , 1.68] |
| **Subtotal (95% CI)** | | | **78.2%** | **1.30 [1.00 , 1.68]** |
| Heterogeneity: Not applicable | | | | |
| Test for overall effect: Z = 1.98 (P = 0.05) | | | | |
| **Total (95% CI)** | | | **100.0%** | **1.37 [1.09 , 1.72]** |
| Heterogeneity: Tau² = 0.00; Chi² = 0.82, df = 1 (P = 0.37); I² = 0% | | | | |
| Test for overall effect: Z = 2.72 (P = 0.006) | | | | |
| Test for subgroup differences: Chi² = 0.82, df = 1 (P = 0.37), I² = 0% | | | | |



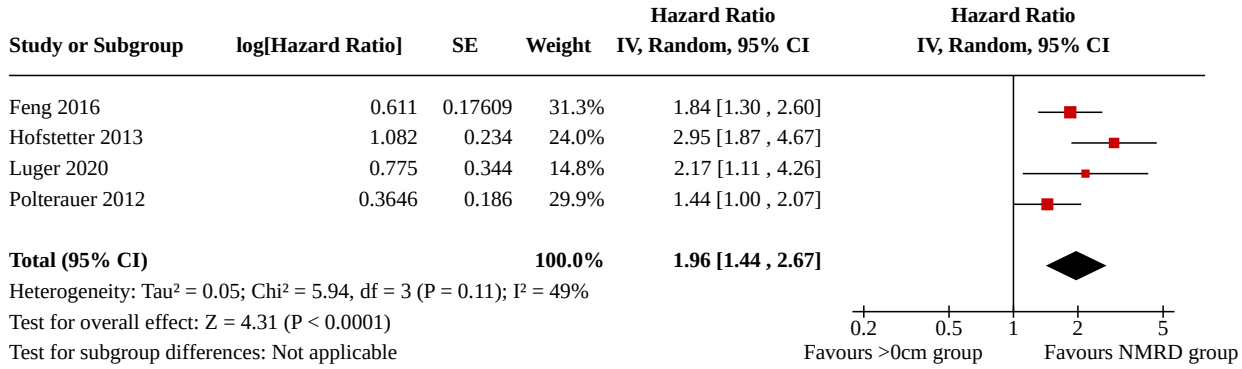
### Analysis 3.5. Comparison 3: PDS: LVRD (> 1 cm) versus SVRD (< 1 cm), Outcome 5: Progression-free survival




## Comparison 4. PDS: RD > 0 cm versus NMRD

| Outcome or subgroup title | No. of studies | No. of participants | Statistical method | Effect size |
|---|---|---|---|---|
| 4.1 Overall survival | 4 | | Hazard Ratio (IV, Random, 95% CI) | 1.96 [1.44, 2.67] |
| 4.2 Progression-free survival | 3 | | Hazard Ratio (IV, Random, 95% CI) | 1.60 [1.36, 1.89] |

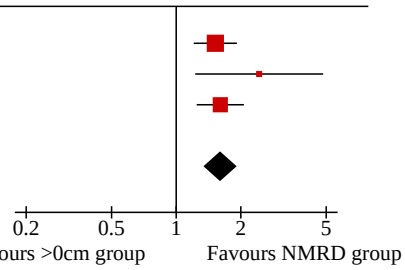
### Analysis 4.1. Comparison 4: PDS: RD > 0 cm versus NMRD, Outcome 1: Overall survival



### Analysis 4.2. Comparison 4: PDS: RD > 0 cm versus NMRD, Outcome 2: Progression-free survival

| Study or Subgroup | log[Hazard Ratio] | SE | Weight | Hazard Ratio IV, Random, 95% CI |
|---|---|---|---|---|
| Feng 2016 | 0.42 | 0.118 | 51.3% | 1.52 [1.21 , 1.92] |
| Luger 2020 | 0.89 | 0.35 | 5.8% | 2.44 [1.23 , 4.84] |
| Polterauer 2012 | 0.47312 | 0.129 | 42.9% | 1.60 [1.25 , 2.07] |
| **Total (95% CI)** | | | **100.0%** | **1.60 [1.36 , 1.89]** |

Heterogeneity: Tau² = 0.00; Chi² = 1.62, df = 2 (P = 0.44); I² = 0%
Test for overall effect: Z = 5.56 (P < 0.00001)
Test for subgroup differences: Not applicable

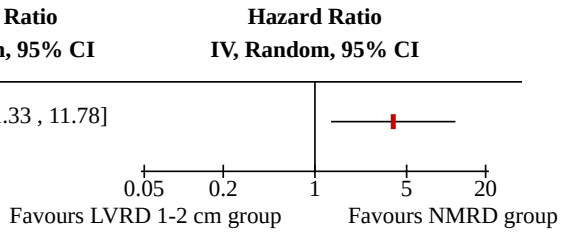



## Comparison 5. PDS: LVRD 1 cm to 2 cm versus NMRD (stage IIIC)

| Outcome or subgroup title | No. of studies | No. of participants | Statistical method | Effect size |
|---|---|---|---|---|
| 5.1 Overall survival | 1 | | Hazard Ratio (IV, Random, 95% CI) | Subtotals only |

### Analysis 5.1. Comparison 5: PDS: LVRD 1 cm to 2 cm versus NMRD (stage IIIC), Outcome 1: Overall survival

| Study or Subgroup | log[Hazard Ratio] | SE | Hazard Ratio IV, Random, 95% CI |
|---|---|---|---|
| Aletti 2006 | 1.3745 | 0.557 | 3.95 [1.33 , 11.78] |

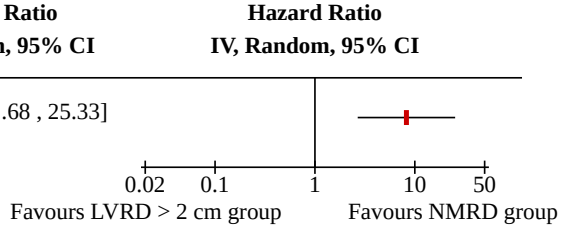



## Comparison 6. PDS: LVRD (> 2 cm) versus NMRD (stage IIIC)

| Outcome or subgroup title | No. of studies | No. of participants | Statistical method | Effect size |
|---|---|---|---|---|
| 6.1 Overall survival | 1 | | Hazard Ratio (IV, Random, 95% CI) | Subtotals only |

### Analysis 6.1. Comparison 6: PDS: LVRD (> 2 cm) versus NMRD (stage IIIC), Outcome 1: Overall survival

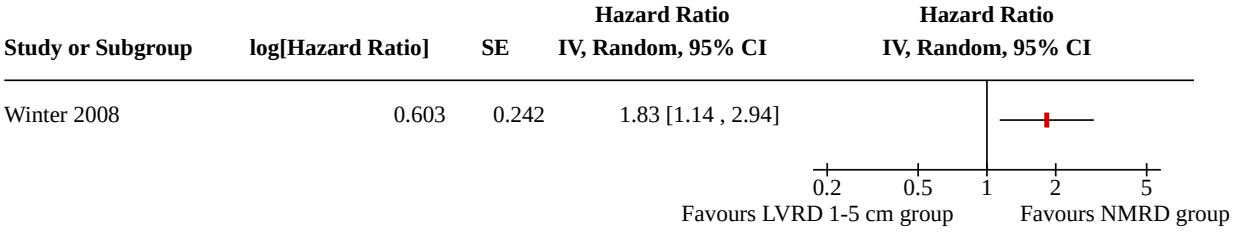
| Study or Subgroup | log[Hazard Ratio] | SE | Hazard Ratio IV, Random, 95% CI |
|---|---|---|---|
| Aletti 2006 | 2.109 | 0.573 | 8.24 [2.68 , 25.33] |



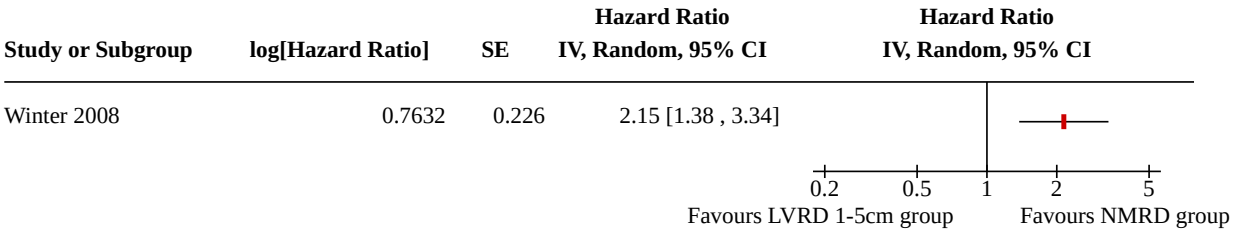
### Comparison 7. PDS: LVRD 1 cm to 5 cm versus NMRD (stage IV disease)

| Outcome or subgroup title | No. of studies | No. of participants | Statistical method | Effect size |
|---|---|---|---|---|
| 7.1 Overall survival | 1 | | Hazard Ratio (IV, Random, 95% CI) | Subtotals only |
| 7.2 Progression-free survival | 1 | | Hazard Ratio (IV, Random, 95% CI) | Subtotals only |

### Analysis 7.1. Comparison 7: PDS: LVRD 1 cm to 5 cm versus NMRD (stage IV disease), Outcome 1: Overall survival



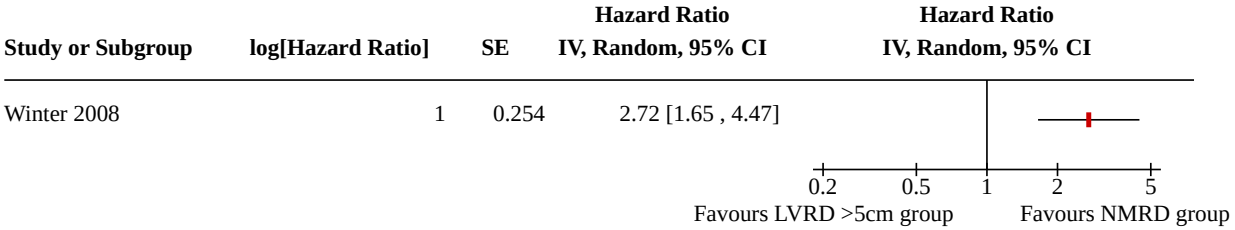

### Analysis 7.2. Comparison 7: PDS: LVRD 1 cm to 5 cm versus NMRD (stage IV disease), Outcome 2: Progression-free survival



