Why i.p. therapy cannot yet be considered as a standard of care for the first-line treatment of ovarian cancer: a systematic review

A. M. C. Swart1,2*, S. Burdett3, J. Ledermann4, P. Mook5 & M. K. B. Parmar2

1Centre for Clinical Pharmacology, Division of Medicine University College; 2Cancer Group; 3Meta-analysis Group, MRC Clinical Trials Unit, London; 4Cancer Research UK, UCL Cancer Trials Centre, London; 5Department of Respiratory Diseases, Health Protection Agency Centre for Infections, London, UK

Received 17 October 2006; revised 14 August 2007; accepted 9 October 2007

A National Cancer Institute (NCI) clinical announcement recommended i.p. therapy for women with optimally debulked ovarian cancer. Its basis was a summary of eight randomised controlled trials and two systematic reviews, which appear to indicate benefit of i.p. therapy. However, the systematic reviews that inform the recommendations have been inappropriately presented and interpreted. The systematic reviews inappropriately pooled results from ‘confounded’ trials in which different drugs and different doses of drugs were given in the control and i.p. treatment arms. Therefore, it is not possible to assess which component of treatment is responsible for improving outcome. In addition, none of the trials use a control arm of the internationally accepted standard of care. Using just the unconfounded trials, indirect comparisons show that the magnitude of benefit observed when i.p. regimens are compared with older i.v. regimens [hazard ratio (HR) for overall survival (OS) 0.75; 95% confidence interval (CI) 0.60–0.92, P = 0.006] is smaller than the magnitude of benefit achieved with modern day standard of i.v. treatment compared with the same i.v. regimen used as control in the unconfounded i.p. trials (HR for OS 0.68; 95% CI 0.58–0.80, P < 0.001). A further difficulty is that the reviews cannot recommend an i.p. regimen for standard use. Drug-related toxicity and catheter complications that occur with i.p. therapy are considerable. The NCI recommendations have major implications for the treatment of women with ovarian cancer and for the next generation of clinical trials. We do not believe that the body of evidence currently available supports the recommendation that i.p. therapy should form part of routine care. The choice of treatment of women with newly diagnosed, optimally debulked, ovarian cancer, where therapy has the best chance of influencing OS, is too important to be left with this uncertainty. A clinical trial that investigates a practical and acceptable regimen which gives some or all chemotherapy by the i.p. route and compares this with standard i.v. chemotherapy should be a priority for those who wish to promote its use.

Key words: intraperitoneal therapy, ovarian cancer, systematic review

introduction

In January 2006, USA National Cancer Institute (NCI) issued a clinical announcement on the use of i.p. chemotherapy in the treatment of women with optimally debulked ovarian cancer [1]. The announcement was timed to coincide with the publication of a small but ‘positive’ randomised clinical trial assessing i.p. therapy [2]. Posted with the announcement was a review of the randomised trials of i.p. chemotherapy. Since 2006, there have been four further systematic reviews published which include an assessment of i.p. chemotherapy for the initial management of primary epithelial cancer [3–6]. We have appraised the methods used in these reviews and highlighted the methodological flaws which affect their validity and conclusions.

background

Worldwide there are >200 000 cases of ovarian cancer per year, just under half of which are reported from developed countries [7]. In North America, there are 25 000 cases per year with 16 000 deaths making ovarian cancer the seventh most common cancer and the leading cause of death from gynaecological cancer [8]. Nearly 70% of women present with disease which has spread beyond the ovary. Primary treatment is surgery with the aim of removing as much of the tumour as possible, with better survival rates reported for women left with small amounts of disease after surgery (optimal debulking) [9]. While there have been considerable improvements in outcome for women with ovarian cancer
Annals of Oncology

