Published in final edited form as:

Cancer Epidemiol. 2018 August ; 55: 8–16. doi:10.1016/j.canep.2018.05.003.

Review of methodological challenges in comparing the effectiveness of neoadjuvant chemotherapy versus primary debulking surgery for advanced ovarian cancer in the United States

Ashley L. Cole, MPH^a, Anna E. Austin, MPH^b, Ryan P. Hickson, PharmD, MPH^a, Matthew S. Dixon, PharmD^a, and Emma L. Barber, MD, MS^c

^aDivision of Pharmaceutical Outcomes and Policy, UNC Eshelman School of Pharmacy, University of North Carolina, Chapel Hill, North Carolina, 27599 USA

^bDepartment of Maternal and Child Health, UNC Gillings School of Global Public Health, University of North Carolina, Chapel Hill, North Carolina, 27599 USA

^cDivision of Gynecologic Oncology, Department of Obstetrics and Gynecology, Northwestern University, Chicago, IL, 60611 USA

Abstract

Randomized trials outside the U.S. have found non-inferior survival for neoadjuvant chemotherapy (NACT) versus primary debulking surgery (PDS) for advanced ovarian cancer (AOC). However, these trials reported lower overall survival and lower rates of optimal debulking than U.S. studies, leading to questions about generalizability to U.S. practice, where aggressive debulking is more common. Consequently, comparative effectiveness in the U.S. remains controversial. We reviewed U.S. comparative effectiveness studies of NACT versus PDS for AOC. Here we describe methodological challenges, compare results to trials outside the U.S., and make suggestions for future research. We identified U.S. studies published in 2010 or later that evaluated the comparative effectiveness of NACT versus PDS on survival in AOC through a PubMed search. Two independent reviewers abstracted data from eligible articles. Nine of 230 articles were eligible for review. Methodological challenges included unmeasured confounders, heterogeneous treatment effects, treatment variations over time, and inconsistent measurement of treatment and survival. Whereas some limitations were unavoidable, several limitations noted across studies were avoidable, including conditioning on mediating factors and immortal time introduced by measuring survival beginning from diagnosis. Without trials in the U.S., non-randomized studies are an important source of evidence for the ideal treatment for AOC. However, several methodological challenges exist when assessing the comparative effectiveness of NACT versus PDS in a non-randomized setting. Future observational studies must ensure that treatment is consistent throughout the study period and that treatment groups are comparable. Rapidly-evolving oncology data networks may allow for identification of treatment intent and other important confounders.

Keywords

neoadjuvant therapy; ovarian neoplasms; comparative effectiveness research; observational study; epidemiology; review; cytoreduction surgical procedures

Introduction

Ovarian cancer is the leading cause of death among gynecologic malignancies, chiefly due to the large proportion of cases diagnosed at advanced stage.^{1–3} Optimal debulking (<1cm residual disease) is an important predictor of survival for patients with advanced ovarian cancer, and an additional survival benefit is conferred for those with resection to no gross residual disease.^{4–14}

The standard treatment for advanced ovarian cancer in the U.S. has traditionally been primary debulking surgery (PDS) – which aims to remove the majority of disease upfront – followed by adjuvant chemotherapy. However, optimal debulking often requires aggressive surgery that is associated with substantial morbidity, including complications that can delay chemotherapy.^{11,15–17}

Neoadjuvant chemotherapy (NACT) followed by interval debulking surgery and adjuvant chemotherapy has become an increasingly accepted alternative to PDS for advanced ovarian cancer.^{18–20} NACT aims to reduce tumor burden and improve performance status preoperatively. This decreases the need for aggressive procedures, reduces post-operative complications, and increases the probability of optimal debulking.^{17,21,22} However, it is unclear whether NACT provides a survival benefit.

In the U.S., the role of NACT in advanced ovarian cancer remains controversial.^{23–29} Although two randomized controlled trials (RCTs) outside the U.S. observed non-inferior survival with NACT versus PDS,^{10,30} the generalizability of these trials to U.S. practice has been questioned due to lower median overall survival and lower rates of optimal debulking compared to U.S. populations.^{28,29} A less aggressive surgical effort compared to that commonly performed in the U.S. has been postulated as the cause of the observed non-inferiority results.^{27–29} Among U.S. gynecologic oncologists surveyed in 2010, 82% did not believe there was sufficient evidence to justify NACT;³¹ conversely, a European survey found that 70% of gynecologic oncologists believed there was sufficient evidence for use of NACT in advanced ovarian cancer.³²

In the absence of U.S. randomized studies on the effectiveness of NACT for advanced ovarian cancer, observational studies represent an important source of evidence. However, non-randomized treatment allocation presents challenges in obtaining unbiased estimates of effectiveness. In this article, we review studies comparing the effectiveness of NACT versus PDS in advanced ovarian cancer, emphasizing epidemiologic methods used, and discuss the strength of the U.S. evidence in the context of existing RCTs, with suggestions for future research.

Materials and Methods

We conducted a PubMed search on February 5th, 2017 using keywords selected based on prior knowledge and examination of previously-identified studies (Figure 1). Eligible studies met the following criteria: (1) compared the effectiveness of NACT and PDS on overall survival for advanced ovarian, fallopian, or primary peritoneal cancer; (2) conducted in the U.S.; (3) published in English, (4) published 2010 or later (reviews published in 2007 and 2011 outline evidence before 2010).^{33,34} We screened titles and abstracts identified through the electronic search for eligibility. The full text of potentially-eligible articles was then reviewed to confirm eligibility.

