Individual Patient Data (IPD) Reviews and Meta-analyses

Lesley Stewart, Jayne Tierney, Claire Vale
IPD Meta-analysis Methods Group


Stewart LA, Tierney JF. To IPD or Not to IPD? Advantages and disadvantages of systematic reviews using individual patient data. Evaluation & the Health Professions 2002;25(1):76-97.
IPD systematic review / meta-analysis

- Described as yardstick and gold standard of systematic review
- Central collection, validation & re-analysis of source data
- Philosophy same as for other Cochrane reviews
- Process differs in terms of data collection and analysis
- Quicker and cheaper than new trial, but longer and more resource intensive than other reviews
- Less common than other types of review but becoming used increasingly
Established history in cardiovascular disease

Established history in range of cancer sites e.g.
- chemotherapy for ovarian cancer
- post-operative radiotherapy for lung cancer
- chemotherapy for bladder cancer
- chemoradiation for cervical cancer

Becoming used in a wide range of fields e.g.
- surgical repair for hernia
- drug treatments for epilepsy
- cholinesterase inhibition for Alzheimer’s disease
- anti-platelet treatments for pre-eclampsia in pregnancy
- compression bandaging for chronic leg ulcers
How IPD meta-analyses are organised

- Carried out by international collaborative group
  - small local secretariat
  - multi-disciplinary advisory group
  - trialists who provide data
- Developing and maintaining this group requires communication and careful management
- Publication in the name of collaborative group
Why IPD?

- Analyses based on published data can give different answers to an IPD meta-analysis e.g.
  - chemotherapy in advanced ovarian cancer
  - radiotherapy in SCLC
  - chemotherapy in NSCLC
  - ovarian ablation in breast cancer
  - immunisation for recurrent miscarriage
  - chemotherapy for head and neck cancer
Why IPD?
Chemotherapy in advanced ovarian example

<table>
<thead>
<tr>
<th></th>
<th>IPD</th>
<th>Published Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trials</td>
<td>11</td>
<td>8</td>
</tr>
<tr>
<td>Patients</td>
<td>1329</td>
<td>788</td>
</tr>
<tr>
<td>Odds Ratio</td>
<td>---</td>
<td>0.71</td>
</tr>
<tr>
<td>Hazard Ratio</td>
<td>0.93</td>
<td>---</td>
</tr>
<tr>
<td>95% confidence interval</td>
<td>0.83-1.05</td>
<td>0.52-0.96</td>
</tr>
<tr>
<td>p-value</td>
<td>0.30</td>
<td>0.027</td>
</tr>
<tr>
<td>Absolute benefit at 30 months</td>
<td>2.5%</td>
<td>7.5%</td>
</tr>
<tr>
<td>Comments</td>
<td>median follow up 6.5 years</td>
<td>point estimate at 30 months</td>
</tr>
</tbody>
</table>

Platinum based combination vs non-platinum single drugs, *Lancet* 1993; 341: 418-422
Ovarian cancer example conclusions

- Differences due to
  - excluded trials, excluded patients, time point of analysis, extra follow up, analysis method
- Published summary data gives a more statistically ‘convincing’ result
- Estimates of effect size are 7.5% and 2.5% improvement in survival at 30 months
- Balanced against other factors, clinical interpretation of results from two approaches may be different
Why IPD?

- Include all trials published and unpublished
- Get round inadequacies in trial reports
  - measure or define patient characteristics differently
  - measure or define outcomes differently
  - selectively report particular outcomes
  - based on different degrees of follow up
  - exclude patients from analyses
  - inappropriate or biased analyses
  - insufficient details of analyses
- Address questions or carry out analyses that cannot be readily achieved with published data
Why IPD?

- Improve data quality
  - all relevant trials and patients
  - all relevant outcomes
  - combine different scales of measurement
  - data checking
- Improve analysis quality
  - include all patients by intention-to-treat
  - appropriate analyses (e.g. time-to-event analysis)
  - long term outcomes
  - patient subgroups
- Improve trial identification, interpretation & dissemination via collaborative approach
Specific reasons for using IPD

- Neo-adjuvant chemotherapy for bladder cancer
  - better estimate of effect on survival
  - effect on different patient subgroups
- Adjuvant chemotherapy for bladder cancer
  - treatment in use, but published data & analyses poor
  - appropriately analyse and rigorously appraise IPD
- Chemoradiation for cervical cancer
  - effect on different patient subgroups
  - detailed analysis of toxicity
- Anti-platelet therapy for pre-eclampsia in pregnancy
  - explore whether effect differs by women’s risk profile
### To IPD or not to IPD?

#### When IPD may be beneficial
- Poor reporting of trials. Information inadequate, selective or ambiguous
- Long-term outcomes
- Time-to-event outcome measures
- Multivariate or other complex analyses
- Differently defined outcome measures
- Subgroup analyses of patient-level characteristics important
- IPD available for high proportion of trials/individuals

#### When IPD may not be beneficial
- Detailed and clear reporting of trials (CONSORT quality)
- Short-term outcomes
- Binary outcome measures
- Univariate or simple analyses
- Outcome measures defined uniformly across trials
- Patient subgroups not important
- IPD available for only a limited number of trials
Doing a systematic review and meta-analysis of IPD
Comparing types of review / meta-analysis

Write protocol: State objectives, searches, inclusion criteria and planned analyses prospectively

Identify all relevant trials

Establish Secretariat, Advisory, Trialist Groups

Assemble the most complete dataset possible

Collect and validate data

Analyse individual studies and perform meta-analysis

Hold Collaborators’ Conference

Prepare structured report

Processes the same for summary data and IPD

Processes are similar for summary data and IPD, but methodology and practical aspects differ

