Methodology for literature review, data mining, modeling of dose-response in pediatric radiation oncology

Søren M Bentzen, Ph.D., D.M.Sc.
Division of Biostatistics and Bioinformatics, UMCGC Shared Biostatistics Service
University of Maryland Greenebaum Cancer Center, Baltimore, MD, USA
sbentzen@som.umaryland.edu

New RT techniques: childhood cancer

Life Years Lost – radiation modalities

- Meningeal infection
- Heart failure
- Nervous system cancer
- Thyroid cancer
- Lung cancer
- Breast cancer
Performance of ‘pediatric’ search filters

- 82.5% of a sample of Cochrane systematic reviews had errors in the reported search strategy that could potentially bias the reviews’ results

- The Cochrane Childhood Cancer Group tested 7 published search filters for identifying pediatric studies in PubMed in a reference data set

- The sensitivity varied from 44.8% to 99.5%

Literature search (i): example of endpoint

- Cochrane style, systematic literature review
- Led by Cecile Ronckers and Leontine Kremer in collaboration with each site-specific review group

EXAMPLE OF TOXICITY: Hearing Loss
- Dizziness OR hearing loss OR deaf, Hearing OR hearing disorder OR hearing disorders OR auditory OR hearing impairment OR hearing impairments OR hearing impairment* OR hear* OR audiologic OR audioling OR audiometry OR audiom* OR audiogram OR audiography OR ototoxicology OR ototoxic* OR hypacusis OR hypacusises OR hypacusa* OR ototoxicity OR deaf* OR cochleotoxicity

Literature search (ii):

- Childhood cancer
  - [(leukaemia OR leukaem* OR leukaemia* OR ALL OR AML OR lymphoma OR lymphom* OR hodgkin* OR hodgkin* OR T-cell OR B-cell OR non-hodgkin OR sarcoma OR sarcom* OR sarcoma, Ewing's OR Ewing* OR osteosarcoma OR osteosarcom* OR wilms tumor OR wilms* OR nephroblastom* OR neuroblastoma OR neuroblastom* OR rhabdomyosarcoma OR rhabdomyosarcom* OR teratoma OR teratom* OR hepatoma OR hepatom* OR hepatoblastoma OR hepatoblastom* OR PNET OR medulloblastoma OR medulloblastom* OR PNET* OR neuroectodermal tumors, primitive OR retinoblastoma OR retinoblastom* OR meningoma OR meningom* OR glioma OR gliom*) OR (pediatric, oncology OR pediatric, oncology)] OR childhood cancer OR childhood tumor OR childhood tumours) OR (brain tumor* OR brain tumour* OR brain neoplasm OR central nervous system neoplasm OR central nervous system tumour OR central nervous system tumour* OR brain cancer* OR brain neoplasm* OR intracranial neoplasm*) OR (leukaemia, lymphocytic, acute) OR (leukaemia, lymphocytic, acute/leukaemia, lymphocytic, acute/leukemia, lymphocytic, acute)}
Literature search (iii)

- Radiotherapy
  - radiometry OR radiation dosage OR radiation dose OR radiation doses OR radiation doses OR radiation dosage* OR radiation dosimetry OR radiation dosimeter* OR dose-response relationship, radiation OR radiotherapy dosage OR radiotherapy(adverse effects OR irradiation dose OR radiotherapy dose OR dose calculation OR near beam dose OR in beam dose OR outside beam dose OR out of beam dose OR radiation/epidemiology OR radiation monitoring OR Organs at risk OR radiation effects(h) OR radiation injury OR radiation injuries OR radiation OR Radiotherapy/complications(Mesh) OR NCTP OR normal tissue complication probability OR DVH OR Dose Volume Histogram OR Radiotherapy Planning OR Conformal/adverse effects OR Dose Response Relationship, radiation OR Organs at Risk/Radiation Effects OR Radiation Injuries/Prevention and Control OR Chemoradiotherapy/Adverse Effects

Pubmed (Sept 23, 2013): n=544 hits → title/abstract screening

Radiation-induced hypothyroidism

- 514 records identified in the base
- 270 records detailed in PubMed
- 97 records after application of filters
- 42 additional records identified by hand-search
- 34 records screened

Mining of ‘big data’ sets

- What data sets???
- ...in pediatric radiation oncology?????
- ‘Big data’ are data sets in the terabyte to exabyte range
- Biomedical data sets are typically several orders of magnitude smaller than the above range... but some are relatively ‘big’
- A number of attractive analytical techniques have been developed for data mining – supervised as well as unsupervised – and there is current interest in Natural Language Processing...
- ...BUT...
Two worlds colliding

**Radiation Biology**
- Relative homogeneous population exposed
- Known exposure: Accurate estimate of dose to relevant target structure
- Specific clinically-biological endpoint
- Simple mathematical dose-response function characterized by two parameters

**Epidemiology**
- Heterogeneous population exposed
- Poorly known exposure: Rough "guess-estimate" of dose to relevant target structure
- Endpoint often not specific to radiation exposure
- Excess relative risk (ERR)/Excess absolute risk (EAR)
  - Ideally as a function of (some) dose

Tamoxifen enhances XRT lung fibrosis

HD: Mantel field vs. Involved node RT

Tamoxifen enhances XRT lung fibrosis

HD: Mantel field vs. Involved node RT

Tamoxifen enhances XRT lung fibrosis

HD: Mantel field vs. Involved node RT

Tamoxifen enhances XRT lung fibrosis

HD: Mantel field vs. Involved node RT
% increase in major coronary events v. dose

- Relative risk estimates are often more impressive than absolute risk estimates.
- And may give rise to false impressions of e.g. risk vs. age.
- The most relevant metric to a patient is ABSOLUTE RISK.

Death of ischemic heart disease after RT

- Treated at age <60 years old.
- No risk fact.

HD: Mantel field vs. Involved node RT

- Maraldo et al. IJROBP 83: 1232 (2012)
- Darby et al. NEJM 368: 987 (2013)
Death of ischemic heart disease after RT

Excess Absolute Risk of coronary events

Excess Absolute Risk of coronary events
Radiation myelopathy after RT for NSCLC

- Is it safe to give 3x6 Gy+5x4 Gy to the spinal cord?
- Linear quadratic model, assuming $\alpha/\beta = 2$ Gy
  - This schedule has an EQD2 = 66 Gy
- Reported incidence of RM: 4±1%
- Actuarial estimate @ 3 years: 30±15%

Kaplan-Meier or cumulative incidence?

CHART Head and Neck
All sites, except larynx

Kaplan-Meier or cumulative incidence?

CHART Head and Neck
All sites, except larynx

Proportion with moderate/severe dryness of mouth

Ataman & Bentzen (unpublished)
## Incidence among 5-year survivors

<table>
<thead>
<tr>
<th></th>
<th>Pts with one or more incident (%)</th>
<th>K-M estimate @ 5 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHART</td>
<td>49/92</td>
<td>53.3%±5.2%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>48.8%±2.7%</td>
</tr>
<tr>
<td>conv.</td>
<td>32/54</td>
<td>59.3%±6.2%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>59.5%±3.4%</td>
</tr>
</tbody>
</table>

2P=0.49  2P=0.02