Each eligible article was reviewed by two independent reviewers. The primary reviewer performed an in-depth review, abstracting information regarding study aims, design, setting, population, methods, overall survival results, and methodological strengths and limitations into an abstraction form designed for the purposes of the study. A second reviewer performed an independent in-depth review, and recorded additional information and/or disagreements into separate fields within the abstraction form. Disagreements between primary and secondary reviewers were discussed and resolved by all reviewers (ALC, AEA, RPH, MSD). References from eligible articles were hand-searched to identify additional studies.

Results

Nine observational studies met eligibility criteria for review (Table 1).^{19,20,35–41} Data from 1991²⁰ through 2013⁴¹ were used to conduct these studies. Most studies included women with stage IIIC-IV ovarian cancer,^{19,35,37,39,41} one included only stage IV,³⁵ and others included earlier stage disease.^{20,38,40} Four studies included only NACT patients who received interval debulking surgery.^{35,37,39,41} Three studies^{20,36,40} used nationally-representative registry data, while the remaining studies used institutional medical record data.

All studies compared overall survival between patients treated with NACT versus PDS using hazard ratios and/or Kaplan-Meier survival estimates, except one study⁴⁰ that used incidence rate ratios (Table 1). Studies used various methods to control for confounding, including multivariable adjustment,^{20,35,39–41} propensity score methods,^{19,20,36} instrumental variable analysis,²⁰ stratification,^{19,36,37,40} and restriction.³⁶ Two studies did not report any confounder control when comparing overall survival.^{37,38} Though measured confounders were listed for most studies, covariates used for confounder control were unclear for some studies.^{20,35,39,41} Four studies conditioned on residual disease, a variable on the causal pathway and one that is not known preoperatively when treatment decisions are being made.^{19,35,39,41}

Some studies attempted to address unmeasured confounding (Table 2). Wright used hospital referral region as an instrumental variable.²⁰ Two studies performed sensitivity analyses to assess the impact of unmeasured confounders, including preoperative disease burden,^{19,36} performance status,³⁶ and breast cancer susceptibility gene (BRCA) status.³⁶ One study

restricted to patients <70 years old with Klabunde-Charlson comorbidity index of 0 to mitigate confounding by health status.³⁶

Four studies^{19,36,37,40} attempted to assess heterogeneous treatment effects (Tables 1–2). Six studies restricted their study populations based on age,^{20,36,38–40} body mass index,³⁷ or comorbidity index.³⁶ One study⁴⁰ stratified survival by a risk score based on age, comorbidity index, and stage. Three studies stratified survival estimates by stage.^{19,36,40}

Five studies had evidence of immortal time bias (Table 2). Of those, three^{20,35,36} introduced immortal time by including person-time from diagnosis to first treatment in the survival calculation, and two^{37,39} introduced immortal time by conditioning survival on events occurring after treatment began (i.e., interval debulking surgery). One study did not provide enough details on survival measurement to make a determination.³⁸ Another study measured survival from interval debulking surgery.⁴¹ No studies reported the length of time between diagnosis and treatment, making the direction and magnitude of immortal time bias difficult to quantify.

Discussion

To our knowledge, this is the first review of comparative effectiveness studies of NACT versus PDS for advanced ovarian cancer focusing on methodological strengths and limitations. We identified several challenges associated with assessing effectiveness in a non-randomized setting. Here we focus on the most relevant and frequently encountered issues, including unmeasured confounders, heterogeneous treatment effects, treatment variation over time, and inconsistent measurement of treatment and survival.

Challenges of assessing comparative effectiveness in an observational setting

Obtaining causal estimates of the effectiveness of NACT versus PDS requires an assumption of exchangeability—that other than the treatment received, those receiving NACT have the same underlying risk of death as those treated with PDS.⁴² In an RCT, this is achieved through treatment randomization; however, in an observational setting, treatment decisions are often based on prognostic factors, resulting in confounding by indication. In this setting, the assumption of exchangeability can still be achieved through conditioning on confounders (e.g., adjustment, restriction, matching) but only if factors influencing treatment decisions and survival are accurately measured and modeled.

The challenge in adequately controlling for confounding by indication is that factors influencing treatment decisions, such as presumed unresectability and poor health/functional status, are often difficult to quantify. Because these factors are more common among patients receiving NACT compared to PDS and are also associated with reduced survival, inadequate confounding control is likely to bias estimates, making survival appear inferior for NACT, compared to PDS (Figure 2a). If the necessary data are available, sensitivity analyses can address unmeasured confounding by approximating the impact of an unmeasured confounder on the effect estimate. This method was used in studies by Meyer¹⁹ and Rauh-Hain.³⁶

Careful consideration of relationships between covariates is crucial, since differences in the adjustment set can change the interpretation of effect estimates. Several reviewed studies conditioned on factors on the causal pathway from treatment to survival (mediators), such as residual disease and surgical complexity (Figure 2b).^{35,39,41} Although it is useful in clinical practice to understand survival differences between these subgroups, conditioning on mediators when assessing effectiveness removes some of the total treatment effect.⁴³ For example, Gill compared patients who received NACT for presumed unresectable disease to matched patients who received suboptimal PDS.⁴¹ Although this study design attempts to identify exchangeable treatment groups by comparing NACT to PDS among patients with unresectable disease, conditioning on a suboptimal outcome in the PDS group makes survival appear inferior for PDS. Additionally, mediators are influenced by prior treatment. For example, NACT patients with no residual disease after debulking likely represent patients with greater tumor burden at diagnosis than PDS patients with no residual disease after debulking; conditioning on residual disease following surgery makes survival appear inferior for NACT.⁴⁴

Heterogeneous treatment effects are an important consideration since the ideal population for NACT is unclear. Potential effect modifiers may be closely related to factors responsible for confounding by indication if suspected heterogeneous treatment effects influences treatment decisions. This makes heterogeneous treatment effects difficult to quantify since it increases the likelihood of non-positivity, where there is a zero probability of receiving one of the two treatment in some subgroups (e.g., young, healthy patients with presumed resectable disease receiving NACT or older patients with high disease burden receiving PDS).⁴²

The changing role of NACT in advanced ovarian cancer over the past 10-15 years presents several challenges for studying the effectiveness of NACT. Historically, NACT was reserved for patients with poor health status or the most severe disease, resulting in substantial channeling bias.⁴⁵ In recent years, particularly after results from RCTs were reported,^{10,46} treatment selection may be less pronounced, and greater clinical equipoise may make comparisons of NACT and PDS in observational studies less subject to this bias.