Only applies to IPD approach
Protocol development

- Introduction
- Objectives
- Trials inclusion criteria
- Identification of trials
- Data collection
- Data analysis
- Publication policy
- Timetable

Consult with Advisory Group as required

- Similar to Cochrane reviews
- More detailed than for Cochrane reviews
Protocol development

Identification of trials
Data collection
Data analysis
Identification of trials

- Any review restricted to published data is at risk of publication bias
- Include all relevant trials published & unpublished
- Unpublished trials not peer reviewed, but
  - trial protocol & IPD allows extensive ‘peer review’
  - can clarify proper randomisation, eligibility
  - quality publication does not guarantee quality data
- Proportion of trials published will vary by
  - disease, intervention, over time
- Extent of unpublished data can be considerable
Identification of trials
Chemoradiation for cervical cancer (initiated 2004)

- Published (76%)
- Abstract only (8%)
- Unpublished (13%)
Identification of trials
Chemoradiation for cervical cancer

- Electronic databases
  - Medline, Cancerlit, LILACS
- Trial Registers
  - e.g. Clinicaltrials.gov, PDQ (cancer.gov), metaRegister, CENTRAL
- Hand search
  - reference lists, conference proceedings
- Experts
  - include preliminary trial list in protocol and ask collaborators to supplement it
Identification of trials
Chemoradiation for cervical cancer

- Medline/Cancerlit (79%)
- LILACS (4%)
- Trials Registers (8%)
- Handsearching (8%)
Which IPD to collect: All patients

- Trial investigators frequently exclude patients from trial analyses and reports
  - ineligibility, patient withdrawal, early outcome, lost to follow-up
- Ad hoc exclusion of patients could introduce bias
- Aim to collect data on all randomised patients
- Also useful to collect data on which patients were excluded and the reasons for their exclusion
  - retention of such data may vary by disease and intervention
Which IPD to collect: All patients

Tierney JF, Stewart LA. Investigating exclusion bias in meta-analysis. *Int J Epidemiol* 34:79-87
Which IPD to collect: All patients
Chemotherapy for soft tissue sarcoma

- Obtained data for 14 trials, 1568 patients
- 341 (22%) of these patients excluded from the investigators’ analyses

<table>
<thead>
<tr>
<th>Patients excluded</th>
<th>Events / Patients</th>
<th>Overall Survival</th>
</tr>
</thead>
<tbody>
<tr>
<td>Investigators</td>
<td>553 / 1227</td>
<td>HR 0.85, P=0.056</td>
</tr>
<tr>
<td>None</td>
<td>709 / 1568</td>
<td>HR 0.90, P=0.157</td>
</tr>
</tbody>
</table>
Which IPD to collect: All patients
Chemotherapy for soft tissue sarcoma

- Pre-specify in the protocol if any patients will be excluded from the analysis
- Assess impact by sensitivity analyses

<table>
<thead>
<tr>
<th>Patients excluded</th>
<th>Events / Patients</th>
<th>Overall Survival</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Locally recurrent &lt;15 years</td>
<td>597 / 1366</td>
<td></td>
<td>0.91</td>
</tr>
<tr>
<td>Metastatic Induction CT</td>
<td></td>
<td></td>
<td>0.278</td>
</tr>
<tr>
<td>None</td>
<td>709 / 1568</td>
<td>0.90</td>
<td>0.157</td>
</tr>
</tbody>
</table>
Which IPD to collect: Variables

- Decision by secretariat in consultation with Advisory Group
- Think about the analyses and work back
- Only want data necessary to carry out these analyses and adequately describe trials
- Publications can indicate
  - which data are feasible (but note there may be more available than reported)
Which IPD to collect: Variables

- Basic identification of patients
  - e.g. anonymous patient ID, centre ID
- Baseline data for descriptive purposes or analyses
  - e.g. age, sex, disease or condition characteristics
- Intervention of interest
  - e.g. date of randomisation, treatment allocated
- Outcomes of interest
  - e.g. survival, toxicity, maternal death, pre-eclampsia, wound healing
- Information on excluded patients
- Include list of variables in meta-analysis protocol
Which IPD to collect: Variables
Chemoradiation for cervical cancer

Baseline characteristics
- Patient ID
- Centre ID
- Patient date of birth or age
- Tumour histology
- Tumour stage
- Tumour grade
- Lymph node involvement
- Patient performance status
- Allocated treatment
- Date of randomisation

Outcome characteristics
- Surgery
- External beam radiotherapy
- Brachytherapy

Outcomes
- Tumour response
- Loco-regional recurrence status
- Date of loco-regional recurrence
- Distant metastases status
- Date of distant metastases
- Survival status
- Date of death or last follow-up
- Acute toxicity
- Late toxicity

Other
- Cause of death
- Whether excluded from analysis
- Reason for exclusion
IPD variable definitions

- Form the basis of the meta-analysis database
- Define variables in way that is unambiguous and facilitates data collection and analysis
- Publications and protocols can indicate
  - how to collect data
## IPD variable definitions

### Chemoradiation for cervical cancer

<table>
<thead>
<tr>
<th>Variable</th>
<th>Type</th>
<th>Width</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
<td>numeric</td>
<td>3</td>
<td>age in years</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>unknown = 999</td>
</tr>
<tr>
<td><strong>Performance status</strong></td>
<td></td>
<td></td>
<td>Accept whatever scale is used, but request details of the system used</td>
</tr>
<tr>
<td><strong>Survival status</strong></td>
<td>numeric</td>
<td>1</td>
<td>0 = Alive</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1 = Dead</td>
</tr>
<tr>
<td><strong>Tumour stage</strong></td>
<td>numeric</td>
<td>1</td>
<td>1 = Stage Ia</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2 = Stage Ib</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3 = Stage IIa</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4 = Stage IIb</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5 = Stage IIIa</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6 = Stage IIIb</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>7 = Stage IVa</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>8 = Stage IVb</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>9 = Unknown</td>
</tr>
<tr>
<td><strong>Date of death</strong></td>
<td>date</td>
<td>-</td>
<td>date in dd/mm/yy format</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>unknown day = --/mm/yy</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>unknown month = --/--/yy</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>unknown date = --/--/--</td>
</tr>
</tbody>
</table>
IPD variable definitions
Anti-platelet therapy for pre-eclampsia in pregnancy