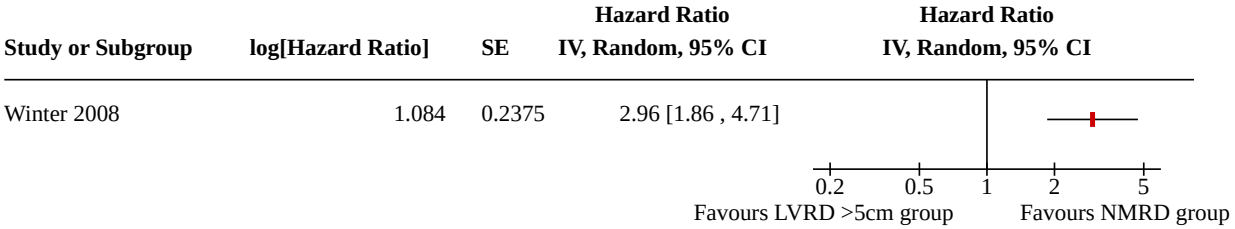

### Comparison 8. PDS: LVRD (> 5 cm) versus NMRD (stage IV disease)

| Outcome or subgroup title | No. of studies | No. of participants | Statistical method | Effect size |
|---|---|---|---|---|
| 8.1 Overall survival | 1 | | Hazard Ratio (IV, Random, 95% CI) | Subtotals only |
| 8.2 Progression-free survival | 1 | | Hazard Ratio (IV, Random, 95% CI) | Subtotals only |

**Analysis 8.1. Comparison 8: PDS: LVRD (> 5 cm) versus NMRD (stage IV disease), Outcome 1: Overall survival**



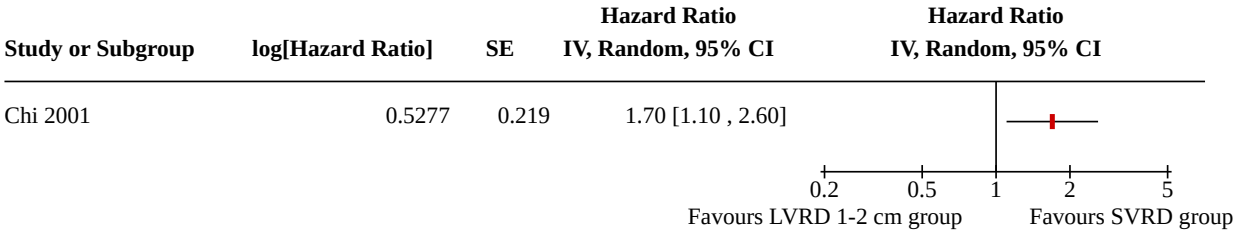

**Analysis 8.2. Comparison 8: PDS: LVRD (> 5 cm) versus NMRD (stage IV disease), Outcome 2: Progression-free survival**




**Comparison 9. PDS: LVRD 1 cm to 2 cm versus SVRD (< 1 cm)**

| Outcome or subgroup title | No. of studies | No. of participants | Statistical method | Effect size |
|---|---|---|---|---|
| 9.1 Overall survival | 1 | | Hazard Ratio (IV, Random, 95% CI) | Subtotals only |

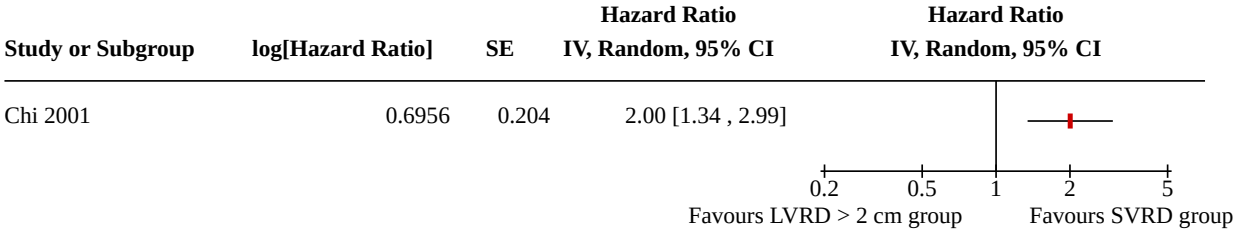
**Analysis 9.1. Comparison 9: PDS: LVRD 1 cm to 2 cm versus SVRD (< 1 cm), Outcome 1: Overall survival**




**Comparison 10. PDS: LVRD (> 2 cm) versus SVRD (< 1 cm)**

| Outcome or subgroup title | No. of studies | No. of participants | Statistical method | Effect size |
|---|---|---|---|---|
| 10.1 Overall survival | 1 | | Hazard Ratio (IV, Random, 95% CI) | Subtotals only |

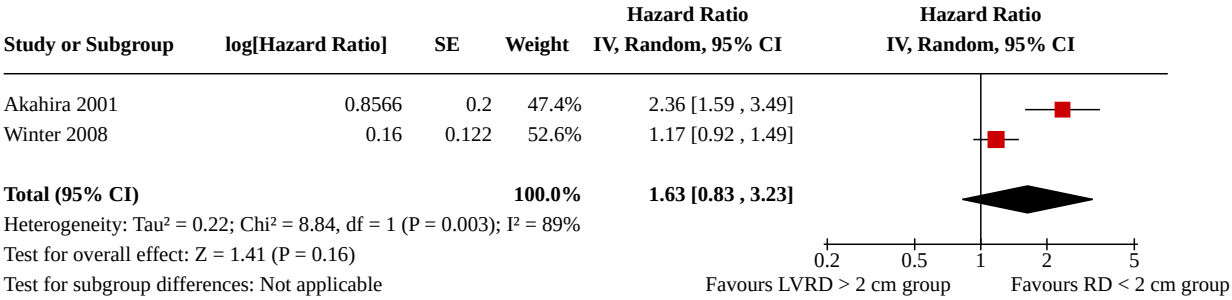
### Analysis 10.1. Comparison 10: PDS: LVRD (> 2 cm) versus SVRD (< 1 cm), Outcome 1: Overall survival



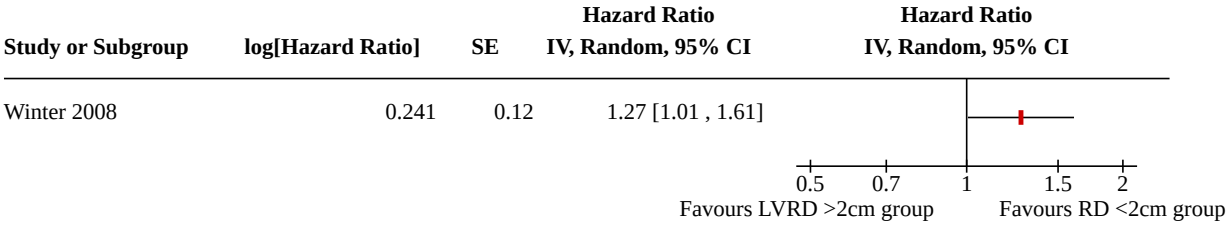

## Comparison 11. PDS: LVRD (> 2 cm) versus RD < 2 cm (stage IV disease)

| Outcome or subgroup title | No. of studies | No. of participants | Statistical method | Effect size |
|---|---|---|---|---|
| 11.1 Overall survival | 2 | | Hazard Ratio (IV, Random, 95% CI) | 1.63 [0.83, 3.23] |
| 11.2 Progression-free survival | 1 | | Hazard Ratio (IV, Random, 95% CI) | Subtotals only |

### Analysis 11.1. Comparison 11: PDS: LVRD (> 2 cm) versus RD < 2 cm (stage IV disease), Outcome 1: Overall survival




### Analysis 11.2. Comparison 11: PDS: LVRD (> 2 cm) versus RD < 2 cm (stage IV disease), Outcome 2: Progression-free survival

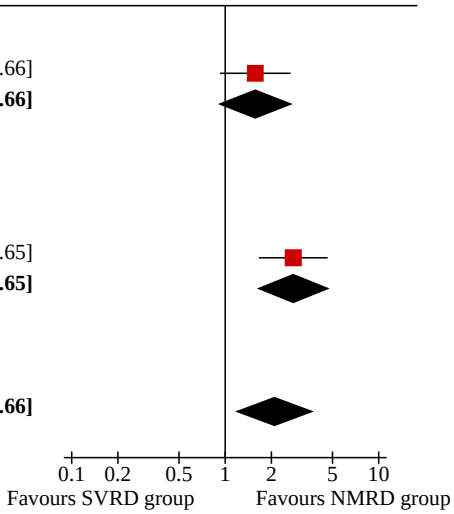


**Comparison 12. IDS: SVRD (< 1 cm) versus NMRD**

| Outcome or subgroup title | No. of studies | No. of participants | Statistical method | Effect size |
|---|---|---|---|---|
| 12.1 Overall survival | 1 | | Hazard Ratio (IV, Random, 95% CI) | 2.09 [1.20, 3.66] |
| 12.1.1 ≤ 4 cycles of NACT | 1 | | Hazard Ratio (IV, Random, 95% CI) | 1.57 [0.93, 2.66] |
| 12.1.2 > 4 cycles of NACT | 1 | | Hazard Ratio (IV, Random, 95% CI) | 2.78 [1.66, 4.65] |
| 12.2 Progression-free survival | 2 | | Hazard Ratio (IV, Random, 95% CI) | 3.03 [0.81, 11.38] |

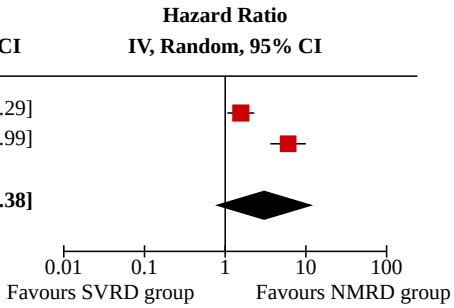
**Analysis 12.1. Comparison 12: IDS: SVRD (< 1 cm) versus NMRD, Outcome 1: Overall survival**

| Study or Subgroup | log[Hazard Ratio] | SE | Weight | Hazard Ratio IV, Random, 95% CI |
|---|---|---|---|---|
| **12.1.1 ≤ 4 cycles of NACT** | | | | |
| Phillips 2018 | 0.4513 | 0.2698 | 49.4% | 1.57 [0.93 , 2.66] |
| **Subtotal (95% CI)** | | | **49.4%** | **1.57 [0.93 , 2.66]** |
| Heterogeneity: Not applicable | | | | |
| Test for overall effect: Z = 1.67 (P = 0.09) | | | | |
| **12.1.2 > 4 cycles of NACT** | | | | |
| Phillips 2018 | 1.021 | 0.263 | 50.6% | 2.78 [1.66 , 4.65] |
| **Subtotal (95% CI)** | | | **50.6%** | **2.78 [1.66 , 4.65]** |
| Heterogeneity: Not applicable | | | | |
| Test for overall effect: Z = 3.88 (P = 0.0001) | | | | |
| **Total (95% CI)** | | | **100.0%** | **2.09 [1.20 , 3.66]** |
| Heterogeneity: Tau² = 0.09; Chi² = 2.29, df = 1 (P = 0.13); I² = 56% | | | | |
| Test for overall effect: Z = 2.60 (P = 0.009) | | | | |
| Test for subgroup differences: Chi² = 2.29, df = 1 (P = 0.13), I² = 56.3% | | | | |