Volume 19 | No. 4 | April 2008

doi:10.1093/annonc/mdm518 | 689

following the introduction of platinum-based chemotherapy in the 1980s, further advances in the last decade have been modest and the mortality rate for the large majority of women who present with advanced stages of this disease remains high. The major trials of platinum-based chemotherapy published over the past 15 years report that the median progression-free survival (PFS) for patients with advanced disease ranges between 14 and 18 months while the median overall survival (OS) lies between 26 and 40 months [10–13]. Two large trials (GOG111 [11] and OV10 [13]) demonstrated superiority of paclitaxel (Taxol; Bristol Myers Squibb) and cisplatin (CP) over cyclophosphamide and CP on response rate, PFS and OS. Two subsequent large trials (du Bois et al. [14] and Ozols et al. [15]) then showed no significant differences in efficacy between the combination of paclitaxel and CP compared with paclitaxel and carboplatin and also noted an advantage for the carboplatin combination in terms of administration, toxicity and quality of life. Data from all these trials set the standard for an internationa...
representing 94% of all patients entered into the relevant RCTs. Data on PFS could be extracted from four RCTs [2, 24, 26, 27] (1052 patients) representing 58% of all patients entered into the relevant RCTs.

The regimens used in the reference arms of the trials reflect changing practice over the 20-year period that the trials were conducted. The third International Gynecologic Cancer Intergroup Ovarian Cancer Consensus Conference in 2004 [16] recommended that the control arm in international clinical trials of systemic treatment of ovarian cancer should be six cycles of combination i.v. chemotherapy with carboplatin [area under the curve (AUC) 5 to AUC 7.5] and paclitaxel (175 mg/m² given as a 3-h infusion) administered every 3 weeks. While all the i.p. regimens did include platinum-based chemotherapy, only two trials used regimens that would now be considered close to the internationally accepted standard of care of carboplatin and paclitaxel [2, 26].

The experimental arms of the randomized trials vary widely with respect to the chemotherapeutic agents used and the dose of platinum and other drugs. All contained platinum-based chemotherapy; the i.p. platinum compound in seven of the trials was CP, only one trial used i.p. carboplatin. Two of the trials used i.p. platinum in combination with i.v. paclitaxel.

The rational behind considering the trials as one homogenous group is that for all or the majority of improvement associated with the experimental arm is due to therapy being given by the i.p. therapy route and not to any other component of therapy (such as an increased dose of platinum in the i.p. arm). With the aim of reducing clinical heterogeneity, we separated the trials into four groups according to similarities between the reference and experimental arms with respect to the chemotherapy drugs used, the dose of platinum and the dose of paclitaxel (Table 1).

We also classified groups as unconfounded or confounded depending on whether more than one major component of treatment (choice of drugs and dose of platinum used) was different.
changed between the reference and experimental arms. In the unconfounded trials, the only difference between the experimental and reference arms is the route of administration.

In the confounded trials, it is not possible to assess which component of treatment is responsible for any difference in outcome.

Figures 1 and 2 show HRs for OS (Figure 1) and PFS (Figure 2) with 95% (inner ticks) and 99% (outer ticks) CIs for individual i.p. trials. Trials are classified (as in Table 1) according to whether or not they are confounded and according to similarities in the reference and experimental arms with respect to the chemotherapy drugs used, the dose of platinum and the dose of paclitaxel. Information on OS is more complete and less biased in these open-label trials than information on PFS. Summary statistics for each subgroup are presented only if more than one trial is included in the group of trials.

In Figures 1 and 2, group A (i) shows the trials with the only unconfounded comparisons, where the only change between the experimental and reference arms of the trial is the route of administration of platinum, i.p. or i.v. The combined HR for the two trials with data on OS is 0.75 (95% CI 0.60–0.92, \( P = 0.006 \)). For PFS, data were only available for one trial. In both trials, the reference arm was i.v. cyclophosphamide plus CP.

The two trials in group A (ii) gave the same drugs in both reference and experimental arms but gave higher doses of platinum in the i.p. arms compared with the i.v. arms. A higher platinum dose i.p. alone seems to yield only a very modest possible improvement in OS, the combined result of these trials for OS was HR 0.87 (95% CI 0.69–1.09, \( P = 0.227 \)). PFS data were only available for one trial.