- **Onset of labour**
  - 1 = spontaneous
  - 2 = induced
  - 3 = pre-labour caesarian
  - 9 = not recorded

- **Sex of baby**
  - 1 = male
  - 2 = female
  - 3 = ambiguous
  - 9 = not recorded

- **Pre-eclampsia**
  - Highest recorded systolic BP in mmHg
  - Highest recorded diastolic BP in mmHg
  - Proteinurea during this pregnancy
    - 0 = no
    - 1 = yes
    - 9 = unknown

  Date when proteinurea first recorded

These variables allow common definition of pre-eclampsia and early onset pre-eclampsia
IPD variable definitions
Anti-platelet therapy for pre-eclampsia in pregnancy

✖ Gestation at randomisation
  In completed weeks
  9 = unknown

Poor choice of code for missing value, woman could be randomised at 9 weeks gestation

✖ Severe maternal morbidity
  1 = none
  2 = stroke
  3 = renal failure
  4 = liver failure
  5 = pulmonary oedema
  6 = disseminated intravascular coagulation
  7 = HELP syndrome
  8 = eclampsia
  9 = not recorded

Collection as a single variable does not allow the possibility of recording more than one event
| Variable definitions |

**Baseline characteristics**

**Date of birth**
- Date in dd/mm/yy or dd/mm/yyyy format.
- Unknown day --/mm/yy
- Unknown month dd/--/yy
- Unknown date --/--

**Histology**
- 1 squamous
- 2 adenosquamous
- 3 adenocarcinoma
- 7 other
- 9 unknown

**Clinical Stage (FIGO)**
- 1 IA
- 2 IB
- 3 IIA
- 4 IIB
- 5 IIIA
- 6 IIIB
- 7 IV A
- 8 IV B
- 9 unknown

**Grade**
- 1 well differentiated
- 2 moderately differentiated
- 3 poorly differentiated / undifferentiated
- 9 unknown

**Performance status**
- Code as convenient, but please supply full details of the system used (e.g. ECOG, Karnofsky, WHO, CMS)

**Pelvic lymph node involvement**
- 0 not involved
- 1 involved
- 9 unknown

**Para-aortic lymph node involvement**
- 0 not involved
- 1 involved
- 9 unknown

**Suggested Coding**

**Local treatment characteristics**

**Surgery**
- 0 no
- 1 hysterectomy
- 2 hysterectomy + pelvic lymphadenectomy
- 3 hysterectomy + pelvic + para-aortic lymphadenectomy
- 7 other
- 9 unknown

**External beam radiotherapy**
- 0 no
- 1 pelvic field
- 2 extended field (pelvic + para-aortic)
- 7 other
- 9 unknown

**Brachytherapy**
- 0 no
- 1 yes
- 9 unknown

**Outcomes**

**Survival Status**
- 0 alive
- 1 dead

**Dates of death or last follow up**
- Date in dd/mm/yy or dd/mm/yyyy format (as for date of birth)

**Loco-regional progression / recurrence status**
- 0 no progression / recurrence
- 1 progression / recurrence

**Date of locoregional progression / recurrence**
- Date in dd/mm/yy or dd/mm/yyyy format (as for date of birth)

**Distant metastases status**
- 0 no metastases
- 1 metastases

**Date of distant metastases**
- Date in dd/mm/yy or dd/mm/yyyy format (as for date of birth)

**Toxicity**

**Acute toxicity data**
- Haematological toxicity (any)
- Haemoglobin toxicity / anaemia
- Thrombocytopenia
- White blood cell toxicity (any)
- Gastrointestinal toxicity (any)
- Genitourinary toxicity (any)
- Skin toxicity (any)
- Other toxicity (any)
- Please supply the most severe grade experienced for each category. Code as convenient giving full details of the grading system used (e.g. CTC, etc).

**Late toxicity data**
- Intestinal toxicity (any)
- Rectal toxicity (any)
- Bladder toxicity (any)
- Vaginal toxicity (any)
- Other toxicity (any)
- Please supply the most severe grade experienced for each category. Code as convenient giving full details of the grading system used (e.g. CTC, etc).

**Other**

**Whether excluded from the analysis**
- 0 no
- 1 yes
- 9 unknown

**Reason for exclusion**
- Supply as convenient but please provide details; for example:
  - ineligible - too old
  - ineligible - metastatic disease found after randomisation
  - protocol violation - clinician withdrew patient
  - lost to follow-up - patient withdrew from trial
  - etc.