**Analysis 12.2. Comparison 12: IDS: SVRD (< 1 cm) versus NMRD, Outcome 2: Progression-free survival**

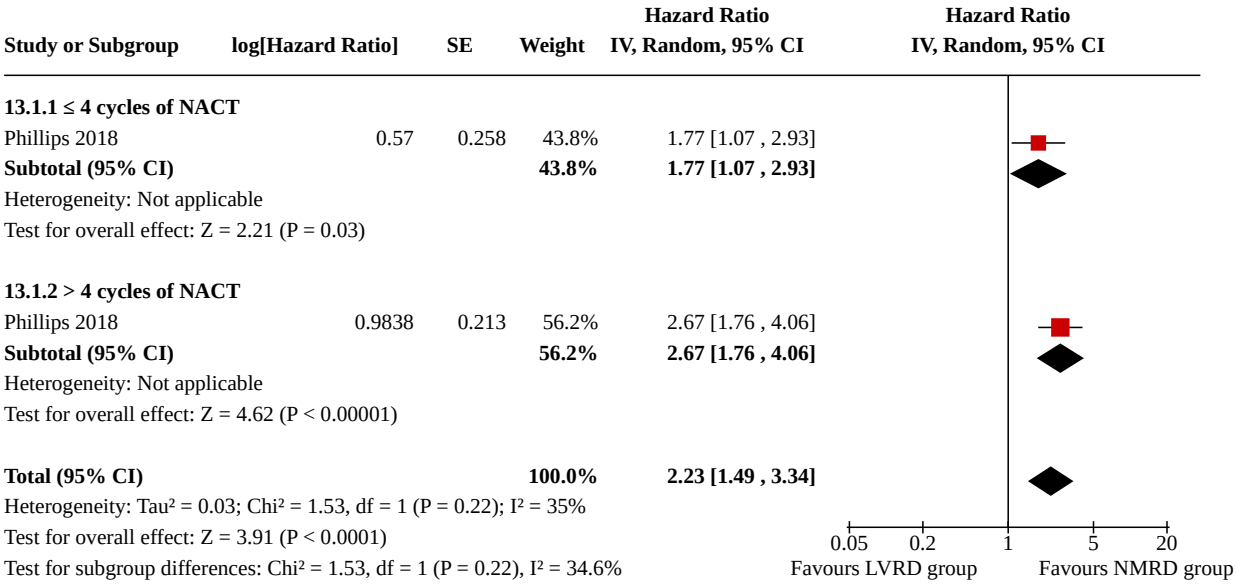
| Study or Subgroup | log[Hazard Ratio] | SE | Weight | Hazard Ratio IV, Random, 95% CI |
|---|---|---|---|---|
| Bixel 2020 | 0.445 | 0.195 | 50.8% | 1.56 [1.06 , 2.29] |
| Liu 2020 | 1.795 | 0.2585 | 49.2% | 6.02 [3.63 , 9.99] |
| **Total (95% CI)** | | | **100.0%** | **3.03 [0.81 , 11.38]** |
| Heterogeneity: Tau² = 0.86; Chi² = 17.38, df = 1 (P < 0.0001); I² = 94% | | | | |
| Test for overall effect: Z = 1.64 (P = 0.10) | | | | |
| Test for subgroup differences: Not applicable | | | | |



### Comparison 13. IDS: LVRD (> 1 cm) versus NMRD

| Outcome or subgroup title | No. of studies | No. of participants | Statistical method | Effect size |
|---|---|---|---|---|
| 13.1 Overall survival | 1 | | Hazard Ratio (IV, Random, 95% CI) | 2.23 [1.49, 3.34] |
| 13.1.1 ≤ 4 cycles of NACT | 1 | | Hazard Ratio (IV, Random, 95% CI) | 1.77 [1.07, 2.93] |
| 13.1.2 > 4 cycles of NACT | 1 | | Hazard Ratio (IV, Random, 95% CI) | 2.67 [1.76, 4.06] |

### Analysis 13.1. Comparison 13: IDS: LVRD (> 1 cm) versus NMRD, Outcome 1: Overall survival



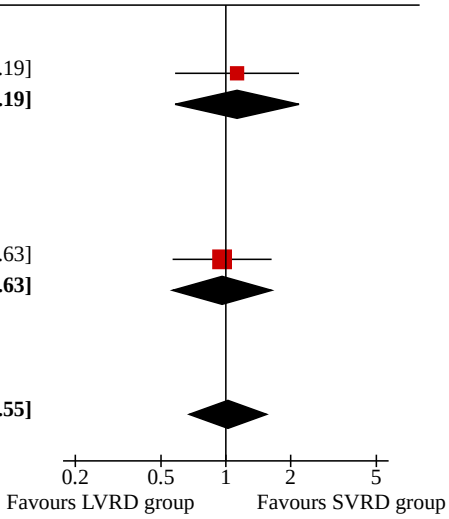

### Comparison 14. IDS: LVRD (> 1 cm) versus SVRD (< 1 cm)

| Outcome or subgroup title | No. of studies | No. of participants | Statistical method | Effect size |
|---|---|---|---|---|
| 14.1 Overall survival | 1 | | Hazard Ratio (IV, Random, 95% CI) | 1.02 [0.68, 1.55] |
| 14.1.1 ≤ 4 cycles of NACT | 1 | | Hazard Ratio (IV, Random, 95% CI) | 1.13 [0.58, 2.19] |
| 14.1.2 > 4 cycles of NACT | 1 | | Hazard Ratio (IV, Random, 95% CI) | 0.96 [0.57, 1.63] |
| 14.2 Overall survival sensitivity analysis including 0 cm | 6 | | Hazard Ratio (IV, Random, 95% CI) | 1.60 [1.21, 2.11] |
| 14.2.1 3 cycles of NACT | 2 | | Hazard Ratio (IV, Random, 95% CI) | 1.80 [0.94, 3.43] |

| Outcome or subgroup title | No. of studies | No. of participants | Statistical method | Effect size |
|---|---|---|---|---|
| 14.2.2 Median 4 cycles of NACT (disease-specific survival) | 1 | | Hazard Ratio (IV, Random, 95% CI) | 1.70 [1.06, 2.75] |
| 14.2.3 Mean/median ~ 6 cycles of NACT | 2 | | Hazard Ratio (IV, Random, 95% CI) | 2.19 [1.07, 4.50] |
| 14.2.4 ≤ 4 cycles of NACT | 1 | | Hazard Ratio (IV, Random, 95% CI) | 1.13 [0.58, 2.19] |
| 14.2.5 > 4 cycles of NACT | 1 | | Hazard Ratio (IV, Random, 95% CI) | 0.96 [0.57, 1.63] |
| 14.3 Overall survival sensitivity analysis excluding Phillips 2018 | 5 | | Hazard Ratio (IV, Random, 95% CI) | 1.84 [1.34, 2.52] |
| 14.3.1 3 cycles of NACT | 2 | | Hazard Ratio (IV, Random, 95% CI) | 1.80 [0.94, 3.43] |
| 14.3.2 Median 4 cycles of NACT (disease-specific survival) | 1 | | Hazard Ratio (IV, Random, 95% CI) | 1.70 [1.06, 2.75] |
| 14.3.3 Mean/median ~ 6 cycles of NACT | 2 | | Hazard Ratio (IV, Random, 95% CI) | 2.19 [1.07, 4.50] |
| 14.4 Progression-free survival | 4 | | Hazard Ratio (IV, Random, 95% CI) | 1.76 [1.23, 2.52] |
| 14.4.1 3 cycles of NACT | 2 | | Hazard Ratio (IV, Random, 95% CI) | 1.68 [0.90, 3.14] |
| 14.4.2 Mean ~ 6 cycles of NACT | 1 | | Hazard Ratio (IV, Random, 95% CI) | 2.32 [1.08, 4.97] |
| 14.4.3 All cycles | 1 | | Hazard Ratio (IV, Random, 95% CI) | 1.82 [1.12, 2.97] |

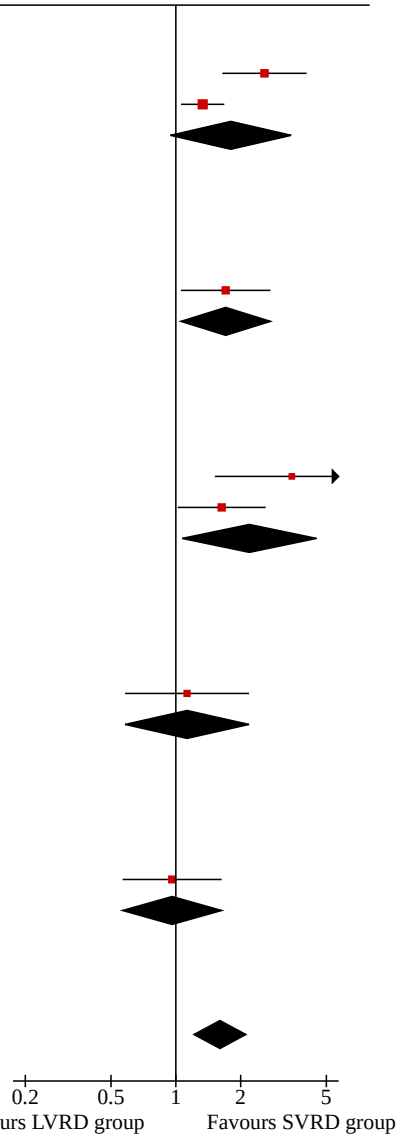
**Analysis 14.1. Comparison 14: IDS: LVRD (> 1 cm) versus SVRD (< 1 cm), Outcome 1: Overall survival**

| Study or Subgroup | log[Hazard Ratio] | SE | Weight | Hazard Ratio IV, Random, 95% CI |
|---|---|---|---|---|
| **14.1.1 ≤ 4 cycles of NACT** | | | | |
| Phillips 2018 | 0.12 | 0.338 | 39.0% | 1.13 [0.58 , 2.19] |
| **Subtotal (95% CI)** | | | **39.0%** | **1.13 [0.58 , 2.19]** |
| Heterogeneity: Not applicable | | | | |
| Test for overall effect: Z = 0.36 (P = 0.72) | | | | |
| **14.1.2 > 4 cycles of NACT** | | | | |
| Phillips 2018 | -0.04 | 0.27 | 61.0% | 0.96 [0.57 , 1.63] |
| **Subtotal (95% CI)** | | | **61.0%** | **0.96 [0.57 , 1.63]** |
| Heterogeneity: Not applicable | | | | |
| Test for overall effect: Z = 0.15 (P = 0.88) | | | | |
| **Total (95% CI)** | | | **100.0%** | **1.02 [0.68 , 1.55]** |
| Heterogeneity: Tau² = 0.00; Chi² = 0.14, df = 1 (P = 0.71); I² = 0% | | | | |
| Test for overall effect: Z = 0.11 (P = 0.92) | | | | |
| Test for subgroup differences: Chi² = 0.14, df = 1 (P = 0.71), I² = 0% | | | | |