In addition to changing beliefs about NACT, treatment itself has changed over time in ways that may violate a key assumption required for causality: that variations in treatment do not differentially affect outcomes.⁴² A clear example is the use of carboplatin and paclitaxel, which became the standard chemotherapeutic regimen for advanced ovarian cancer in 2003.⁴⁷ Five studies using data from the 1990s are therefore studying two different interventions within the NACT group.^{20,35,38–40} A less quantifiable example is the shift toward more aggressive surgery in the early- to mid-2000s, as the prognostic value of complete cytoreduction became established.^{4,12–14} For these reasons, restriction to later years is likely necessary to obtain valid estimates of effectiveness for NACT in a contemporary U.S. setting.

Definitions of treatment and survival can introduce biases, such as immortal time (Figure 3). Immortal time occurs when the study design imposes a period during which patients must survive in order to be included in the study and inappropriately attributes this survival time

to treatment, thus overestimating survival in one or both groups.⁴⁸ Given the short overall survival in advanced ovarian cancer and small survival differences between NACT and PDS, even a short period of immortal time could alter conclusions about relative effectiveness.

Without information on the intended treatment, classifying NACT using only initial chemotherapy results in potential misclassification of patients receiving palliative chemotherapy, thus underestimating survival in this group. Some studies attempted to address this issue by requiring NACT patients to receive interval debulking surgery, but this approach creates additional methodological problems. Immortal time introduced by measuring survival from initial diagnosis or treatment erroneously inflates survival estimates for NACT patients (Figure 3b-c), but measuring survival from interval debulking surgery underestimates survival time relative to PDS (Figure 3d). Additionally, conditioning on receipt of surgery for the NACT group inflates survival estimates; patients who were intended to be treated with interval debulking surgery but did not respond to adjuvant chemotherapy are excluded from the analysis. Additional problems arise when covariates are measured at interval debulking surgery, since covariates may be affected by prior treatment. For instance, ascites and pleural effusion that are alleviated by chemotherapy will be underestimated when measured at interval debulking surgery, resulting in inadequate confounding control that makes survival appear inferior for NACT.³⁹

Measuring survival from diagnosis introduces immortal time in both treatment groups (Figure 3a). The magnitude and direction of bias will depend on the relative length of time from diagnosis to initial treatment in each group. When time from diagnosis to treatment varies by treatment group, patients can be matched on time-to-treatment or a landmark analysis can be conducted. A landmark analysis defines a fixed period to classify treatment, with survival in both groups measured from the end of that period (Figure 3e). While this allows time to distinguish treatment intent among NACT patients, it also excludes patients who die during the high-risk period shortly after PDS, making PDS appear superior.

Current state of the evidence and future studies

Two RCTs completed outside the U.S. observed non-inferior survival with NACT versus PDS for advanced ovarian cancer.^{10,30} The main criticisms of these RCTs are lower than expected overall survival and lower rates of optimal debulking among PDS patients, compared to U.S. studies.²⁷⁻²⁹ Potential explanations for these findings include less aggressive surgical practices than U.S. settings and selective referral of patients with more extensive disease into trials, both of which create generalizability challenges for clinicians attempting to identify the ideal target population for NACT. These explanations also raise the possibility that PDS performed in an aggressive surgical setting among patients with lower disease burden may lead to a greater survival benefit than NACT.²⁷⁻²⁹

Because no RCTs have been performed in the U.S., it is important to weigh the potential biases associated with existing U.S. evidence with the limited generalizability of non-U.S. RCTs. Of the U.S. studies reviewed, those we judged to be methodologically strongest reported survival advantages with PDS of 5-10 months for stage IIIC disease and 1-5 months for stage IV disease, compared to NACT.^{19,36,40} However, these observational studies were unable to fully address biases introduced by unmeasured confounders^{19,35,40} and ambiguous

treatment intent,^{19,36} which would make survival appear inferior for NACT. Therefore, although non-U.S. RCT results may reflect lower overall survival than would be observed in an aggressive surgical setting, survival differences in U.S. observational studies are likely overestimated. These large differences may not be observed with truly exchangeable treatment groups.

To ensure that treatment groups are exchangeable, future observational studies must consider the appropriate study population in which to compare the effectiveness of NACT and PDS. Recently published treatment guidelines suggest consideration of NACT for patients with stage IIIC-IV disease for whom optimal debulking is unlikely;⁴⁹ however, in an earlier survey, only 29% of gynecologic oncologists reported that they could accurately make this determination preoperatively.³¹ Furthermore, providers' assessment of optimal resectability may vary by other factors, such as surgical skills. Use of an objective, pre-treatment measure of tumor burden, such as laparoscopic scoring methods, may help ensure exchangeability across treatment groups. However, though numerous methods have been developed and validated, none is universally accepted or consistently used across institutions in the U.S., which may make conditioning on this measure in retrospective studies challenging.^{18,50–56} Incorporation of a standard measure into guideline-concordant care may improve both treatment decisions and future studies on this topic.