**Exploratory analysis of haemoglobin**

**Pre-treatment haemoglobin**
- Please definitions and coding will be supplied after the trial is completed.
Planning analyses

- Range of possibilities
  - Main analyses of outcomes
  - Subset analyses by trial group
  - Subgroup analyses by patient characteristics (patient treatment interactions)
    - realistically only possible with IPD
  - Sensitivity analyses
  - Exploratory analyses (e.g. prognostic factors, baseline risk etc.)
  - Time-to-event analysis
- Pre-specify all in protocol
Planning analyses
Chemoradiation for cervical cancer

Main analyses of outcomes
- survival, local and distant disease-free survival, response, acute and late toxicity

Subset analyses by
- chemotherapy type, dose intensity & scheduling
- radiotherapy dose and duration

Subgroup analyses by
- patient age and performance status, tumour histology, stage and grade and lymph node involvement
Planning analyses
Chemoradiation for cervical cancer

- Sensitivity analysis
  - by trial design
- Exploratory analysis of
  - relationship between treatment, haemoglobin levels and outcome
Collecting Data
Initiating collaboration with trialists

- Initial letter inviting collaboration, but not yet asking for data explaining
  - main aims and objectives
  - importance of the collaborative group
  - publication policy
  - collaborative group policy
  - confidentiality of data
- Ask specific questions relating to trial eligibility
Trial level data collection

- Data needed to adequately describe the trial
  - Trial ID and trial title
  - Method of randomisation & allocation concealment
  - Planned treatments
  - Recruitment and stopping information
  - Other information that is not clear from trial report

- Obtaining the trial protocol can also be valuable in describing a trial

- Use to clarify eligibility
  - Establish table of included studies
Trial level data collection

- Principal contact details
- Data contact details
- Up to date trial publication information
- Other trials of relevance
- Whether willing to take part in meta-analysis
- Preferred method of data transfer

This information can be collected on forms accompanying the meta-analysis protocol
Example form

<table>
<thead>
<tr>
<th>Meta-analysis of Concomitant Chemoradiotherapy for Locally Advanced Cancer of the Uterine Cervix</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name: ____________________________________________________________________________________</td>
</tr>
<tr>
<td>Name of trial: __________________________________________________________________________</td>
</tr>
<tr>
<td>Are you willing to take part in this meta-analysis?  Yes ☐ No ☐  If yes, please can you supply a copy of the trial protocol and forms when you return this form.</td>
</tr>
<tr>
<td>Trial Design</td>
</tr>
<tr>
<td>Was informed consent obtained from each patient?  Yes ☐ No ☐  Date trial opened to accrual: ___________</td>
</tr>
<tr>
<td>Date trial opened to accrual: ___________  Date trial closed to accrual: ___________</td>
</tr>
<tr>
<td>What method of randomisation was used?  Simple ☐ Permuted blocks ☐ Minimisation ☐ Other ___________</td>
</tr>
<tr>
<td>What method was used to conceal randomisation?  Sealed envelope ☐ Central telephone ☐ Other ___________</td>
</tr>
<tr>
<td>What, if any, stratification factors were used? ______________________________________________________________________________</td>
</tr>
<tr>
<td>What proportions was the trial designed to have in each arm (e.g. 1:1)? __________________________________________________________________</td>
</tr>
<tr>
<td>Early Stopping</td>
</tr>
<tr>
<td>Did the trial have a target for patient accrual?  Yes ☐ No ☐  Did the trial reach its target accrual?  Yes ☐ No ☐  Was the trial stopped early?  Yes ☐ No ☐  Was a formal stopping rule used?  Yes ☐ No ☐</td>
</tr>
<tr>
<td>If a formal stopping rule was not used, what was the reason for stopping the trial? __________________________________________________________________</td>
</tr>
<tr>
<td>Data Transfer</td>
</tr>
<tr>
<td>Please provide data on all patients randomised. You may complete the data forms provided or supply your data as a computer printout, on floppy disk (formatted for PC) or by e-mail. Data can be in almost any format (ASCII, Excel, Dbase, FoxPro, etc.), but please indicate which format has been used. Data files should be encrypted. It would be helpful if you used the coding suggested. However, you may code the data in the way that is most convenient to you. Please supply us with full details of the data coding system used.  Are you able to use the suggested coding?  Yes ☐ No ☐  Which method of data encryption would you prefer? (e.g. WinZip etc) __________________________________________________________________</td>
</tr>
<tr>
<td>Guarantee of Confidentiality of Individual Trial Results</td>
</tr>
<tr>
<td>Data will remain the property of the trial investigator who supplied it and will not be used, circulated or distributed in any way that allows access to individual trial data, without first seeking the permission of the trial investigator.  I want my data to remain confidential  Yes ☐ No ☐</td>
</tr>
</tbody>
</table>


Example form

<table>
<thead>
<tr>
<th>Meta-analysis of Concomitant ChemoRadiotherapy for Locally Advanced Cancer of the Uterine Cervix</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Name:</strong> ____________________________</td>
</tr>
<tr>
<td><strong>Telephone:</strong> ________________________</td>
</tr>
</tbody>
</table>

If different from above, please give details of the appropriate contact for the collection of your trial data.

**Name:** ____________________________ | **Address:** ____________________________ |
| **Telephone:** ________________________ | **E-mail:** ____________________________ |

Are the details of your trial correct? [ ] Yes [ ] No

Is the most recent publication of your trial listed in Appendix A of the protocol? [ ] Yes [ ] No

If no, please give details

Do you know of any other relevant trials not listed in Appendix A of the protocol? [ ] Yes [ ] No

If yes, please give details

---

**Which of the following data would you be able to supply for each patient randomised?**

<table>
<thead>
<tr>
<th>Baseline characteristics</th>
<th>Yes</th>
<th>No</th>
<th>Performance status</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient identifier (preferably not patient name)</td>
<td>[ ] Yes [ ] No</td>
<td>Pelvic lymph node involvement</td>
<td>[ ] Yes [ ] No</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Centre identifier</td>
<td></td>
<td>Para-aortic lymph node involvement</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Date of birth or age at randomisation</td>
<td></td>
<td>Iliac lymph node involvement</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Histology</td>
<td></td>
<td>Date of randomisation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clinical Stage (FIGO)</td>
<td></td>
<td>Allocated treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Local treatment characteristics</th>
<th>Yes</th>
<th>No</th>
<th>Other</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surgery</td>
<td>[ ] Yes [ ] No</td>
<td>Whether excluded from the analysis</td>
<td>[ ] Yes [ ] No</td>
<td></td>
<td></td>
</tr>
<tr>
<td>External beam radiotherapy</td>
<td></td>
<td>Reason for exclusion</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brachytherapy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Yes</th>
<th>No</th>
<th>Survival status</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tumour response</td>
<td></td>
<td>Date of death or last follow-up</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Locoregional progression/recurrence status</td>
<td></td>
<td>Cause of death</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Date of locoregional recurrence/progression</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---
### Meta-analysis of Concomitant Chemoradiotherapy for Locally Advanced Cancer of the Uterine Cervix