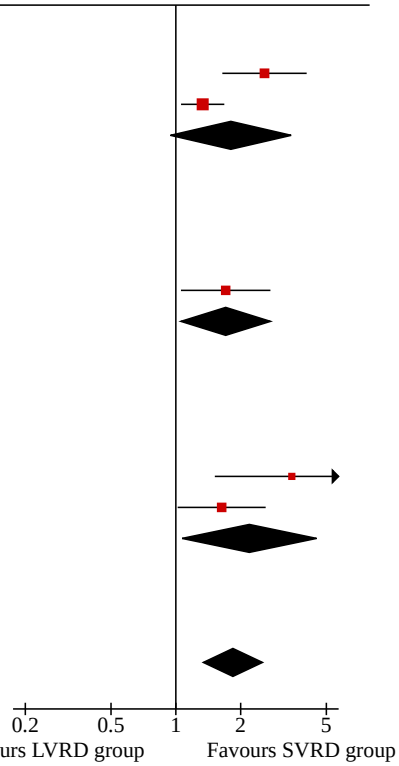
## Analysis 14.2. Comparison 14: IDS: LVRD (> 1 cm) versus SVRD (< 1 cm), Outcome 2: Overall survival sensitivity analysis including 0 cm

| Study or Subgroup | log[Hazard Ratio] | SE | Weight | Hazard Ratio IV, Random, 95% CI |
|---|---|---|---|---|
| **14.2.1 3 cycles of NACT** | | | | |
| Zhang 2018 | 0.948 | 0.23 | 15.6% | 2.58 [1.64 , 4.05] |
| Zhu 2016 | 0.287 | 0.118 | 22.5% | 1.33 [1.06 , 1.68] |
| **Subtotal (95% CI)** | | | **38.1%** | **1.80 [0.94 , 3.43]** |
| Heterogeneity: Tau² = 0.19; Chi² = 6.54, df = 1 (P = 0.01); I² = 85% | | | | |
| Test for overall effect: Z = 1.79 (P = 0.07) | | | | |
| **14.2.2 Median 4 cycles of NACT (disease-specific survival)** | | | | |
| Davidson 2019 | 0.533 | 0.244 | 14.8% | 1.70 [1.06 , 2.75] |
| **Subtotal (95% CI)** | | | **14.8%** | **1.70 [1.06 , 2.75]** |
| Heterogeneity: Not applicable | | | | |
| Test for overall effect: Z = 2.18 (P = 0.03) | | | | |
| **14.2.3 Mean/median ~ 6 cycles of NACT** | | | | |
| Cioffi 2018 | 1.24 | 0.42 | 7.9% | 3.46 [1.52 , 7.87] |
| Kaban 2017 | 0.49 | 0.24 | 15.1% | 1.63 [1.02 , 2.61] |
| **Subtotal (95% CI)** | | | **23.0%** | **2.19 [1.07 , 4.50]** |
| Heterogeneity: Tau² = 0.16; Chi² = 2.40, df = 1 (P = 0.12); I² = 58% | | | | |
| Test for overall effect: Z = 2.14 (P = 0.03) | | | | |
| **14.2.4 ≤ 4 cycles of NACT** | | | | |
| Phillips 2018 | 0.12 | 0.338 | 10.6% | 1.13 [0.58 , 2.19] |
| **Subtotal (95% CI)** | | | **10.6%** | **1.13 [0.58 , 2.19]** |
| Heterogeneity: Not applicable | | | | |
| Test for overall effect: Z = 0.36 (P = 0.72) | | | | |
| **14.2.5 > 4 cycles of NACT** | | | | |
| Phillips 2018 | -0.04 | 0.27 | 13.5% | 0.96 [0.57 , 1.63] |
| **Subtotal (95% CI)** | | | **13.5%** | **0.96 [0.57 , 1.63]** |
| Heterogeneity: Not applicable | | | | |
| Test for overall effect: Z = 0.15 (P = 0.88) | | | | |
| **Total (95% CI)** | | | **100.0%** | **1.60 [1.21 , 2.11]** |
| Heterogeneity: Tau² = 0.07; Chi² = 14.34, df = 6 (P = 0.03); I² = 58% | | | | |
| Test for overall effect: Z = 3.33 (P = 0.0009) | | | | |
| Test for subgroup differences: Chi² = 5.02, df = 4 (P = 0.28), I² = 20.4% | | | | |

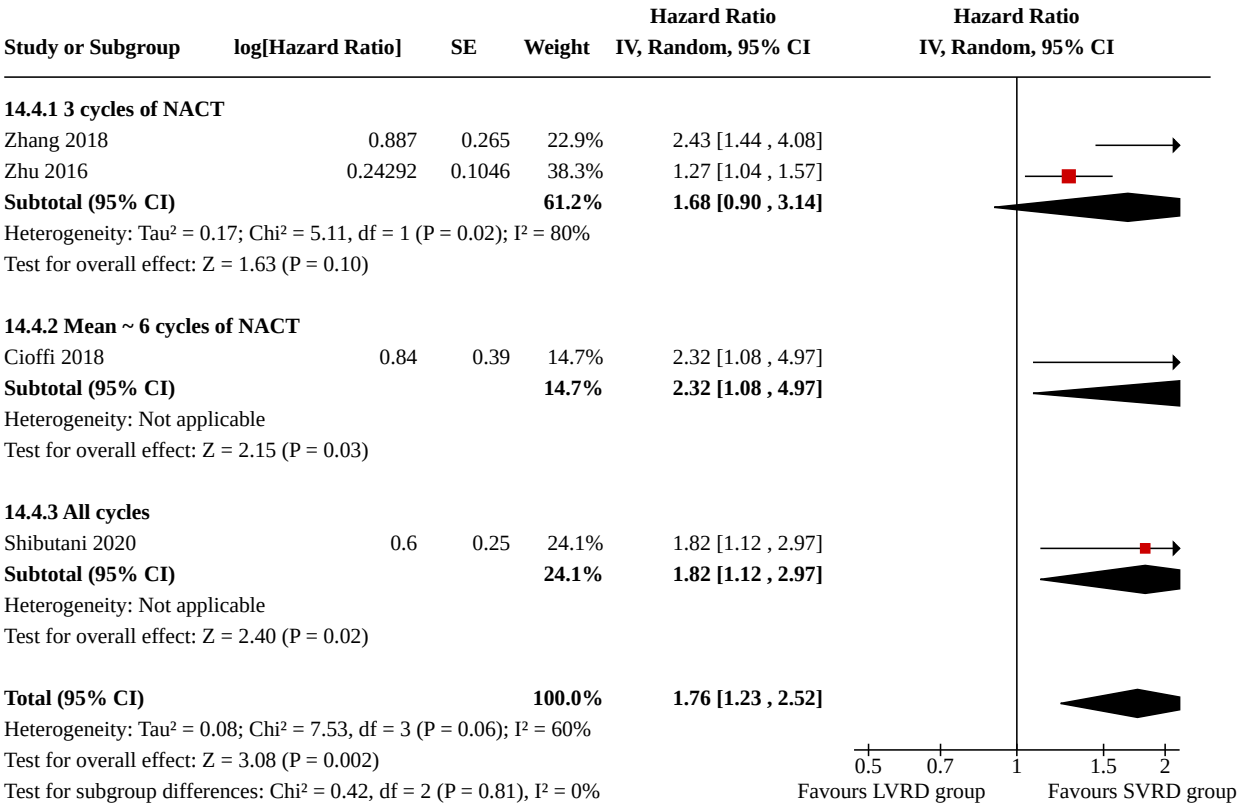


## Analysis 14.3. Comparison 14: IDS: LVRD (> 1 cm) versus SVRD (< 1 cm), Outcome 3: Overall survival sensitivity analysis excluding Phillips 2018

| Study or Subgroup | log[Hazard Ratio] | SE | Weight | Hazard Ratio IV, Random, 95% CI |
|---|---|---|---|---|
| **14.3.1 3 cycles of NACT** | | | | |
| Zhang 2018 | 0.948 | 0.23 | 20.6% | 2.58 [1.64 , 4.05] |
| Zhu 2016 | 0.287 | 0.118 | 29.7% | 1.33 [1.06 , 1.68] |
| **Subtotal (95% CI)** | | | **50.3%** | **1.80 [0.94 , 3.43]** |
| Heterogeneity: Tau² = 0.19; Chi² = 6.54, df = 1 (P = 0.01); I² = 85% | | | | |
| Test for overall effect: Z = 1.79 (P = 0.07) | | | | |
| **14.3.2 Median 4 cycles of NACT (disease-specific survival)** | | | | |
| Davidson 2019 | 0.533 | 0.244 | 19.5% | 1.70 [1.06 , 2.75] |
| **Subtotal (95% CI)** | | | **19.5%** | **1.70 [1.06 , 2.75]** |
| Heterogeneity: Not applicable | | | | |
| Test for overall effect: Z = 2.18 (P = 0.03) | | | | |
| **14.3.3 Mean/median ~ 6 cycles of NACT** | | | | |
| Cioffi 2018 | 1.24 | 0.42 | 10.4% | 3.46 [1.52 , 7.87] |
| Kaban 2017 | 0.49 | 0.24 | 19.8% | 1.63 [1.02 , 2.61] |
| **Subtotal (95% CI)** | | | **30.2%** | **2.19 [1.07 , 4.50]** |
| Heterogeneity: Tau² = 0.16; Chi² = 2.40, df = 1 (P = 0.12); I² = 58% | | | | |
| Test for overall effect: Z = 2.14 (P = 0.03) | | | | |
| **Total (95% CI)** | | | **100.0%** | **1.84 [1.34 , 2.52]** |
| Heterogeneity: Tau² = 0.07; Chi² = 10.14, df = 4 (P = 0.04); I² = 61% | | | | |
| Test for overall effect: Z = 3.79 (P = 0.0002) | | | | |
| Test for subgroup differences: Chi² = 0.33, df = 2 (P = 0.85), I² = 0% | | | | |