Instrumental variable analysis is an ideal method for comparative effectiveness in the presence of unmeasured confounding by indication, since it estimates the treatment effect within the range of clinical equipoise (e.g., “borderline-resectable” disease).⁵⁷ This method uses the variation in treatment that is due to a factor (the instrument) otherwise unrelated to the outcome, circumventing the effect of prognostic factors on treatment choice, thus avoiding confounding bias (Figure 2c). Wright and colleagues used hospital referral region as an instrument to estimate the effectiveness of NACT among patients whose treatment choice results from institutional treatment preferences.²⁰ However, the untestable assumption that the instrument is not associated with the outcome, except through treatment choice, is difficult to support. It is plausible that survival may differ across hospital referral regions due to factors other than treatment choice. Additionally, because instrumental variable analysis estimates effectiveness for the subgroup of patients whose treatment choice is driven only by the instrument (i.e., not prognostic factors), the ability to define that subpopulation is key to understanding to whom results are generalizable. This is particularly problematic given the likelihood of heterogeneous treatment effects when comparing NACT and PDS in advanced ovarian cancer.

In the absence of a compelling instrumental variable, data on treatment intent is critical to avoiding misclassification of palliative chemotherapy as NACT. Rapidly evolving oncology data networks may increase the feasibility of identifying treatment intent while also obtaining an adequate sample size. Alternatively, information on recommendations for surgery available in existing registry data may be useful in differentiating chemotherapy with surgical versus palliative intent.⁵⁸ Although these data lack information on other important confounders—such as performance status and distribution of disease—they may be useful for informing sensitivity analyses to estimate the impact of misclassification of NACT.

Assessing the effectiveness of NACT in a non-randomized setting is challenging. However, because there are no ongoing U.S. RCTs, observational studies represent an important source of evidence. Many authors acknowledged the methodological challenges we discuss here, but were limited by the data available. However, other studies introduced avoidable biases — including conditioning on mediating factors and immortal time introduced by measuring survival beginning from diagnosis — with implications that were sometimes difficult to predict based on reported information. For this reason, we made no attempts to aggregate study results or make causal inferences about the comparative effectiveness of NACT versus PDS based on existing U.S. literature.

Currently, two RCTs are underway that may inform the ongoing debate on this topic. The Surgical Complications Related to Primary or Interval Debulking in Ovarian Neoplasms (SCOPRION) trial in Italy is using well-defined criteria for resectability and performing aggressive debulking more similar to the U.S.¹⁷ The Trial on Radical Upfront Surgical in Advanced Ovarian Cancer (TRUST) is also comparing NACT and PDS in an aggressive surgical setting.⁵⁹ Until results from these trials are available, we hope this review provides insight to help researchers design future non-experimental studies in this area.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Til Stürmer, MD, PhD for input and feedback on this project.

Abbreviations

NACT	neoadjuvant chemotherapy
PDS	primary debulking surgery
RCT	randomized controlled trial

References

- 1The role of the generalist obstetrician-gynecologist in the early detection of ovarian cancer. *Gynecologic oncology*. 2002; 87(3):237–239. [PubMed: 12492084]
- 2Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. *CA: a cancer journal for clinicians*. 2016; 66(1):7–30. [PubMed: 26742998]
- 3SEER Stat Fact Sheets: Ovarian Cancer <http://seer.cancer.gov/statfacts/html/ovary.html>. Accessed 20 Mar 2017
- 4Aletti GD, Dowdy SC, Gostout BS, et al. Aggressive surgical effort and improved survival in advanced-stage ovarian cancer. *Obstet Gynecol*. 2006; 107(1):77–85. [PubMed: 16394043]
- 5Winter WE 3rd, Maxwell GL, Tian C, et al. Prognostic factors for stage III epithelial ovarian cancer: a Gynecologic Oncology Group Study. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2007; 25(24):3621–3627. [PubMed: 17704411]
- 6Winter WE, Maxwell GL, Tian C, et al. Tumor residual after surgical cytoreduction in prediction of clinical outcome in stage IV epithelial ovarian cancer: a Gynecologic Oncology Group Study. *J Clin Oncol*. 2008; 26(1):83–89. [PubMed: 18025437]