**Baseline characteristics**

<table>
<thead>
<tr>
<th>Date of birth</th>
<th>Date in dd/mm/yy or dd/mm/yyyy format. Unknown day --/mm/yy Unknown month dd/--/yy Unknown date --/--/--</th>
</tr>
</thead>
<tbody>
<tr>
<td>Histology</td>
<td>squamous 1 adenosquamous 2 adenocarcinoma 3 other 7 unknown 9</td>
</tr>
<tr>
<td>Clinical Stage (FIGO)</td>
<td>IA 1 IB 2 IIA 3 IIB 4 IIIA 5 IIIB 6 IVa 7 IVb 8 unknown 9</td>
</tr>
<tr>
<td>Grade</td>
<td>well differentiated 1 moderately differentiated 2 poorly differentiated / undifferentiated 3 unknown 9</td>
</tr>
<tr>
<td>Performance status</td>
<td>Code as convenient, but please supply full details of the system used (e.g. ECOG, Karnofsky, WHO, CMS)</td>
</tr>
<tr>
<td>Pelvic lymph node involvement</td>
<td>0 not involved 1 involved 9 unknown</td>
</tr>
<tr>
<td>Para-aortic lymph node involvement</td>
<td>0 not involved 1 involved 9 unknown</td>
</tr>
</tbody>
</table>

**Suggested Coding**

| Surgery | 0 no 1 hysterectomy 2 hysterectomy + pelvic lymphadenectomy 3 hysterectomy + pelvic + para-aortic lymphadenectomy 7 other 9 unknown |
| External beam radiotherapy | 0 no 1 pelvic field 2 extended field (pelvic + para-aortic) 7 other 9 unknown |
| Brachytherapy | 0 no 1 yes 9 unknown |

**Local treatment characteristics**

| Outcomes | |
| Survival Status | 0 alive 1 dead |
| Dates of death or last follow up | Date in dd/mm/yy or dd/mm/yyyy format (as for date of birth) |
| Loco-regional progression / recurrence status | 0 no progression / recurrence 1 progression / recurrence |
| Date of loco-regional progression / recurrence | Date in dd/mm/yy or dd/mm/yyyy format (as for date of birth) |
| Distant metastases status | 0 no metastases 1 metastases |
| Date of distant metastases | |

**Toxicity**

| Acute toxicity data | Haematological toxicity (any) Haemoglobin toxicity / anaemia Thrombocytopenia White blood cell toxicity (any) Gastrintestinal toxicity (any) Genitourinary toxicity (any) Skin toxicity (any) Other toxicity (any) Please supply the most severe grade experienced for each category. Code as convenient giving full details of the grading system used (e.g. CTC, etc.) |
| Late toxicity data | Intestinal toxicity (any) Rectal toxicity (any) Bladder toxicity (any) Vaginal toxicity (any) Other toxicity (any) Please supply the most severe grade experienced for each category. Code as convenient giving full details of the grading system used (e.g. CTC, etc.) |
| Other | Whether excluded from the analysis 0 no 1 yes 9 unknown |
| Reason for exclusion | Supply as convenient but please provide details, for example: ineligible - too old ineligible - metastatic disease found after randomisation protocol violation - clinician withdrew patient lost to follow-up - patient withdrew from trial etc. |
| Exploratory analysis of haemoglobin | |
| Pre-treatment haemoglobin | Precise definitions and coding will be established in future studies |
Initiating collaboration with trialists

- Barriers
  - Practical (tracing people, language differences)
    - e-mail, web-sites, directories, search engines
  - Unfamiliar with methods
    - protocol, good communication
  - Political (difficult people, powerful groups)
    - protocol, good communication, intermediaries
  - Financial (money for data or preparing data)
    - ???
Maintaining contact with trialists

- Important to maintain good communication throughout
  - regular correspondence
  - newsletters
  - e-mails

- Often deal with more than one person per trial
  - clinical coordinator, statistician, data centre
  - keep everyone informed & no crossed wires
Concomitant chemoradiotherapy for cervical cancer: A meta-analysis of individual patient data

Meta-analysis progress

Since the last newsletter in February, we have received more responses to our invitation to participate. Fourteen investigators have now replied. Five have already provided their data; seven more are able to participate and some are already preparing their data. Three investigators have replied that they are keen to participate, but will need to have appropriate permissions to supply data to us. We will try to help these investigators wherever we can.

We are now keen to hear from the remaining investigators. Please let us know if you are able to participate or not. We might be able to help.

Missing Trialists: Can you help?