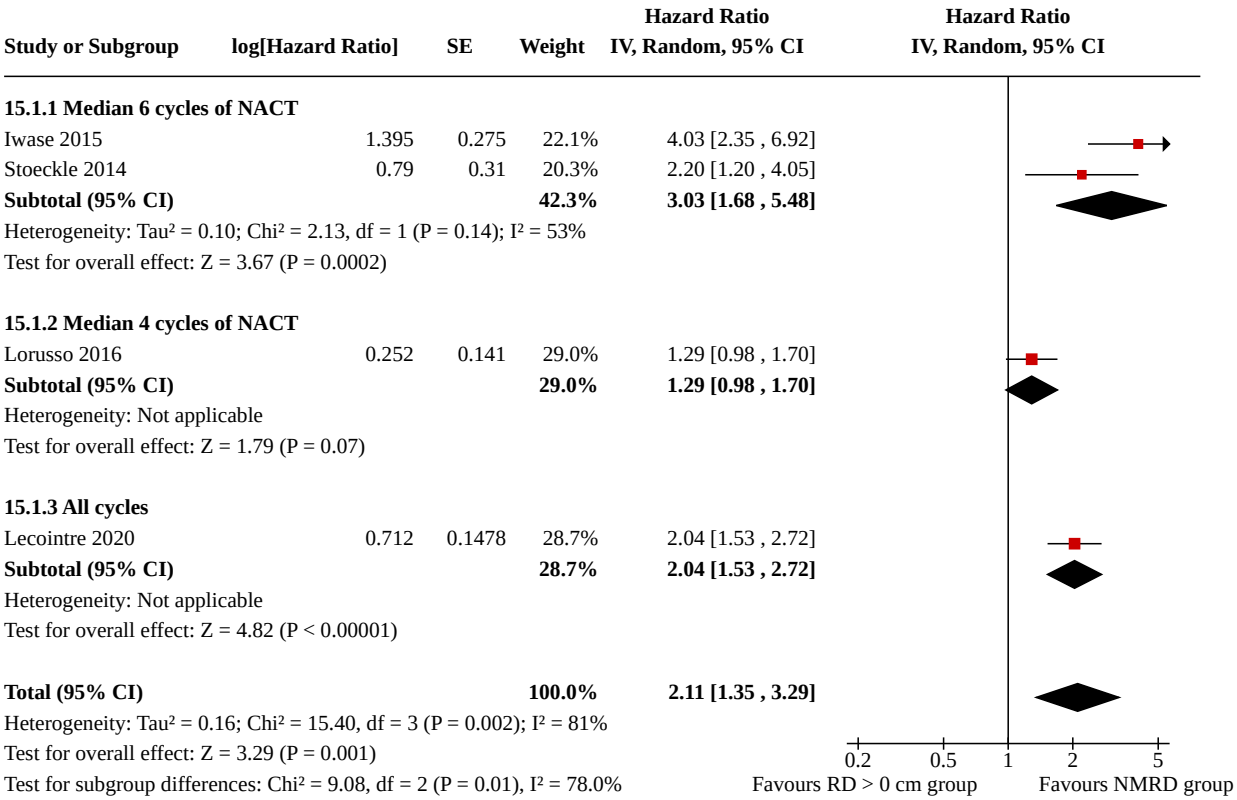
**Analysis 14.4. Comparison 14: IDS: LVRD (> 1 cm) versus SVRD (< 1 cm), Outcome 4: Progression-free survival**



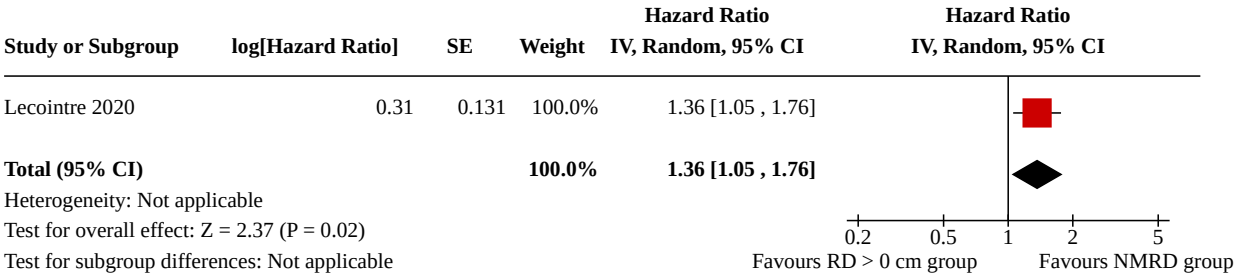

## Comparison 15. IDS: RD > 0 cm versus NMRD

| Outcome or subgroup title | No. of studies | No. of participants | Statistical method | Effect size |
|---|---|---|---|---|
| 15.1 Overall survival | 4 | | Hazard Ratio (IV, Random, 95% CI) | 2.11 [1.35, 3.29] |
| 15.1.1 Median 6 cycles of NACT | 2 | | Hazard Ratio (IV, Random, 95% CI) | 3.03 [1.68, 5.48] |
| 15.1.2 Median 4 cycles of NACT | 1 | | Hazard Ratio (IV, Random, 95% CI) | 1.29 [0.98, 1.70] |
| 15.1.3 All cycles | 1 | | Hazard Ratio (IV, Random, 95% CI) | 2.04 [1.53, 2.72] |
| 15.2 Progression-free survival | 1 | | Hazard Ratio (IV, Random, 95% CI) | 1.36 [1.05, 1.76] |

## Analysis 15.1. Comparison 15: IDS: RD > 0 cm versus NMRD, Outcome 1: Overall survival




## Analysis 15.2. Comparison 15: IDS: RD > 0 cm versus NMRD, Outcome 2: Progression-free survival



## ADDITIONAL TABLES

**Table 1. Summary of stage and residual disease in included upfront primary debulking surgery (PDS) studies**

| Study | No. | Stage | | Optimal | Suboptimal | Median follow-up | Median age in years | Setting |
|---|---|---|---|---|---|---|---|---|
| | n | III n (%) | IV n (%) | n (%) | n (%) | Months | (Range) | |
| Akahira 2001 | 225 | 0 (0) | 225 (100) | < 2: 70 (31) | > 2: 155 (69) | 47.5 (13 to 112) | 54 (26 to 85) | Japan |
| Aletti 2006 | 194 | 194 (100) | 0 (0) | 0: 46 (24)<br>< 1: 85 (44) | 1 to 2: 22 (11)<br>> 2: 41 (21) | 32.4 (0.2 to 126) | 64 (24 to 87) | USA |
| Ataseven 2016 | 326 | 0 (0) | 326 (100) | 0: 157 (55)<br>< 1: 88 (31) | > 1: 41 (14)<br>NS: n = 40 exc. | 34 (IQR: 12 to 70) | < 65: 205 (63)<br>> 65: 121 (37) | Germany<br>Austria |
| Bristow 2011 | 405 | 405 (100) | 0 (0) | 0: 209 (52)<br>< 1: 196 (48) | | 33.0 | 59 (Range not reported) | USA |
| Chan 2003 | 104 | 84 (81) | 20 (19) | < 1: 71 (68) | > 1: 33 (32) | 33 (6 to 142) | Mean was 50.5 years and 61 years for younger and older women, respectively (Range: 22 and 85) | USA |
| Chang 2012a | 203 | 189 (93) | 14 (7) | 0: 63 (31)<br>< 1: 77 (38) | > 1: 63 (31) | 43 (1 to 124) | 54 (30 to 78) | South Korea |
| Chang 2012b | 191 | 189 (100) | 0 (0) | 0: 61 (32)<br>< 1: 67 (36) | > 1: 61 (32) | Not reported | 54 (30 to 78) | South Korea |
| Chi 2001 | 282 | 216 (77) | 66 (23) | < 1: 71 (25)<br>1 to 2: 73 (26) | > 2: 137 (49) | 32 (1 to 139) | 59 (22 to 87) | USA |
| Chi 2006 | 465 | 465 | 0 | 0: 67 (14) | > 1: 229 (49) | 38 | 60 | USA |

**Table 1. Summary of stage and residual disease in included upfront primary debulking surgery (PDS) studies** *(Continued)*

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | (100) | (0) | < 1: 169 (37) | | (1 to 199) | (22 to 87) | |
| Cuylan 2018 | 218 | 218<br>(100) | 0<br>(0) | 0: 55 (25)<br>< 1: 163 (75) | | 31.5 | 54 (18 to 78) | Turkey |
| Eisenkop 2003 | 408 | 408<br>(100) | 0<br>(0) | 0: 351 (86)<br>< 1: 41 (10) | > 1: 16 (4) | 32.8 | 62.8<br>(24 to 91) | USA |
| Feng 2016 | 625 | n = 567 (91) stage III/IV | | 0: 209 (33) | > 0: 416 (67) | 29 (3 to 100) | 56 (30 to 84) | China |
| Hofstetter 2013 | 191 | 158<br>(83) | 33 (17) | 0: 121 (63) | > 0: 70 (37) | 42 | < 57: 98<br>> 57: 93 | Europe |
| Kahl 2017 | 793 | 428<br>(54) | 365<br>(46) | 0: 482 (61)<br>< 1: 226 (39) | > 1: 85 | 47<br>(IQR: 18 to 87) | 60 (19 to 88) | Germany |
| Klar 2016 | 5055 | 4488/5130 (87.5)<br>stage III/IV; n = 4850 in RD analysis | | 0: 1779 (37)<br>< 1: 1442 (30) | > 1: 1629 (33) | 0 to 144 | Mean: 57.4<br>(SD 10.53) | Germany<br>France<br>Denmark |
| Langstraat 2011 | 280 | 210<br>(76) | 67<br>(24) | 0: 61 (22)<br>< 1: 120 (43) | > 1: 95 (35) | 3.2 years<br>(0 to 15.8) | Mean: 73.5<br>(65 to 89) | USA |
| Luger 2020 | 178 | 91 (51) | 87 (49) | 0: 133 (75) | > 0: 45 (25) | 49.6<br>(IQR 32.9 to 66.3) | 64.6 years (IQR 50.8 to 72.7) | Austria |
| McGuire 1995 | 458 | 305 (67) | 153 (33) | All sub-optimal | 1 to 2 cm:<br>85 (18.6) | > 2 cm:<br>373 (81.4) | Not reported | USA |
| Melamed 2017a | 307 | 241<br>(78) | 66<br>(22) | 0: 141 (59)<br>< 1: 77 (32) | > 1: 23 (9)<br>n = 66 missing | 34.1 | < 60: 200 (65)<br>> 60: 107 (35) | USA |
| Melamed 2017b | 6013 | 4954 | 1506 | 0: 2048 (46) | > 1: 546 (12) | | < 60: 2803 (47) | |