- 7Bristow 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):278–287. [PubMed: 10053096]
- 8Bristow 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 : official journal of the American Society of Clinical Oncology*. 2002; 20(5):1248–1259. [PubMed: 11870167]
- 9Eisenkop SM, Spirtos NM, Friedman RL, Lin WC, Pisani AL, Peticucci 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):390–396. [PubMed: 12893206]
- 10Vergote I, Tropé CG, Amant F, et al. Neoadjuvant chemotherapy or primary surgery in stage IIIC or IV ovarian cancer. *N Engl J Med*. 2010; 363(10):943–953. [PubMed: 20818904]
- 11Aletti 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):676.e671–677. [PubMed: 18060979]
- 12Aletti GD, Dowdy SC, Gostout BS, et al. Quality improvement in the surgical approach to advanced ovarian cancer: the Mayo Clinic experience. *J Am Coll Surg*. 2009; 208(4):614–620. [PubMed: 19476798]
- 13Chi DS, Eisenhauer EL, Zivanovic O, et al. Improved progression-free and overall survival in advanced ovarian cancer as a result of a change in surgical paradigm. *Gynecologic oncology*. 2009; 114(1):26–31. [PubMed: 19395008]
- 14Eisenhauer EL, Abu-Rustum NR, Sonoda Y, et al. The addition of extensive upper abdominal surgery to achieve optimal cytoreduction improves survival in patients with stages IIIC-IV epithelial ovarian cancer. *Gynecologic oncology*. 2006; 103(3):1083–1090. [PubMed: 16890277]
- 15Wright JD, Herzog TJ, Neugut AI, et al. Effect of radical cytoreductive surgery on omission and delay of chemotherapy for advanced-stage ovarian cancer. *Obstet Gynecol*. 2012; 120(4):871–881. [PubMed: 22996105]
- 16Khuri SF, Henderson WG, DePalma RG, et al. Determinants of long-term survival after major surgery and the adverse effect of postoperative complications. *Ann Surg*. 2005; 242(3):326–341. discussion 341-323. [PubMed: 16135919]
- 17Fagotti A, Ferrandina G, Vizzielli G, 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 (Oxford, England : 1990)*. 2016; 59:22–33.
- 18Vergote I, De Wever I, Tjalma W, Van Gramberen M, Decloedt J, van Dam P. Neoadjuvant chemotherapy or primary debulking surgery in advanced ovarian carcinoma: a retrospective analysis of 285 patients. *Gynecol Oncol*. 1998; 71(3):431–436. [PubMed: 9887245]
- 19Meyer LA, Cronin AM, Sun CC, et al. Use and Effectiveness of Neoadjuvant Chemotherapy for Treatment of Ovarian Cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2016
- 20Wright JD, Ananth CV, Tsui J, et al. Comparative effectiveness of upfront treatment strategies in elderly women with ovarian cancer. *Cancer*. 2014; 120(8):1246–1254. [PubMed: 24443159]
- 21Hou JY, Kelly MG, Yu H, et al. Neoadjuvant chemotherapy lessens surgical morbidity in advanced ovarian cancer and leads to improved survival in stage IV disease. *Gynecologic oncology*. 2007; 105(1):211–217. [PubMed: 17239941]
- 22Kang S, Nam BH. Does neoadjuvant chemotherapy increase optimal cytoreduction rate in advanced ovarian cancer? Meta-analysis of 21 studies. *Annals of surgical oncology*. 2009; 16(8):2315–2320. [PubMed: 19517192]
- 23Vergote I, , Amant F, , Leunen K. *Gynecol Oncol Vol. 119*. United States: 2010 Neoadjuvant chemotherapy in advanced ovarian cancer: what kind of evidence is needed to convince US gynaecological oncologists?; 12
- 24Vergote 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 : official journal of the American Society of Clinical Oncology. 2011; 29(31):4076–4078. [PubMed: 21931032]
- 25Vergote I, Leunen K, Amant F. Primary surgery or neoadjuvant chemotherapy in ovarian cancer: what is the value of comparing apples with oranges? *Gynecologic oncology*. 2012; 124(1):1–2. [PubMed: 22153125]
- 26Vergote 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(1):6–11. [PubMed: 23006973]
- 27Chi DS, Bristow RE, Armstrong DK, Karlan BY. Is the easier way ever the better way? *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2011; 29(31):4073–4075. [PubMed: 21931018]
- 28Chi DS, Musa F, Dao F, et al. An analysis of patients with bulky advanced stage ovarian, tubal, and peritoneal carcinoma treated with primary debulking surgery (PDS) during an identical time period as the randomized EORTC-NCIC trial of PDS vs neoadjuvant chemotherapy (NACT). *Gynecologic oncology*. 2012; 124(1):10–14. [PubMed: 21917306]
- 29Schorge JO, Clark RM, Lee SI, Penson RT. Primary debulking surgery for advanced ovarian cancer: are you a believer or a dissenter? *Gynecologic oncology*. 2014; 135(3):595–605. [PubMed: 25316177]
- 30Kehoe S, Hook J, Nankivell M, et al. Primary chemotherapy versus primary surgery for newly diagnosed advanced ovarian cancer (CHORUS): an open-label, randomised, controlled, non-inferiority trial. *Lancet (London, England)*. 2015; 386(9990):249–257.
- 31Dewdney SB, Rimel BJ, Reinhart AJ, et al. The role of neoadjuvant chemotherapy in the management of patients with advanced stage ovarian cancer: survey results from members of the Society of Gynecologic Oncologists. *Gynecologic oncology*. 2010; 119(1):18–21. [PubMed: 20673970]
- 32Cornelis S, Van Calster B, Amant F, Leunen K, van der Zee AG, Vergote I. Role of neoadjuvant chemotherapy in the management of stage IIIC-IV ovarian cancer: survey results from the members of the European Society of Gynecological Oncology. *Int J Gynecol Cancer*. 2012; 22(3):407–416. [PubMed: 22367320]
- 33Onda T, Yoshikawa H. Neoadjuvant chemotherapy for advanced ovarian cancer: overview of outcomes and unanswered questions. *Expert review of anticancer therapy*. 2011; 11(7):1053–1067. [PubMed: 21806329]
- 34Bristow RE, Eisenhauer EL, Santillan A, Chi DS. Delaying the primary surgical effort for advanced ovarian cancer: a systematic review of neoadjuvant chemotherapy and interval cytoreduction. *Gynecologic oncology*. 2007; 104(2):480–490. [PubMed: 17166564]
- 35Rauh-Hain JA, Rodriguez N, Growdon WB, et al. Primary debulking surgery versus neoadjuvant chemotherapy in stage IV ovarian cancer. *Annals of surgical oncology*. 2012; 19(3):959–965. [PubMed: 21994038]
- 36Rauh-Hain JA, Melamed A, Wright A, et al. Overall Survival Following Neoadjuvant Chemotherapy vs Primary Cytoreductive Surgery in Women With Epithelial Ovarian Cancer: Analysis of the National Cancer Database. *JAMA oncology*. 2017; 3(1):76–82. [PubMed: 27892998]
- 37Worley MJ Jr, Guseh SH, Rauh-Hain JA, et al. What is the optimal treatment for obese patients with advanced ovarian carcinoma? *American journal of obstetrics and gynecology*. 2014; 211(3):231.e231–239. [PubMed: 24704062]
- 38McLean KA, Shah CA, Thompson SA, Gray HJ, Swensen RE, Goff BA. Ovarian cancer in the elderly: outcomes with neoadjuvant chemotherapy or primary cytoreduction. *Gynecologic oncology*. 2010; 118(1):43–46. [PubMed: 20398918]
- 39Glasgow MA, Yu H, Rutherford TJ, et al. Neoadjuvant chemotherapy (NACT) is an effective way of managing elderly women with advanced stage ovarian cancer (FIGO Stage IIIC and IV). *Journal of surgical oncology*. 2013; 107(2):195–200. [PubMed: 22648987]
- 40Thrall MM, Gray HJ, Symons RG, Weiss NS, Flum DR, Goff BA. Neoadjuvant chemotherapy in the Medicare cohort with advanced ovarian cancer. *Gynecol Oncol*. 2011; 123(3):461–466. [PubMed: 21945309]