We are still not sure if we have been able to reach all of the investigators and so we may not have received our previous correspondence. We would therefore like to check whether our contact details for the investigators of the following trials are out of date. Can you help us by letting us know if you have up-to-date contact information for any of the following:

- Wong LC, Choi SY, Chey DJ, Sham JST, Maa HK, University of Hong Kong (Cytogenetic Oncology 1989; 35:159-63)
- Tsang CD, Cheng TY, Choy SY, Ty As part of the data collection.
- Singh TT, Singh BV, Sharma DT, Singh NL, Regional Institute of Medical Sciences, Imphal, Manipur State, India (Indian J Cancer 2003; 40:3:10-16)
- Ayala Hernandez Jr, de la Huerta RS, Canfield FM, Drezec AP, Hospital de Oncologia, CAW, INSS, Mexico City (Cytogenetic and Obstetric 1999; 25:258-62)
- Wong LC, Ng KNYS, Cheung ANY, Cheng WK, Ng TY, Chey TDK, University of Hong Kong (Journal of Clinical Oncology 1999; 17(17):2005-60)
- Fernandez DJ, Vidjasagar MS, Kao JCI, Shenoy A, Kasturi DR, Kastura Medical College, Manulal, Karnataka State, India (Proc 16th Annual Meeting of the Assoc, Radiation Oncologists of India 1995; Kargil: 97-105)
- Builes R, Rivera R, Hospital San Felipe, Tegucigalpa, Honduras (Premsa Med Argentina 1986; 72(3):100-5)

Consumer involvement

In this project we will involve women from the UK who have been affected by cervical cancer as Research Partners. They will work with the Meta-analysis Secretariat on a variety of aspects of the project. By working together we may better understand and interpret side effects data; disseminate the findings widely and identify where more research can be done to address issues that are important to women affected by cervical cancer.

On April 1st, the first meeting of our Reference Group for consumer involvement took place. The meeting highlighted that we need to make it clearer what we expect the Research Partners to do and think carefully about what support they may need. We have now drafted Information and Terms for the Reference Group, which is currently reviewing.

We have already recruited two women to be Research Partners and hope to recruit 8-14 more. We aim to hold a first meeting of the Research Partners in the next month or so.

Collaborators’ Meeting

The results of the meta-analysis will be presented for the first time at a meeting of the Collaborative Group. We are currently looking at potential venues for this meeting, which we aim to hold this end of 2002.

In order to do this, we need to complete our analyses by late Autumn. For this reason we have set a deadline of September 2005 for receipt of any data. Please let us know if this deadline may cause you any difficulty.

Chemotherapy for locally advanced bladder cancer

A meta-analysis of individual patient data

We are very pleased to report an excellent response to the meta-analysis so far. We have either received or been promised data from 80% of the trials we identified. However, we are still waiting for agreement to collaborate from a further 4 trials. If you can help us to contact the investigators responsible for these trials, we would be very grateful.

Since the last newsletter, we have received data from two more trials and would like to thank all those responsible for their work. We are currently collating this information to ensure that we represent the data accurately prior to inclusion in the meta-analysis.

Collaborators’ Conference

An important feature of all our individual patient data meta-analyses is that we hold a Collaborators’ Conference, where we present and discuss the results for the first time. For this meta-analysis, we will be holding the Collaborators’ Conference at the Naiicon Hall Hotel, Birmingham, UK on 21-22 February 2002. This will coincide with the European Urological Association meeting, which will be taking place from 29-27 February 2002 at the National Exhibition Centre, Birmingham. All of those investigators who provide data for the meta-analysis will be invited to attend.

The meeting will follow a fairly informal format to allow open and detailed discussion of the results. All collaborators will enjoy spacious, en-suite rooms within the hotel and access to a range of leisure facilities. The hotel has a reputation for excellent food and high standards and we hope the hall will be an ideal venue for our meeting.

Naiicon Hall Hotel

Naiicon Hall is a fourteenth-century Italianate house, set amongst 15 acres of countryside. It is just 10 minutes drive from the National Exhibition Centre and Birmingham International Airport. The meeting will be held in the hotel’s conference centre over two days, which will enable us to discuss the results fully whilst giving you time to relax and enjoy the facilities at the hotel.

We would also like to thank those investigators who are currently preparing their data to submit to us and arrange them to be ready as soon as possible. Ideally we need to have collected in all of the data by the end of November 2001 to ensure that we have enough time to thoroughly check and verify the data with you before we begin the analyses. Success of the meta-analysis and Collaborators’ Conference relies on your data. We very much hope that you will hear from us soon.
Data collection: Principles

- Flexible data formats
  - data forms, database printout, flat text file (ASCII), spreadsheet (e.g. Excel), database (e.g. Dbase, Foxpro), other (e.g. SAS dataset)
  - Accept transfer by electronic or other means

- Security issues
  - request anonymous patient IDs
  - encrypt electronic data

- Accept the trialists coding, secretariat can re-code
  - but suggest data coding

- Offer assistance
  - site visit, financial ??
Data collection: Method of data transfer

- Chemotherapy for ovarian cancer (initiated 1989)
  - 44% on paper, 39% on disk, 17% by e-mail

- Chemotherapy for bladder cancer (initiated 2001)
  - 10% on paper, 10% on disk, 80% by e-mail

- Chemoradiation for cervical cancer (initiated 2004)
  - 10 data sets received so far, 100% by e-mail
Data collection: Time to assemble data

Neoadjuvant chemotherapy for locally advanced cervix cancer

- Protocol and searches
  May 98 - Jan 99
- Invite to collaborate
  Mar 1999
- Collaborators’ meeting
  Sep 2000

Neoadjuvant chemotherapy for locally advanced bladder cancer

- Protocol and searches
  Dec 00 - May 01
- Invite to collaborate
  Jun 2001
- Collaborators’ meeting
  Feb 2002
Set up meta-analysis database
Retain copy of trial data as supplied
Convert data formats (ASCII, spreadsheet, database, etc.) to database format
  • Excel, Dbase, Access, Foxpro, SPSS, SAS, Stata
  • software more compatible now
Data collection: Managing trial data

- Re-code data to meta-analysis coding
  - calculate or transform derived variables e.g.
    - calculate survival time from date of death / last follow-up and date of randomisation
    - derive disease-free survival from recurrence / progression / metastases and survival variables