In group A (iii), doses of both platinum and paclitaxel are different in the two arms with a higher dose of both platinum and paclitaxel given in the i.p. therapy arm. There is only one trial in this group, but it is the most recently reported publication and was the motivation for the NCI clinical announcement. HR for OS was 0.75 (95% CI 0.58–0.97, \( P = 0.028 \)) and HR for PFS 0.80 (95% CI 0.64–1.00, \( P = 0.50 \)). There are other methodological difficulties with this trial including the exclusion of more randomised patients and more patients lost to follow-up in the i.p. arm. There is also a larger improvement in PFS compared with OS which is counterintuitive and a suggestion that the estimates of median OS may be unreliable at the extremes of the Kaplan–Meier plot because there are relatively few patients at risk at this time. In addition, there are questions about feasibility and acceptability of a treatment in which 42% of patients received less than two cycles of i.p. therapy.
In the two trials in group A (iv), different drugs as well as different doses of platinum were included in the reference and experimental arms. Outcome measures were only available for one of the two trials in this group. The overall results shown in Figures 1 and 2, across all trials using i.p. therapy (whether unconfounded or confounded) give an overall summary result of HR 0.79 (95% CI 0.69–0.90, \(P = 0.003\)) and HR 0.81 (95% CI 0.70–0.93, \(P < 0.001\)) for OS and PFS, respectively. It is not possible to interpret the meaning of the summary statistic with respect to the improvement in outcome that could be obtained with the introduction of i.p. therapy against current international standards because of the wide range of reference and experimental arms across the trials. In none of the trials did the i.p. therapy arm receive less platinum, less paclitaxel or a simpler regimen. In fact, in all but the three unconfounded trials in group A (i), patients receiving i.p. therapy received more treatment than the reference arm. The only trials which can be considered to give a reliable summary of the relative value of i.p. therapy over i.v. therapy are the three trials (of which we have data from the two largest for OS) in the unconfounded comparison. The estimate of the effect of i.p. therapy on OS (HR 0.75; 95% CI 0.60–0.92) obtained for the unconfounded trials is similar to the overall (uninterpretable) estimate, but there are other reasons why i.p. therapy can still not be considered as an appropriate standard therapy.

**choice of control arm and relevance of the results of i.p. trials to current practice**

The most important observation in our review is that none of the trials which consider i.p. therapy compare it against the international standard reference arm of carboplatin and paclitaxel or other widely used regimens (e.g. carboplatin alone). It is therefore not possible to say how the variety of i.p. regimens tested would compare with these regimens. Since no such direct comparison is possible, the only way to consider this issue is to make an indirect comparison with appropriate trials. To allow indirect comparison of the results of the i.p. trials with current standard treatments, we summarise the results of the two trials which compare i.v. therapy with cyclophosphamide and CP with i.v. paclitaxel and CP (Table 2 and Figures 1 and 2—group B). The reference arms of these pivotal trials (GOG111 [11] and OV-10 [13]) which investigated the efficacy of platinum plus paclitaxel were CP 75 mg/m² and cyclophosphamide 750 mg/m². The experimental arm of GOG111 was CP 75 mg/m² and paclitaxel 135 mg/m² >24 h and that of OV10 was CP 75 mg/m² and paclitaxel 175 mg/m² >3 h. Both trials gave six cycles every 3 weeks. Although such indirect comparisons should be treated with caution, the magnitude of improvement in OS seen with i.v. CP and paclitaxel compared with i.v. CP in these two trials (HR 0.68; 95% CI 0.58–0.80) is actually
slightly larger than the magnitude of improvement seen with the unconfounded i.p. trials where the reference arms were i.v. CP and i.v. CP plus doxorubicin (which has been shown to be approximately equivalent to i.v. CP [23]) (HR 0.75; 95% CI 0.60–0.92).