- 41Gill SE, McGree ME, Weaver AL, Cliby WA, Langstraat CL. Optimizing the treatment of ovarian cancer: Neoadjuvant chemotherapy and interval debulking versus primary debulking surgery for epithelial ovarian cancers likely to have suboptimal resection. *Gynecologic oncology*. 2017; 144(2):266–273. [PubMed: 27916269]
- 42Hernán MA, , Robins JM. Causal Inference Boca Raton, FL: Chapman & Hall/CRC; 2017forthcoming
- 43Richiardi L, Bellocco R, Zugna D. Mediation analysis in epidemiology: methods, interpretation and bias. *Int J Epidemiol*. 2013; 42(5):1511–1519. [PubMed: 24019424]
- 44Vergote IB, Van Nieuwenhuysen E, Vanderstichele A. How to Select Neoadjuvant Chemotherapy or Primary Debulking Surgery in Patients With Stage IIIC or IV Ovarian Carcinoma. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2016
- 45Petri H, Urquhart J. Channeling bias in the interpretation of drug effects. *Stat Med*. 1991; 10(4): 577–581. [PubMed: 2057656]
- 46Vergote I, , Trope CG, , Amant F, , Kristensen GB, , Sardi JE, , Ehlen T. 12th Biennial Meeting of the International Gynecologic Cancer Society Bangkok, Thailand: 2008 EORTC-GCG/ NCIC-CTG randomized trial comparing primary debulking surgery with neoadjuvant chemotherapy in stage IIIC-IV ovarian, fallopian tube, and peritoneal cancer.
- 47Ozols RF, Bundy BN, Greer BE, 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. *J Clin Oncol*. 2003; 21(17):3194–3200. [PubMed: 12860964]
- 48Suisse S. Immortal time bias in pharmaco-epidemiology. *Am J Epidemiol*. 2008; 167(4):492–499. [PubMed: 18056625]
- 49Wright AA, Bohlke K, Armstrong DK, et al. Neoadjuvant Chemotherapy for Newly Diagnosed, Advanced Ovarian Cancer: Society of Gynecologic Oncology and American Society of Clinical Oncology Clinical Practice Guideline. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2016; 34(28):3460–3473. [PubMed: 27502591]
- 50Brun JL, Rouzier R, Uzan S, Daraï E. External validation of a laparoscopic-based score to evaluate resectability of advanced ovarian cancers: clues for a simplified score. *Gynecol Oncol*. 2008; 110(3):354–359. [PubMed: 18572226]
- 51Fagotti A, Ferrandina G, Fanfani F, et al. Prospective validation of a laparoscopic predictive model for optimal cytoreduction in advanced ovarian carcinoma. *Am J Obstet Gynecol*. 2008; 199(6): 642.e641–646. [PubMed: 18801470]
- 52Horowitz NS, Miller A, Rungruang B, et al. Does aggressive surgery improve outcomes? Interaction between preoperative disease burden and complex surgery in patients with advanced-stage ovarian cancer: an analysis of GOG 182. *J Clin Oncol*. 2015; 33(8):937–943. [PubMed: 25667285]
- 53Baccarani M, Cortes J, Pane F, et al. Chronic myeloid leukemia: an update of concepts and management recommendations of European LeukemiaNet. *J Clin Oncol*. 2009; 27(35):6041–6051. [PubMed: 19884523]
- 54Bristow RE, Duska LR, Lambrou NC, et al. A model for predicting surgical outcome in patients with advanced ovarian carcinoma using computed tomography. *Cancer*. 2000; 89(7):1532–1540. [PubMed: 11013368]
- 55Andikyan V, Kim A, Gretz HF, et al. Laparoscopic Assessment to Determine the Likelihood of Achieving Optimal Cytoreduction in Patients Undergoing Primary Debulking Surgery for Ovarian, Fallopian Tube, or Primary Peritoneal Cancer. *Am J Clin Oncol*. 2018
- 56Chereau E, Lavoue V, Ballester M, et al. External validation of a laparoscopic-based score to evaluate resectability for patients with advanced ovarian cancer undergoing interval debulking surgery. *Anticancer Res*. 2011; 31(12):4469–4474. [PubMed: 22199317]
- 57Greenland S. An introduction to instrumental variables for epidemiologists. *Int J Epidemiol*. 2000; 29(4):722–729. [PubMed: 10922351]
- 58Thrall MM, Gray HJ, Symons RG, Weiss NS, Flum DR, Goff BA. Trends in treatment of advanced epithelial ovarian cancer in the Medicare population. *Gynecologic oncology*. 2011; 122(1):100–106. [PubMed: 21496889]
- 59Mahner S, Trillsch F, Chi D, et al. Neoadjuvant chemotherapy in ovarian cancer revisited. *Ann Oncol*. 2016; 27(Suppl 1):i30–i32. [PubMed: 27141067]

- 60VanderWeele TJ, Hernán MA. Causal Inference Under Multiple Versions of Treatment. *J Causal Inference*. 2013; 1(1):1–20. [PubMed: 25379365]
- 61Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology*. 1999; 10(1):37–48. [PubMed: 9888278]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 1.
Selection process for articles comparing the effectiveness of neoadjuvant chemotherapy versus primary debulking surgery for advanced ovarian cancer.