**Table 1. Summary of stage and residual disease in included upfront primary debulking surgery (PDS) studies** *(Continued)*

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | (77) | (23) | < 1: 1848 (42) | 1571 missing | | > 60: 3210 (53) | |
| Paik 2018 | 419 | 370 (88) | 49 (12) | 0: 107 (26)<br>< 1: 147 (35) | > 1: 165 (39) | 43 (3 to 164) | Mean = 54.5 (SD 10.3) | South Korea |
| Peiretti 2010 | 259 | 199 (76) | 60 (24) | 0: 115 (44)<br>< 1: 83 (32) | 1 to 2: 18 (7)<br>> 2: 43 (17) | 29.8 | 58 (22 to 77) | Spain<br>Italy |
| Peiretti 2012 | 238 | 180 (76) | 58 (24) | 0: 99 (41)<br>< 1: 106 (44) | > 1: 32 (15) | Not reported | 59.7 (22 to 85) | Italy<br>USA |
| Polterauer 2012 | 226 | II: 15 (7)<br>III: 174 (77) | 37 (16) | 0: 157 (69) | > 0: 69 (31) | 25.0 (1 to 49) | Mean: 57.5 (SD 11.9) | Europe |
| Shim 2016 | 276 | III/IV (n = 276) | | Not reported | Not reported | Not reported | 54 (20 to 80) | South Korea |
| Tewari 2016 | 1718 | 1241 (72) | 477 (28) | 0: 85 (5)<br>< 1: 701 (41) | > 1: 932 (54) | Not reported | 58.5 to 60.2 for 0 to > 1 cm RD | USA |
| Tseng 2018 | 978 | 794 (81) | 184 (19) | 0: 408 (42)<br>< 1: 378 (39) | > 1: 192 (19) | 77.7 (1 to 198) | 61 (19 to 95) | USA |
| Van Geene 1996 | 219 | 180 (82) | 39 (18) | < 2 cm | < 2 cm: 92 (42) | > 2 cm: 127 (58) | 57 (24 to 75) | UK |
| Wimberger 2010 | 573 | 573 (100) | 0 (0) | 0: 70 (12)<br>< 1: 168 (29) | > 1: 335 (59) | Not reported | 59 (19 to 83) | Germany<br>France |
| Winter 2007 | 1895 | 1895 (100) | 0 (0) | 0: 437 (23)<br>< 1: 791 (42) | > 1: 667 (35) | 43 | 57 (16 to 86) | USA |
| Winter 2008 | 360 | 0 | 360 | 0 cm | 0 cm: 29 (8)<br>< 1 cm: 79 (22)<br>Total: 108 (30) | 28 | 59 (24 to 86) | USA |

IQR: interquartile range; RD: residual disease; SD: standard deviation

**Table 2. Summary of stage and residual disease in included interval debulking surgery (IDS) studies**

| Study | No. | Stage | | Optimal | Suboptimal | Median follow-up | Median age in years | Setting |
|---|---|---|---|---|---|---|---|---|
| | | III n (%) | IV n (%) | n (%) | n (%) | Months | (Range) | |
| Cioffi 2018 | 102 | 64 (63) | 38 (37) | 0: 37 (44)<br>< 1: 20 (23)† | ≥ 1: 28 (33)† | Not reported | Mean age<br>≥ 70 years: 74.5 (41%)<br>< 70 years: 58.3 (59%) | Italy |
| Davidson 2019 | 282 | IIIC: 114 (40)<br>IV: 101 (36)<br>Assumed AOC: 57 (20)<br>Unknown: 10 (4) | | 0: 165 (59)‡<br>≤ 1: 63 (22)‡ | > 1 to 2: 6 (2)‡<br>> 2: 37 (13)‡ | Not reported | 63.9 (34.1 to 84.8) | USA |
| Iwase 2015 | 124 | IIIB: 6 (5)<br>IIIC: 77 (62) | 41 (33) | < 1: 113 (91) | ≥ 1: 11 (9) | 39.5 (5 to 142) | 58 (29 to 83) | Japan |
| Kaban 2017 | 203 | Not reported | | ≤ 1: 165 (81)§ | > 1: 36 (19)§ | 34.5 (1 to 124) | 59 (28 to 84) | Turkey |
| Lecuru 2019 | 188 | Not reported | | Not reported | | 42.6 | Not reported | France |
| Lorusso 2016 | 193 | Not reported | | Not reported | | Not reported | Not reported | Italy |
| Petrillo 2014 | 322 | 251 (78) | 72 (22) | No definition of optimal given<br>0: 236 (73)<br>≤ 1: 36 (11)<br>> 1: 50 (16) | | 47 (3 to 181) | ≤ 65: 226 (70%)<br>> 65: 96 (30%) | Italy |
| Phillips 2018 | 398 | 273 (69) | 123 (31) | 0: 255 (64)<br>< 1: 55 (14) | ≥ 1: 88 (22) | Not reported | Mean: 63.9<br>(95% CI 42.2 to 85.6) | UK |
| Stoeckle 2014 | 118 | 82 (69) | 36 (31) | 0: 80 (68) | ≥ 1: 7 (6) | 37 | 64 (37 to 88) | France |

**Table 2. Summary of stage and residual disease in included interval debulking surgery (IDS) studies** *(Continued)*

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | < 1: 31 (26) | | | | |
| Zhang 2018 | 200 | 169 (85) | 31 (15) | 0: 59 (30)<br>< 1: 38 (19) | 1 to 2: 8 (4)<br>> 2: 30 (15) | 43.5 (IQR 38.5 to 56.2) | 61 (38 to 80) | China |
| Zhu 2016 | 672 | 564 (84) | 108 (16) | ≤ 1: 486 (72) | > 1: 186 (28) | 38 (5 to 103) | 55 (30 to 70) | China |

†85/102 participants underwent debulking surgery following neoadjuvant chemotherapy.
‡Residual disease data available for n = 271/282.
§Residual disease data available for n = 201/203.
**AOC:** advanced ovarian cancer; **CI:** confidence interval; **IQR:** interquartile range

**Table 3. Risk of bias assessment according to QUIPS (Quality in Prognostic Studies) for overall survival (OS) in primary debulking surgery (PDS) studies**

| Study | Study participation | Study attrition | Prognostic factor measurement | Outcome measurement | Adjustment for other prognostic factors | Statistical analysis and reporting |
|---|---|---|---|---|---|---|
| Akahira 2001 | Low | Unclear | Unclear | Low | High | Unclear |
| Aletti 2006 | Low | Unclear | Low | High | Low | Unclear |
| Ataseven 2016 | Low | Unclear | Low | Low | Low | High |
| Bristow 2011 | Low | Unclear | Low | Low | High | High |
| Chan 2003 | Low | Unclear | Unclear | Low | Low | High |
| Chang 2012a | Low | Unclear | Low | Low | Low | High |
| Chang 2012b | Low | Unclear | Low | Low | Unclear | High |
| Chi 2001 | Low | Unclear | Low | Low | Low | High |
| Chi 2006 | Low | Unclear | Low | Low | Low | High |
| Cuylan 2018 | Low | Unclear | Unclear | Low | Low | High |
| Eisenkop 2003 | Low | Unclear | Low | Low | High | High |
| Feng 2016 | Low | Unclear | Low | Low | Unclear | High |
| Hofstetter 2013 | Unclear | Unclear | Low | Low | Unclear | High |
| Kahl 2017 | Low | Unclear | Unclear | Low | Unclear | High |
| Klar 2016 | Low | Unclear | Unclear | Low | Unclear | High |
| Langstraat 2011 | Low | Unclear | Low | Low | Unclear | High |
| Luger 2020 | Low | Unclear | Low | Low | Low | High |
| McGuire 1995 | Low | Unclear | Low | Low | Unclear | High |
| Melamed 2017a | Low | Unclear | Low | Low | High | Unclear |
| Melamed 2017b | Low | Unclear | Low | Low | High | Unclear |
| Paik 2018 | Low | Unclear | Low | Low | Unclear | High |
| Peiretti 2012 | Low | Unclear | Unclear | Low | Low | High |
| Petrillo 2014 | Low | Unclear | Low | Low | High | High |
| Polterauer 2012 | Low | Unclear | Unclear | Low | Low | Unclear |

**Table 3. Risk of bias assessment according to QUIPS (Quality in Prognostic Studies) for overall survival (OS) in primary debulking surgery (PDS) studies** *(Continued)*

| | | | | | | |
|---|---|---|---|---|---|---|
| Tewari 2016 | Low | Unclear | Low | Low | Low | High |
| Tseng 2018 | Low | Unclear | Low | Low | Low | High |
| Van Geene 1996 | Unclear | Unclear | Unclear | Unclear | Unclear | Unclear |
| Wimberger 2010 | Low | Unclear | Low | Low | Unclear | High |
| Winter 2007 | Low | Unclear | Low | Low | Unclear | Unclear |
| Winter 2008 | Low | Unclear | Low | Low | Unclear | Unclear |