For PFS, data are more limited for the unconfounded comparison of i.p. therapy and so the overall estimate (HR 0.81; 95% CI 0.70–0.93) is used. The magnitude of improvement in PFS with i.v. platinum and paclitaxel compared with CP is also larger than the improvement seen with i.p. therapy (HR 0.71; 95% CI 0.62–0.81, \(P = 0.001\)).

The HR for OS with i.p. therapy of 0.75 would translate into an improvement in median OS of ~12 months (for optimally debulked patients), but with an overall HR of 0.68 for the i.v. CP plus paclitaxel trials, a larger improvement of ~16 months would be expected for similar patients.

**toxicity**

Summary measures of toxicity are difficult to obtain and even when reported are difficult to interpret. This is because few trials reported toxicity in detail, different scales of measurement were used and information was frequently combined by site or grade. Interpretation of comparisons across trials is difficult as toxicity is highly dependent on chemotherapy regimen and dose as well as method of administration. Consideration was therefore given to a number of approaches. The method used in the Cochrane review [3] calculated a pooled odds ratio for types of reported toxicity classified in broad common toxicity criteria categories. Pooled estimates of toxicity favoured i.v. therapy in the adverse outcomes classified as fever, fatigue, gastrointestinal, infection, metabolic and pain. No pooled effect was observed for i.p. therapy. There was considerable heterogeneity as a result of the different regimens that were used. Data published with the NCI alert included a summary of proportions of patients reported with different types and grades of toxicity. Much of the reported toxicity for i.v. therapy is haematological which is easily managed and has a limited impact on quality of life. Interestingly in the one trial where a high proportion of patients in the i.v. therapy reported abdominal symptoms, closer examination of the trial report revealed that 80% of

---

**Table 2.** Intravenous taxane trials using similar reference arms to unconfounded i.p. trials

<table>
<thead>
<tr>
<th>Trial</th>
<th>Published Year</th>
<th>No. of patients</th>
<th>Reference arm</th>
<th>Experimental arm</th>
<th>Outcome measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>GOG111 [11]</td>
<td>1996</td>
<td>410</td>
<td>Q 3–4 × 6 cycles; Cis 75 mg/m² i.v.; Cyclo 750 mg/m² i.v.</td>
<td>Q 3–4 × 6 cycles; Cis 75 mg/m² i.v.; Paclitaxel 135 mg/m² &gt;24 h i.v.</td>
<td>OS, PFS</td>
</tr>
<tr>
<td>OV10 [13]</td>
<td>2000</td>
<td>680</td>
<td>Q 3 × 6 cycles; Cis 75 mg/m² i.v.; Cyclo 750 mg/m² i.v.</td>
<td>Q 3 × 6 cycles; Cis 75 mg/m² i.v.; Paclitaxel 175 mg/m² &gt;3 h i.v.</td>
<td>OS, PFS</td>
</tr>
</tbody>
</table>

Cis, cisplatin; Cyclo, cyclophosphamide; OS, overall survival; PFS, progression-free survival.