Figure 2.

Directed acyclic graph (DAG)* depicting selected methodological issues in observational comparative effectiveness studies of NACT versus PDS for advanced ovarian cancer.

NACT, neoadjuvant chemotherapy; PDS, primary debulking surgery

*DAGs are used to represent causal relationships and inform appropriate confounding control⁶¹.

**Confounding by indication occurs when these factors are unmeasured/uncontrolled.

Figure 3.

Definitions of treatment and outcome in comparative effectiveness studies of neoadjuvant chemotherapy versus primary debulking surgery for advanced ovarian cancer.

Index date=diagnosis; treatment identified at first treatment in regimen. If time from diagnosis to treatment is balanced, results in nondifferential overestimation of survival in both groups^{20,36}.

Index date=diagnosis; treatment identified at surgery. Overestimates survival in both groups, but greater overestimation in NACT group³⁵.

Index date=first treatment; treatment identified at surgery. Overestimates survival for NACT^{37,39}.

Index date=surgery; treatment identified at surgery. Excluding NACT patients who die prior to IDS overestimates survival for NACT. Measuring survival farther from diagnosis underestimates survival for NACT relative to PDS⁴¹.

Landmark analysis: index date=predetermined time post-diagnosis. Anyone who receives treatment between diagnosis and end of landmark period is included in study. May exclude high-risk person-time immediately following PDS, overestimating survival for PDS¹⁹.

Index date=first treatment; treatment identified at first treatment. If time from diagnosis to treatment is imbalanced, consider matching on this person-time^{19,40}.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Characteristics of studies evaluating overall survival for patients treated with neoadjuvant chemotherapy versus primary debulking surgery for advanced ovarian cancer

Table 1

Study	Study Sample	Data Source	Main Comparison	Analysis Methods/Index Date*	Results (NACT vs. PDS)**
McLean, 2010 ⁴³	Age 65, ovarian cancer, 7/1997-3/2007	Seattle Cancer Care Alliance Tumor Registry	NACT (N=34) vs. PDS (N=141)	KM curves; Index date unspecified	Median OS=29vs.34, p=0.7
Thrall, 2011 ⁴⁵	Age 65, stage III-IV EOC, 1995-2005	SEER-Medicare	NACT (N=958) vs. PDS (N=4,827)	KM curves, multivariable Poisson regression; Index date=first treatment	Stage III: Median OS=23vs.28, p<0.001; 1-yr RR=0.90, 95%CI=0.74-1.10; 2-yr RR=1.16, 95%CI=1.01-1.34; Stage IV: Median OS=19vs.18, p=0.4; 1-yr RR=0.69, 95%CI=0.57-0.83; 2-yr RR=0.85, 95%CI=0.73-0.99
Rauh-Hain, 2012 ⁴⁰	Stage IV EOC, 1995-2007	Massachusetts General Hospital, Brigham and Women's Hospital cancer registry databases	NACT+IDS (N=45) vs. PDS (N=176)	KM curves, multivariable CPH regression; Index date=diagnosis	Median OS=33vs.29, p=0.1; HR=1.00, 95%CI=0.60-1.60
Glasgow, 2013 ⁴⁴	Age 70, stage III-IV epithelial ovarian, fallopian, or primary peritoneal cancer, 1/1996-9/2009	Yale-New Haven Hospital Gynecologic Oncology Tumor Board Database	NACT+IDS (N=42) vs. PDS (N=62)	KM curves, CPH regression; Index date=first treatment	Median OS=25vs.39, p=0.9
Wright, 2014 ²⁴	Age 65, stage II-IV EOC, 1991-2007	SEER-Medicare	NACT (N=2,238) vs. PDS (N=5,345)	Multivariable CPH regression, PS-matching, IPTW, IVA; Index date=diagnosis	CPH: Median OS=16vs.29; HR=1.27, 95%CI=1.19-1.35; PS-matching: Median OS=19vs.24; HR=1.24, 95%CI=1.15-1.34; IPTW: Median OS=21vs.27; HR=1.30, 95%CI=1.25-1.35; IVA: Median OS=24vs.24; HR=1.04, 95%CI=0.67-1.60
Worley, 2014 ⁴²	BMI 30, stage III-IV epithelial ovarian, fallopian, primary peritoneal cancer, 2005-2010	Medical records (unspecified institution)	NACT+IDS (N=22) vs. PDS (N=95)	KM curves; Index date=first treatment	Median OS=53vs.32, p=0.04; BMI 30-34.9: Median OS=32vs.58, p=0.03; BMI 35: No significant differences (unspecified)
Meyer, 2016 ²³	Stage III-IV ovarian, fallopian, primary peritoneal cancer, 2003-2012	NCCN Ovarian Cancer Database (6 NCI-designated cancer centers)	NACT vs. PDS PS-matched pairs: N=174 stage III-C, N=123 stage IV	PS-matched KM curves and CPH regression; Index date=first treatment	Stage III-C: Median OS=33vs.43; HR=1.40 95%CI=1.11-1.77; Stage IV: Median OS=31vs.36; HR=1.16, 95%CI=0.89-1.52
Rauh-Hain, 2017 ⁴¹	Age<70, CCI=0, stage III-IV EOC, 2003-2011	National Cancer Database	NACT (N=2,935) vs. PDS (N=2,935)	PS matching, KM curves, multivariable CPH regression; Index date=diagnosis	Median OS=32vs.37; HR=1.18, 95%CI=1.11-1.26; Stage III-C: Median OS=38vs.47; HR=1.24, 95%CI=1.11-1.37; Stage IV: Median OS=28vs.31; HR=1.13, 95%CI=1.04-1.23
Gill, 2017 ⁴⁶	Stage III-IV ovarian, fallopian, primary peritoneal cancer, 2007-2013 (NACT	Mayo Clinic medical records	Age- and stage-matched NACT+IDS (N=45) vs. suboptimal PDS (N=45)	KM curves, multivariable CPH regression; Index date=surgery	Median OS (years)=2.5vs.1.4; HR=0.53, 95%CI=0.31-0.89