- Keep records of all changes to trial data

- Check, query and verify data with trialist
  - improved software automates more tasks

- Then append trial to meta-analysis database
### Example individual patient data

<table>
<thead>
<tr>
<th>Patient ID</th>
<th>Date of randomisation</th>
<th>Treatment allocated</th>
<th>Age</th>
<th>Stage</th>
<th>Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>001</td>
<td>23 June 1990</td>
<td>Control</td>
<td>46</td>
<td>2b</td>
<td>poor</td>
</tr>
<tr>
<td>002</td>
<td>19 Oct 1988</td>
<td>Treatment</td>
<td>39</td>
<td>4</td>
<td>moderate</td>
</tr>
<tr>
<td>003</td>
<td>01 Feb 1991</td>
<td>Treatment</td>
<td>51</td>
<td>2a</td>
<td>good</td>
</tr>
<tr>
<td>004</td>
<td>09 April 1987</td>
<td>Control</td>
<td>32</td>
<td>3</td>
<td>moderate</td>
</tr>
<tr>
<td>203</td>
<td>11 Nov 1989</td>
<td>Control</td>
<td>40</td>
<td>2b</td>
<td>good</td>
</tr>
<tr>
<td>204</td>
<td>03 Jan 1990</td>
<td>Treatment</td>
<td>35</td>
<td>2a</td>
<td>poor</td>
</tr>
<tr>
<td>205</td>
<td>15 Mar 1992</td>
<td>Control</td>
<td>56</td>
<td>3</td>
<td>moderate</td>
</tr>
</tbody>
</table>
Example individual patient data

<table>
<thead>
<tr>
<th>PatID</th>
<th>DOR</th>
<th>Arm</th>
<th>Age</th>
<th>Stage</th>
<th>Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>001</td>
<td>23/06/1990</td>
<td>2</td>
<td>46</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>002</td>
<td>19/10/1988</td>
<td>1</td>
<td>39</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>003</td>
<td>01/02/1991</td>
<td>1</td>
<td>51</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>004</td>
<td>09/04/1987</td>
<td>2</td>
<td>32</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>203</td>
<td>11/11/1989</td>
<td>2</td>
<td>40</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>204</td>
<td>03/01/1990</td>
<td>1</td>
<td>35</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>205</td>
<td>15/03/1992</td>
<td>1</td>
<td>56</td>
<td>5</td>
<td>2</td>
</tr>
</tbody>
</table>
Data checking: Rationale

- IPD enables detailed data checking, not easily achieved with any other approach
- Reasons for checking
  - not to centrally police trials or to expose fraud
  - improve accuracy of data
  - improve follow-up
  - ensure appropriate analysis
  - ensure all randomised patients are included
  - ensure no non-randomised patients are included
Data checking: Types

- **Standard**
  - missing data, excluded patients
  - internal consistency and range checks
  - compare with publication

- **Randomisation**
  - balance across arms and baseline factors
  - pattern of randomisation

- **Follow-up**
  - up-to-date and equal across arms

- **Verification**
  - send tables, data list and trial analysis to trialist
Data checking: Pattern of randomisation
Chemoradiation for cervical cancer

![Graph showing the number of patients randomised over time (1991 to 1996). The x-axis represents the date of randomisation, and the y-axis represents the number of patients randomised. Two lines are depicted: one for Chemoradiation (red) and one for Control (dashed). The number of patients increases over time, with a steady increase in both groups.](image)
Data checking: Pattern of randomisation
Radiotherapy vs Chemotherapy in Multiple Myeloma
Data checking: Weekday randomised Chemotherapy for bladder cancer

ARM
- Neoad CT
- Control

Number of randomisations

- MONDAY
- TUESDAY
- WEDNESDAY
- THURSDAY
- FRIDAY

Number of randomisations: 40, 30, 20, 10
Data checking: Weekday randomised
Chemoradiation for cervical cancer

Number of randomisations

ARM
CTRT
Control

SUNDAY  MONDAY  TUESDAY  WEDNESDAY  FRIDAY  SATURDAY
Data checking: Weekday randomised
Post-operative radiotherapy in lung cancer

Number of randomisations

Arm
- RT
- Control
Data checking: Follow up
Chemotherapy for bladder cancer

‘Reverse’ survival curve - take patients event-free, use censoring as event
Data checking: Follow up
Chemotherapy for bladder cancer

![Cumulative Survival Graph]

- ARM
- Control
- Neoadj CT
Analysing data
Analysis: General principles

- Most commonly, 2-stage analysis
  - same summary statistics used
    - odds ratio, relative risk risk difference, mean difference and standardised mean
  - derived from IPD for each trial
  - combined in meta-analysis, stratified by trial

- Less commonly, 1 stage analysis
  - regression/modelling approach
  - all patients are combined into a single ‘mega’ trial (not appropriate)

Benefits of IPD approach to analysis

- IPD can improve analysis quality
- Use the IPD to **re-do the analyses** from scratch, in the same way in all trials, **correcting any problems** in original analyses
Benefits of IPD approach to analysis

- E.g Adjuvant bladder cancer - previous systematic reviews based on published data raised concerns about some trials
  - did not use conventional log rank tests to compare treatment and control arms
  - did not conduct intention-to-treat analyses
  - did not clearly define / report outcomes
- Outcomes re-defined from IPD and analyses re-done appropriately
Major benefit of IPD is that it allows time-to-event analysis, which takes account of:
- whether an event happens
- the time at which it happens

For some diseases just the ability to do such an analysis justifies the IPD approach:
- cure is not likely, prolongation of survival
- time to onset of disease, time free of symptoms
Analysis: Time-to-event

- Individual patient data
  - uses individual times at which each event takes place & takes account of censoring
  - uses log rank O-E & V
  - summarises entire survival experience
  - estimate hazard ratio (HR)
  - allows survival curves
Exploring trial-level differences