**Table 3.** Compliance with protocol treatment i.v. taxane trials and i.p. trials

<table>
<thead>
<tr>
<th>Reference arm (i.v.)</th>
<th>Experimental arm (i.p.)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A. Intrapertioneal chemotherapy trials</strong></td>
<td></td>
</tr>
<tr>
<td>A (i) Unconfounded trials—same chemotherapy drugs; same doses of platinum</td>
<td></td>
</tr>
<tr>
<td>SWOG8501/GOG104 [22]</td>
<td>58%</td>
</tr>
<tr>
<td>Polyzos et al. [23]</td>
<td>NR</td>
</tr>
<tr>
<td>GONO et al. [24]</td>
<td>96%</td>
</tr>
<tr>
<td>A (ii) Confounded trials—same chemotherapy drugs; different dose of platinum</td>
<td></td>
</tr>
<tr>
<td>Yen et al. [25]</td>
<td>32%</td>
</tr>
<tr>
<td>GOG114/SWOG9227 [26]</td>
<td>86%</td>
</tr>
<tr>
<td>A (iii) Confounded trials same chemotherapy drugs; different doses of platinum and different doses of paclitaxel</td>
<td></td>
</tr>
<tr>
<td>GOG1172 [2]</td>
<td>83%</td>
</tr>
<tr>
<td>A (iv) Confounded trials—different chemotherapy drugs; different doses of platinum</td>
<td></td>
</tr>
<tr>
<td>Kirmani et al. [27]</td>
<td>60%</td>
</tr>
<tr>
<td>Zylberberg et al. [28]</td>
<td>NR</td>
</tr>
<tr>
<td>B. Intravenous taxane trials</td>
<td></td>
</tr>
<tr>
<td>Proportion of patients completed all cycles</td>
<td></td>
</tr>
<tr>
<td>Reference arm (Cis/Cyclo)</td>
<td>Reference arm (Cis/paclitaxel)</td>
</tr>
<tr>
<td>GOG111 [11]</td>
<td>78%</td>
</tr>
<tr>
<td>OV10 [13]</td>
<td>79%</td>
</tr>
</tbody>
</table>

NR, not reported; GONO, Gruppo Oncologico Nord-Ovest; Cis, cisplatin; Cyclo, cyclophosphamide.
patients in the i.v. group had an i.p. catheter placed for peritoneal fluid sampling.

Compliance with planned chemotherapy can be regarded as a composite measure of toxicity and efficacy as trial treatment is stopped if patients are failing therapy which can be due to either a toxicity or disease progression or a combination of the two. This outcome measure was reported consistently for the majority of trials and is therefore the approach we used. The proportion of patients receiving all cycles of planned chemotherapy is summarised in Table 3. Overall, compliance and deliverability of i.p. regimens are generally worse than i.v. reference regimens in the i.p. trials and worse than i.v. CP and paclitaxel (85%) in the taxane trials.

For the i.p. trials, data were available for two of the three unconfounded trials [group A (i)]. In one trial, the same proportion of patients in both arms completed all cycles (58%) and in the other a higher proportion of patients allocated to i.v. received all cycles (96%) than those allocated to i.p. (64%). In the confounded trials which gave the same chemotherapy drugs but different dose of platinum [group A (ii)], higher proportions of patients in both trials allocated to i.v. chemotherapy received all six cycles.

In the confounded trial where different doses of platinum and paclitaxel were given [group A (iii)], only 42% of patients completed all cycles of the i.p. regimen compared with 83% of patients assigned to i.v. This was the only trial where patient-assessed outcomes were measured. Quality of life measured by Functional Assessment of Cancer Therapy- Ovarian module (adjusted for baseline imbalances) was worse in patients receiving i.p. therapy before cycle 4 and 3–6 weeks after completion of treatment although at 12 months after the sixth cycle, no significant difference was detected. Toxicity was only reported in one of the trials in group A (iv), the trials where very different regimens were given. Gastrointestinal toxicity was very high in both groups but 78% of patients in the i.v. group assigned to i.v. therapy had peritoneal catheters in situ which were used for peritoneal cytology sampling and two of these patients developed peritoneal-vaginal fistulas.

**relative toxicity of i.p. therapy and i.v. carboplatin and paclitaxel**

No trials compare i.p. therapy with the current standard of i.v. carboplatin and paclitaxel. Indirect comparisons from Table 2, however, show that overall, giving some drugs via the i.p. route (platinum alone or platinum plus paclitaxel), therapy is more toxic and less deliverable than therapy with i.v. platinum and paclitaxel. Catheters cause abdominal pain, nausea and vomiting and there is an increased risk of infection and fever [26]. Intrapерitoneal paclitaxel has been associated with increased haematological, metabolic and neurological toxicity [2].