Study	Study Sample	Data Source	Main Comparison	Analysis Methods/Index Date*	Results (NACT vs. PDS)**
	+IDS), 2003-2011 (PDS)				

NACT, neoadjuvant chemotherapy; PDS, primary debulking surgery; KM, Kaplan-Meier; OS, overall survival; EOC, epithelial ovarian cancer; SEER, Surveillance, Epidemiology, and End Results; IDS, interval debulking surgery; CPH, Cox proportional hazards; CCI, Charlson Comorbidity Index; PS, propensity score; IPTW, inverse probability of treatment weighting; IVA, instrumental variable analysis; BMI, body mass index; NCCN, National Cancer Comprehensive Network; NCI, National Cancer Institute

* Index date is the time from which survival is measured

** OS presented in months unless otherwise indicated

Table 2

Strengths and limitations of eligible studies evaluating comparative effectiveness of neoadjuvant chemotherapy versus primary debulking surgery for advanced ovarian cancer*

Study	Major strengths	Major limitations	Addressed unmeasured confounding	Assessed HTE [†]	Did not condition on mediator [‡]	Appropriate index date [‡]	Treatment variation irrelevance ^{§§}
McLean 2010	<ul style="list-style-type: none">• Treatment decisions made by gyn/onc• Treatment intent was known• Assessed treatment effects within elderly women	<ul style="list-style-type: none">• No control for confounders• Unspecified index date• Small sample size• Poor description of methods• Treatment changes over time not considered in the analysis	✓	✓	✓		
Thrall 2011	<ul style="list-style-type: none">• Index date was 1st treatment• Attempted to identify treatment intent• Large, nationally representative sample• Stratified by stage, surgical risk	•Surgical intent variable not validated		✓	✓	✓	
Rauh-Hain 2012	<ul style="list-style-type: none">• Large sample size• Stratified by metastatic site• Assessed treatment effects within stage IV disease	<ul style="list-style-type: none">• Conditioned NACT on IDS• Index date was diagnosis date• Adjusted for residual disease• Treatment changes over time not considered in the analysis• Treatment intent unknown	✓				
Glasgow 2013	<ul style="list-style-type: none">• Treatment decisions made by gyn/onc• Many important confounders measured• Assessed treatment effects within elderly women• Treatment intent was known	<ul style="list-style-type: none">• Treatment indication information not used• Small sample size• Treatment changes over time not considered in the analysis• Conditioned NACT on IDS; measured survival from 1st treatment• Performance score measured after receipt of NACT	✓	✓	✓		

Study	Major strengths	Major limitations	Addressed unmeasured confounding	Assessed HTE [†]	Did not condition on mediator [‡]	Appropriate index date [‡]	Treatment variation irrelevance [‡]
		• May have adjusted for mediators					
Worley 2014	<ul style="list-style-type: none">• Treatment decisions made by gyn/onc• Assessed treatment effects within obese patients• Stratified by obesity level	<ul style="list-style-type: none">• No control for confounders• Small sample size• Treatment changes over time not considered in the analysis• Conditioned NACT on IDS; measured survival from 1st treatment		✓	✓		
Wright 2014	<ul style="list-style-type: none">• Large, nationally representative sample• PS and IVA used to address confounding• Strong instrument• Most variables balanced between treatment groups• Considered calendar time in analysis	<ul style="list-style-type: none">• Requiring 6-month survival may introduce selection bias• Instrument may be related to survival outside of treatment• Did not address differences between IVA and PS results• Index date was diagnosis• Cohorts remained imbalanced by calendar year	✓		✓		
Meyer 2016	<ul style="list-style-type: none">• PS matching with well-balanced groups• Many important confounders measured• Sensitivity analyses to address unmeasured confounding• Index date was 1st treatment	• Treatment intent was unknown	✓	✓	✓	✓	✓
Gill2017	<ul style="list-style-type: none">• Measured treatment indication• Treatment intent was known• Attempted to attain exchangeability by comparing NACT to suboptimal PDS	<ul style="list-style-type: none">• Small sample size• Many confounders measured at IDS• Conditioned PDS on inferior outcome• Conditioned NACT on IDS; measured survival from IDS					
Rauh-Hain 2017	<ul style="list-style-type: none">• PS matching, including year	<ul style="list-style-type: none">• Index date was diagnosis	✓	✓	✓		✓

Study	Major strengths	Major limitations	Addressed unmeasured confounding	Assessed HTE [†]	Did not condition on mediator [‡]	Appropriate index date [‡]	Treatment variation irrelevance ^{§§}
	<ul style="list-style-type: none">• Attempted to attain exchangeability through restriction• Sensitivity analyses to address unmeasured confounding• Large, nationally-representative sample• Re-estimation of PS in subgroup analyses	<ul style="list-style-type: none">• Treatment intent was unknown					

^{*} Table S1 defines criteria used for this table.

[†]Through stratification or restriction

[‡]In primary analyses

[§]Changes in treatment over time unrelated to outcomes⁶⁰

NACT, neoadjuvant chemotherapy; PDS, primary debulking surgery; HTE, heterogeneous treatment effects; gyn/onc, gynecological oncologist; IDS, interval debulking surgery; PS, propensity score; IVA, instrumental variable analysis.