Subset analysis
- Or ‘subgroup’ analysis by trial characteristics
- Group by trial treatments, methodology, quality etc.
  - drug type, treatment scheduling
  - drug dose
- Compares the size of treatment effect on outcome across different trial groups
- Easy to do with published summary data or IPD
- May have more trial level data when collecting IPD
Subset analysis
Chemotherapy for bladder cancer

<table>
<thead>
<tr>
<th></th>
<th>CT</th>
<th>Control</th>
<th>O-E</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Single agent platinum</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wallace [2]</td>
<td>59/83</td>
<td>50/76</td>
<td>2.74</td>
<td>27.18</td>
</tr>
<tr>
<td>Martinez-Pineiro [3]</td>
<td>43/62</td>
<td>38/59</td>
<td>0.33</td>
<td>20.11</td>
</tr>
<tr>
<td><strong>Sub-total</strong></td>
<td>136/186</td>
<td>125/190</td>
<td>8.92</td>
<td>63.80</td>
</tr>
<tr>
<td><strong>Platinum-based combinations</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cortesi unpublished</td>
<td>43/82</td>
<td>41/71</td>
<td>-1.87</td>
<td>20.84</td>
</tr>
<tr>
<td>MRC/EORTC [9]</td>
<td>275/491</td>
<td>301/485</td>
<td>-23.69</td>
<td>143.61</td>
</tr>
<tr>
<td>Sherif [7]</td>
<td>79/158</td>
<td>90/159</td>
<td>-6.37</td>
<td>42.18</td>
</tr>
<tr>
<td>Sengeløv [8]</td>
<td>70/78</td>
<td>60/75</td>
<td>1.79</td>
<td>31.96</td>
</tr>
<tr>
<td><strong>Sub-total</strong></td>
<td>686/1220</td>
<td>744/1213</td>
<td>-55.67</td>
<td>355.65</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>822/1406</td>
<td>869/1403</td>
<td>-46.75</td>
<td>419.45</td>
</tr>
</tbody>
</table>

Hazard Ratio

- HR=1.15 p=0.264
- HR=0.86 p=0.003
- HR=0.89 p=0.022

Interaction p=0.029
Exploring patient-level differences

Subgroup analyses

- Group by type of patient
  - age, sex, tumour stage, tumour grade
  - previous hypertensive disorders of pregnancy, previous SGA infant

- Compares size of treatment effect on outcome (not prognosis) across patient subgroups
Exploring patient-level differences

- Difficult to do with published summary data
  - trial-level summaries of patient-level information e.g. mean age
  - rarely report outcome according to patient subgroups

- Easy to do with IPD which allows
  - many combinations of subgroups and outcomes
  - consistent definition of subgroups across trials
Subgroup analysis
Post-operative radiotherapy for lung cancer

Hazard Ratio

Age
- <=54
- 55-59
- 60-64
- >=65

Test for trend
\( \chi^2_{(1)} = 0.929, \ p=0.335 \)

Sex
- Female
- Male

Test for interaction
\( \chi^2_{(1)} = 0.005, \ p=0.944 \)

Histology
- Adenocarcinoma
- Squamous
- Other

Test for interaction
\( \chi^2_{(1)} = 0.572, \ p=0.751 \)
Subgroup analysis
Post-operative radiotherapy for lung cancer

Hazard Ratio

Stage
1
2
3

Nodal Status
0
1
2

Test for trend
$\chi^2_{(1)} = 13.194, \ p=0.0003$

Test for trend
$\chi^2_{(1)} = 5.780, \ p=0.016$
Analysis: Exploratory/sensitivity

- Assess the robustness of the main IPD results e.g.
  - with and without a particular trial
  - with or without particular types of patients excluded in a consistent way across all trials
  - compared to published data when IPD could not be obtained

- Explore additional hypotheses
  - adjustment for imbalances in baseline characteristics
  - prognostic factor analysis
Most IPD groups use own software
  - ours (SCHARP) does 2-stage analyses and produces graphical output for
  - re-developed version available next year

Input into RevMan
  - primary analysis needs to be done elsewhere
  - for time-to-event outcomes use “IPD” or “generic inverse variance” outcome type
  - for other outcomes use appropriate RevMan outcome types (e.g. “dichotomous” etc)
  - not easy to enter (patient) subgroup analyses
Collaborators’ Meeting

- Integral part of IPD approach
- IPD MA a collaborative project
- Incentive to collaborate
- Trialists have opportunity
  - to discuss results
  - to challenge the analysis
  - to discuss interpretation & implication of results
  - Suggest new research
- Sets a deadline to which secretariat and trialists have to work
Resources required

- Likely to be more costly and time-consuming
  - need empirical data
  - but technology advances to cut costs/time
- But differences between IPD and other types of systematic review may not be so great
- IPD projects can be run concurrently
- Practical / political issues
- Cost of Collaborators’ Conference not encountered in other types of review
Getting started

- Contact IPD Meta-analysis Methods Group
  - Administrator: Larysa Rydewska (lhr@ctu.mrc.ac.uk)
  - Website (http://www.ctu.mrc.ac.uk/ukcccr/ipd/home.asp)
    - Database of ongoing and planned IPD reviews
    - Database of methodological projects
    - Reference lists, FAQ,s etc
- Cochrane handbook (to be updated)
- Mentoring - work with someone who has already completed an IPD meta-analysis
To IPD or not to IPD?

- Many benefits particularly
  - improved data and analysis quality
  - improved trial identification, interpretation and dissemination
  - collaboration on further research

- Some benefits possible through collection of additional summary data, but
  - re-doing analyses, re-classifying data etc. may be as much or more work for trialists?

- So why not collect IPD?