**Table 4. Risk of bias assessment according to QUIPS (Quality in Prognostic Studies) for overall survival (OS) in interval debulking surgery (IDS) studies**

| Study | Study participation | Study attrition | Prognostic factor measurement | Outcome measurement | Adjustment for other prognostic factors | Statistical analysis and reporting |
|---|---|---|---|---|---|---|
| Cioffi 2018 | Low | Unclear | Low | Low | Low | Unclear |
| Davidson 2019 | Low | Unclear | Unclear | High | High | Unclear |
| Iwase 2015 | Unclear | Unclear | Low | Low | Low | High |
| Kaban 2017 | Unclear | Unclear | Low | Low | Unclear | High |
| Lecointre 2020 | Low | Unclear | Unclear | Low | High | Unclear |
| Lecuru 2019 | High | Unclear | Low | Low | High | High |
| Liu 2020 | Low | Unclear | Low | Low | High | High |
| Lorusso 2016 | High | Unclear | Low | Low | High | High |
| Petrillo 2014 | Low | Unclear | Low | Low | High | High |
| Phillips 2018 | Low | Unclear | Low | Low | High | High |
| Stoeckle 2014 | Low | Unclear | Low | Low | Unclear | Unclear |
| Zhang 2018 | Low | Unclear | Low | Low | Unclear | Unclear |
| Zhu 2016 | Low | Unclear | Unclear | Low | High | High |

**Table 5. Risk of bias assessment according to QUIPS (Quality in Prognostic Studies) for progression-free survival (PFS) in primary debulking surgery (PDS) studies**

| Study | Study participation | Study attrition | Prognostic factor measurement | Outcome measurement | Adjustment for other prognostic factors | Statistical analysis and reporting |
|---|---|---|---|---|---|---|
| Chang 2012a | Low | Unclear | Low | Unclear | Low | High |
| Chang 2012b | Low | Unclear | Low | Unclear | Unclear | High |
| Cuylan 2018 | Low | Unclear | Unclear | Unclear | Low | High |
| Feng 2016 | Low | Unclear | Low | Unclear | Unclear | High |
| Klar 2016 | Low | Unclear | Unclear | Unclear | Unclear | High |
| Luger 2020 | Low | Unclear | Low | Unclear | Low | High |
| McGuire 1995 | Low | Unclear | Low | Unclear | Unclear | High |
| Paik 2018 | Low | Unclear | Low | Unclear | Unclear | High |
| Peiretti 2010 | Low | Unclear | Low | Unclear | High | High |
| Polterauer 2012 | Low | Unclear | Unclear | Unclear | Low | Unclear |
| Shim 2016 | High | Unclear | Low | Unclear | High | High |
| Tewari 2016 | Low | Unclear | Low | Unclear | Low | High |
| Tseng 2018 | Low | Unclear | Low | Unclear | Low | High |
| Wimberger 2010 | Low | Unclear | Low | Unclear | Unclear | High |
| Winter 2007 | Low | Unclear | Low | Unclear | Unclear | Unclear |
| Winter 2008 | Low | Unclear | Low | Unclear | Unclear | Unclear |

**Table 6. Risk of bias assessment according to QUIPS (Quality in Prognostic Studies) for progression-free survival (PFS) in interval debulking surgery (IDS) studies**

| Study | Study participation | Study attrition | Prognostic factor measurement | Outcome measurement | Adjustment for other prognostic factors | Statistical analysis and reporting |
|---|---|---|---|---|---|---|
| Bixel 2020 | Low | Unclear | Unclear | Unclear | High | High |
| Cioffi 2018 | Low | Unclear | Low | Unclear | Low | Unclear |
| Lecointre 2020 | Low | Unclear | Unclear | Unclear | High | Unclear |

**Table 6. Risk of bias assessment according to QUIPS (Quality in Prognostic Studies) for progression-free survival (PFS) in interval debulking surgery (IDS) studies** *(Continued)*

| | | | | | | |
|---|---|---|---|---|---|---|
| Lecuru 2019 | High | Unclear | Low | Unclear | High | High |
| Liu 2020 | Low | Unclear | Low | Unclear | High | High |
| Petrillo 2014 | Low | Unclear | Low | Unclear | High | High |
| Shibutani 2020 | Low | Unclear | Low | Unclear | Low | High |
| Zhang 2018 | Low | Unclear | Low | Unclear | Unclear | Unclear |
| Zhu 2016 | Low | Unclear | Unclear | Unclear | High | High |

## APPENDICES

### Appendix 1. MEDLINE search strategy

1. exp Ovarian Neoplasms/
2. (ovar* adj5 cancer*).mp.
3. (ovar* adj5 neoplas*).mp.
4. (ovar* adj5 carcinom*).mp.
5. (ovar* adj5 malignan*).mp.
6. (ovar* adj5 tumor*).mp.
7. (ovar* adj5 tumour*).mp.
8. 1 or 2 or 3 or 4 or 5 or 6 or 7
9. exp Surgical Procedures, Operative/
10. surg*.mp.
11. "surgery".fs.
12. 9 or 10 or 11
13. debulk*.mp.
14. cytoreduc*.mp.
15. 13 or 14
16. 8 and 12 and 15
17. "randomized controlled trial".pt.
18. "controlled clinical trial".pt.
19. randomized.ab.
20. randomly.ab.
21. trial.ab.
22. groups.ab.
23. exp Cohort Studies/
24. cohort*.mp.
25. (case adj series).mp.
26. 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25
27. 16 and 26
28. Animals/
29. Humans/
30. 28 not (28 and 29)
31. 27 not 30

## Appendix 2. EMBASE search strategy

1. exp Ovary Tumor/
2. (ovar* adj5 (cancer* or neoplas* or carcinom* or malignan* or tumor* or tumour*)).mp.
3. 1 or 2
4. exp Surgery/
5. surg*.mp.
6. su.fs.
7. 4 or 5 or 6
8. (debulk* or cytoreduc*).mp.
9. 3 and 7 and 8
10. exp Controlled Clinical Trial/
11. crossover procedure/
12. double-blind procedure/
13. randomized controlled trial/
14. single-blind procedure/
15. random*.mp.
16. factorial*.mp.
17. (crossover* or cross over* or cross-over*).mp.
18. placebo*.mp.
19. (double* adj blind*).mp.
20. (singl* adj blind*).mp.
21. assign*.mp.
22. allocat*.mp.
23. volunteer*.mp.
24. exp cohort analysis/
25. cohort*.mp.
26. series.mp.
27. 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26
28. 9 and 27

key: mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer name, fs=floating subheading

## Appendix 3. CENTRAL search strategy

1. MeSH descriptor Ovarian Neoplasms explode all trees
2. ovar* near/5 cancer*
3. ovar* near/5 neoplas*
4. ovar* near/5 carcinom*
5. ovar* near/5 malignan*
6. ovar* near/5 tumor*
7. ovar* near/5 tumour*
8. (#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7)
9. MeSH descriptor Surgical Procedures, Operative explode all trees
10. surg*
11. Any MeSH descriptor with qualifier: SU
12. (#9 OR #10 OR #11)
13. debulk*
14. cytoreduc*
15. (#13 OR #14)
16. (#8 AND #12 AND #15)

## Appendix 4. Risk of bias and applicability assessment

Risk of bias and applicability assessment tool to assess risk of bias and applicability of prognostic factor studies (Riley 2019). Signalling questions and risk of bias ratings are listed in bullet points.

Domain 1: Participant selection

Risk of bias:

- Adequate participation in the study by eligible persons
- Description of the target population or population of interest
- Description of the baseline study sample
- Adequate description of the sampling frame and recruitment
- Adequate description of the period and place of recruitment
- Adequate description of inclusion and exclusion criteria

Risk of bias ratings:

- High: the relationship between the PF and outcome is very likely to be different for participants and eligible non-participants
- Moderate: the relationship between the PF and outcome may be different for participants and eligible non-participants
- Low: the relationship between the PF and outcome is unlikely to be different for participants and eligible non-participants

Applicability:

Are there concerns that the included women do not match the review question?

Domain 2: Study attrition

Risk of bias:

- Adequate response rate for study participants
- Description of attempts to collect information on participants who dropped out
- Reasons for loss to follow-up are provided
- Adequate description of participants lost to follow-up
- There are no important differences between participants who completed the study and those who did not

Risk of bias ratings:

- High: the relationship between the PF and outcome is very likely to be different for completing and non-completing participants
- Moderate: the relationship between the PF and outcome may be different for completing and non-completing participants
- Low: the relationship between the PF and outcome is unlikely to be different for completing and non-completing participants

Domain 3: Prognostic factor measurement

Risk of bias:

- A clear definition or description of the PF is provided
- Method of PF measurement is adequately valid and reliable
- Continuous variables are reported or appropriate cutpoints are used
- The method and setting of measurement of PF is the same for all study participants
- Adequate proportion of the study sample has complete data for the PF
- Appropriate methods of imputation are used for missing PF data

Risk of bias ratings:

- High: the measurement of the PF is very likely to be different for different levels of the outcome of interest
- Moderate: the measurement of the PF may be different for different levels of the outcome of interest
- Low: the measurement of the PF is unlikely to be different for different levels of the outcome of interest

Applicability:

Are there concerns that residual disease, the way that it is measured, or the way that it is interpreted, differ from the review question?

Domain 4: Outcome measurement

Risk of bias:

- A clear definition of the outcome is provided
- Method of outcome measurement used is adequately valid and reliable
- The method and setting of outcome measurement is the same for all study participants

Risk of bias ratings:

- High: the measurement of the outcome is very likely to be different related to the baseline level of the PF
- Moderate: the measurement of the outcome may be different related to the baseline level of the PF
- Low: the measurement of the outcome is unlikely to be different related to the baseline level of the PF

Applicability:

Are there concerns that outcome does not match the review question or that follow-up was not of sufficient duration?

Domain 5: Adjustment for other prognostic factors

Risk of bias:

- All other important PFs are measured
- Clear definitions of the important PFs measured are provided
- Measurement of all important PFs is adequately valid and reliable
- The method and setting of PF measurement are the same for all study participants
- Appropriate methods are used to deal with missing values of PFs, such as multiple imputation
- Important PFs are accounted for in the study design
- Important PFs are accounted for in the analysis

Applicability:

Did the prognostic factors adjusted for match the review question?