**conclusions**

The reviews of i.p. therapy and the recommendations that they have generated have major implications for the practical management for a large proportion of women with ovarian cancer who are having to make treatment decisions on the basis of incomplete data. They also have important implications for the design of the next generation of clinical trials which will set the standard treatments for the future. Before addressing the challenge of incorporating the newer angiogenesis inhibitors, with potential problems of wound healing and bowel perforations, into regimens which include i.p. therapy, it is very important that we know that i.p. regimens are genuinely better than standard i.v. therapy. The reviews that have been carried out have indicated that i.p. therapy should be considered as part of the standard approach to treating optimally debulked International Federation of Gynecology and Obstetrics (FIGO) stage III ovarian cancer and the NCI recommendation is that ‘after primary surgery, women with optimally debulked FIGO stage III ovarian cancer should be counselled about the clinical benefit associated with combined IV and IP administration of chemotherapy’. The recommendations are included on the NCI’s Web site with a designation of IIiA for the level of evidence [33]. This level of evidence is designated for randomised (level 1), unblinded (ii) trials with mortality as an end point (A). The designation does not take into account deficiencies in both the control and the experimental arms of the i.p. trials. A closer and more critical and appropriate consideration of the data, however, shows that the apparently ‘clear’ body of evidence in favour of i.p. therapy is not as strong as indicated. Compared with current standard of care, the benefits of i.p. therapy may have been overestimated. We believe that it is impossible at this stage to give women accurate information on the risks and benefits of i.p. therapy compared with the modern standard of care. The clinical heterogeneity of the regimens in the trials has led the NCI to recommend as standard practice i.p. CP (100 mg/m²) and a taxane (whether given by an i.v. only or i.v. and i.p.). It is not clear whether giving a taxane i.v. only or a taxane i.v. and i.p. is equivalent in terms of toxicity and efficacy, or how the choice should be made for individual patients. The regimen in the i.p. arm of GOG172 [2] included i.p. CP and i.v. and i.p. paclitaxel. The only other trial which included i.p. CP and i.v. paclitaxel also included two cycles of high-dose i.v. carboplatin in the experimental arm. Thus, it is not clear what the standard i.p. regimen should be. Any i.p. CP/paclitaxel regimen that deviates from that used in GOG172 cannot be said to have been assessed in any trial. The NCI recommendation is a direct consequence of the lack of clarity in the design of the individual trials and the inappropriate presentation and interpretation of systematic reviews. No new treatment would be approved by regulatory authorities or should be adopted as standard clinical practice on the basis of this evidence. Allied to this, i.p. regimens are considerably more toxic than the current standard of care and indirect comparisons strongly indicate that they may actually be less effective than a current standard of care of i.v. platinum and paclitaxel. Others have attempted to use indirect approaches to assess the benefit of the i.p. regimen used in GOG172 and concluded that if the i.v. therapy reference arm in GOG172 had been treated with i.v. carboplatin and paclitaxel instead of i.v. CP and paclitaxel, the difference in OS may have been considerably reduced (8.2 months, instead of 15.9 months) [34].

694 | Swart et al. | Volume 19 | No. 4 | April 2008
Consequently, we do not believe that i.p. therapy should form part of the routine care of women with ovarian cancer. Until problems with drug-related toxicity and catheter complications are resolved, even proposals for further randomised trials of i.p. therapy against modern standard of care could be considered as speculative. The choice of treatment of women with newly diagnosed, optimally debulked, ovarian cancer, where therapy has the best chance influencing OS, is too important to be left with these uncertainties.

funding

MRC Clinical Trials Unit to A.M.S., S.B., P.M. and M.P.; University College London to J. L.

references

17. Bookman MA, for the Gynecologic Cancer InterGroup (GOIG). GOG0182-ICON5: 5-arm phase II randomized trial of paclitaxel (P) and carboplatin (C) vs combinations with gemcitabine (G), PEG-liposomal doxorubicin (D), or topotecan (T) in patients (pts) with advanced-stage epithelial ovarian (EOC) or primary peritoneal (PPC) carcinoma. J Clin Oncol 2006; 24 (18S): 5002.