Risk of bias ratings:

- High: the observed effect of the PF on the outcome is very likely to be distorted by another factor related to PF and outcome
- Moderate: the observed effect of the PF on outcome may be distorted by another factor related to PF and outcome
- Low: the observed effect of the PF on outcome is unlikely to be distorted by another factor related to PF and outcome

Domain 6: Statistical analysis and reporting

Risk of bias:

- Sufficient presentation of data to assess the adequacy of the analytic strategy
- Strategy for model building is appropriate and is based on a conceptual framework or model
- The selected statistical model is adequate for the design of the study
- There is no selective reporting of results

Risk of bias ratings:

- High: the reported results are very likely to be spurious or biased related to analysis or reporting
- Moderate: the reported results may be spurious or biased related to analysis or reporting
- Low: the reported results are unlikely to be spurious or biased related to analysis or reporting

## Appendix 5. Domains to be considered when judging the strength of the body of evidence

We considered the following domains when we assessed the strength of the body of evidence, based on the GRADE approach (Guyatt 2008):

- Risk of bias: Based on the results of the risk of bias assessments, we downgraded confidence in the evidence base if most evidence was from studies that we judged to be at high risk of bias.
- Indirectness: We downgraded confidence in the evidence base if we had concerns that the study sample, the prognostic factor, the outcome and/or the other factors in the models in the primary studies did not reflect the review question.

- Inconsistency: We downgraded confidence in the evidence base if there was unexplained heterogeneity or variability in results across studies.
- Imprecision: We downgraded confidence in the evidence base if the estimate of the effect size from a meta-analysis was not precise or, if no meta-analysis was performed, if the estimate of the size of effect from individual studies was not precise.
- Publication bias: Studies showing no association are likely to be unpublished, unless part of a larger study that specifically aimed to compare tests. We downgraded our confidence in the evidence base if we had reason to suspect publication bias from our assessments of reporting bias.
- Size of effect: We upgraded our confidence in the evidence base if the size of effect was moderate or large. If a meta-analysis was not possible, we upgraded if the size of effect was moderate or large for most included studies.

## Appendix 6. Factors included in multivariate analysis in upfront primary debulking (PDS) studies

| Citation | Factors included in multivariable (multivariate) analysis |
|---|---|
| Akahira 2001 | Residual disease, histology and performance status |
| Aletti 2006 | Residual disease, age, American Society of Anesthesiology (ASA) score, histological grade, operative time and aggressive surgery |
| Ataseven 2016 | Age, performance status, residual tumour, tumour stage and ascites |
| Bristow 2011 | Race, tumour grade 3, non-serous histology, ASA score > 3, surgical complexity score, serum albumin < 3.0 g/dL, platinum-based therapy, residual disease and perioperative morbidity |
| Chan 2003 | Residual disease, age (older versus younger), stage (IV versus III) and performance status (1 to 2 versus 0) |
| Chang 2012a | Stage (IV), surgical procedure, residual disease and age |
| Chang 2012b | Residual disease, type of surgery, performance of lymphadenectomy and age |
| Chi 2001 | Residual disease, age, stage (IIIC and IV versus IIIA/IIIB) and ascites (yes versus no) |
| Chi 2006 | Residual disease, age and ascites |
| Cuylan 2018 | Age, maximal cytoreduction and stage |
| Eisenkop 2003 | Residual disease and sum of rankings |
| Feng 2016 | Age, FIGO stage, residual disease and TTC |
| Hofstetter 2013 | Interval from surgery to start of chemotherapy (≤ 28 versus < 28 days), stage (III versus IV), residual disease, age and extent of surgery |
| Kahl 2017 | ACCI, ECOG PS, FIGO stage, surgical complexity score, blood loss, residual disease and duration of surgery |
| Klar 2016 | Age, ECOG, BMI, stage, grading, residual tumour and histology |
| Langstraat 2011 | Creatinine > 1.2 mg/dL, surgical complexity score, residual disease, stage IV disease and age |
| Luger 2020 | Age (cat), CA-125, paraaortic nodes, FIGO, cardiophrenic lymph nodes dimension, residual disease |
| McGuire 1995 | Residual disease, age, GOG performance status, histological subtype, stage or residual disease and measurable disease |

(Continued)

| | |
|---|---|
| Melamed 2017a | Age, race/ethnicity, stage, region, insurance status, treating facility type, hospital annual ovarian cancer volume, residual disease and presence of comorbidities |
| Melamed 2017b | |
| Paik 2018 | Age, CA-125 level (U/mL), FIGO stage, residual disease and normal-sized ovary |
| Peiretti 2010 | Age, stage IV vs IIIC and any residual disease |
| Peiretti 2012 | Age, stage, histology, grade, presence of ascites and residual tumour at end of surgery |
| Polterauer 2012 | Tumour stage, residual tumour, histological grade, histological type and age |
| Shim 2016 | Not reported (abstract) |
| Tewari 2016 | Age, race/ethnicity, performance status, grade, stage, histology,ascites, CA 125 (µg/ml), tumour residual and time from surgery to initiation of chemotherapy |
| Tseng 2018 | Age, albumin, stage, ASA score, histology, BRCA status, OR Tumour Index, residual disease and postop IP chemotherapy |
| Van Geene 1996 | Residual disease, performance status and pattern of spread |
| Wimberger 2010 | Age, performance status, histology, residual tumour size, peritoneal carcinomatosis and stage IV disease site |
| Winter 2007 | Residual disease, age (discrete), race, GOG performance status, histology and tumour grade |
| Winter 2008 | Residual disease, histology and stage IV disease site |

ACCI: age-adjusted Charlson Comorbidity Index; ASA: American Society of Anaesthesiologists; BMI: body mass index; BRCA: breast cancer gene; ECOG: Eastern Cooperative Oncology Group; FIGO: International Federation of Gynecology and Obstetrics; GOG: Gynaecologic Oncology Group; IP: intraperitoneal; PS: performance score; TTC: time to chemotherapy

## Appendix 7. Factors included in multivariate analysis for each study on neoadjuvant chemotherapy (NACT) and interval debulking surgery (IDS)

| Citation | Factors included in multivariable (multivariate) analysis |
|---|---|
| Bixel 2020 | Residual disease, NACT cycles, route of chemotherapy administration (intraperitoneal or intravenous), maintenance therapy (yes/no) |
| Cioffi 2018 | Residual disease, age, number of neoadjuvant chemotherapy courses, debulking surgery, ASA score, hypoalbuminaemia (defined as albuminaemia < 32 g/L), FIGO stage, presence of ascites ≥ 500 mL, high tumour dissemination and Charlson comorbidity score |
| Davidson 2019 | Residual disease, ASA score, age, SCS and major morbidity |
| Iwase 2015 | Residual disease, FIGO stage, histological subtype, NACT cycles, NACT regimen, systematic lymphadenectomy, excision of other organ(s), ascites cytology, lymph node metastasis |
| Kaban 2017 | Residual disease, age, lymphadenectomy, macroscopic tumour in omentum, number of chemotherapy cycles |

(Continued)

| | |
|---|---|
| Lecointre 2020 | Residual disease, number of NACT cycles (≤4, > 4), age (cat), Charlson index, FIGO, lymph node status (N+ vs N0), response to NACT |
| Lecuru 2019 | Complete cytoreduction, ECOG, ascites, neutrophil/lymphocyte ratio, PCI at baseline, RECIST ORR (response rate at end of NACT), pCR and treatment arm (nintedanib vs placebo) |
| Liu 2020 | Residual disease, age (cont) |
| Lorusso 2016 | Residual disease, ECOG and number of NACT cycles† |
| Petrillo 2014 | Residual disease, age, carcinomatosis at diagnosis, CA-125, pathological response to NACT |
| Phillips 2018 | Residual disease, FIGO stage, chemotherapy regime (carbo/taxol vs carboplatin) |
| Shibutani 2020 | Residual disease, age (cat), performance status, FIGO, disease type, histology, NACT cycles, NACT regimens |
| Stoeckle 2014 | Residual disease, tumour grade, WHO performance status, ASA, bowel surgery (yes/no), FIGO stage |
| Zhang 2018 | Residual disease, Pre-operative ascites, number of tumour sites, number of NAC cycles, CA-125 at diagnosis, CA-125 decreasing kinetics |
| Zhu 2016 | Residual disease, FIGO stage, chemosensitivity, Glasgow prognostic score |

†Full list of variables in multivariate analysis not explicitly mentioned.

ECOG: Eastern Cooperative Oncology Group; FIGO: International Federation of Gynecology and Obstetrics; NACT: neoadjuvant chemotherapy; PCI: Peritoneal Cancer Index; WHO: World Health Organization

## HISTORY

Protocol first published: Issue 9, 2021

## CONTRIBUTIONS OF AUTHORS

AE, BWR, KG and RN drafted the clinical and discussion sections of the review; AB, SH and PK data extracted items for inclusion in the review; AB drafted the methodological, results and discussion sections of the review. DC and LV reconciled the methodological and results sections of the review and contributed to the discussion. All authors agreed the final version.

## DECLARATIONS OF INTEREST

- Andrew Bryant: none known
- Ahmed Elattar: none known
- Patience Kunonga: none known
- Brett A Winter-Roach: none known
- Shaun Hiu: none known
- Dawn Craig: none known
- Luke Vale: none known
- Ketankumar Gajjar: none known
- Raj Naik: none known

## SOURCES OF SUPPORT

### Internal sources

- None, Other

### External sources

- National Institute for Health Research (NIHR), via Cochrane infrastructure funding to Cochrane Gynaecological, Neuro-oncology and Orphan Cancers, UK

## DIFFERENCES BETWEEN PROTOCOL AND REVIEW

Three studies included a small proportion of women with early-stage (predominantly stage II) or unknown disease. Although not stringently part of our initial inclusion criteria, we included a study if the proportion with unknown or early-stage disease in the entire cohort was small. The proportion of women with early or unknown stage of disease in Feng 2016 (9.3%), Polterauer 2012 (6.6%) and Klar 2016 (12.5%) was not going to affect the applicability of the results.

The definitions of RD < 1 cm and RD > 1 cm were changed from near-optimal and suboptimal in the published protocol to small-volume residual disease (SVRD) and large-volume residual disease (LVRD), respectively. It was felt that this would make it easier to read for the non-clinical reader, as a combination of numbers and letters is more challenging and Cochrane Reviews have a large lay audience.

## INDEX TERMS

### Medical Subject Headings (MeSH)

Carcinoma, Ovarian Epithelial [drug therapy] [surgery]; Chemotherapy, Adjuvant [methods]; *Clinical Decision-Making; Neoadjuvant Therapy [methods]; Neoplasm, Residual; *Ovarian Neoplasms [drug therapy] [pathology] [surgery]; Prognosis; Uncertainty

### MeSH check words

Female; Humans