Quality of evidence in rare cancers

Paolo Bruzzi
Former Head, Clinical Epidemiology
IRCCS Policlinico San Martino – Genova
Retired
Evidence Based Medicine

The integration of best research evidence with clinical expertise and patient values

Sackett DL et al
Cookbook medicine: “anyone can play the doctor! Just read the guidelines!”
Paternalistic Medicine: “I know what is best for you”
Clinical Expertise

Eminence Based Medicine: “in my opinion....”

Patient values & Preferences
Evidence Based Medicine

The integration of best research evidence with clinical expertise and patient values = RCT, Systematic Review of RCT’s

Most effective therapy
Most Effective Therapy for disease Z?

• The treatment which produces the best effects on patients

• ?

• The treatment **MOST LIKELY** to produce

...the DESIRED EFFECTS...

... on the largest **PROPORTION** of patients
Efficacy: 2 Probabilities

1. **Likely**: Probability that treatment Y is the most effective (on average)

2. **Proportion**: Probability that patient X is one of those who actually benefit from treatment Y
Efficacy: 2 Probabilities

1. **Likely**: Probability that treatment Y is the most effective (on average)

Trial results -> Average effect: Health administrator, Doctor
Efficacy: 2 Probabilities

1. **Likely**: Probability that treatment Y is the most effective (on average)

2. **Proportion**: Probability that patient X is one of those who actually benefit from treatment Y

Trial: results -> **PROBABILITY** distribution of **LIKELY** effects: THE PATIENT!
Efficacy or Effectiveness?

Patient: *Efficacy*

Health administrator: *Effectiveness*
Efficacy is the extent to which an intervention does more good than harm under ideal circumstances

Patient:
- Which is the best therapy for me?
- (where can I get it?)
Standard Definition

**Effectiveness** assesses whether an intervention does more good than harm when provided under usual circumstances of healthcare practice.

Health administrator: In the setting (hospital, state, etc.) I am governing (BUDGET, skills, facilities, population, geography, etc.) which choice is going to produce the most benefits?
Standard Definitions

Effectiveness → Pragmatic trials

Efficacy → Explanatory Trials
## Explanatory and pragmatic trials

<table>
<thead>
<tr>
<th>Trial</th>
<th>Explanatory</th>
<th>Pragmatic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aim</td>
<td>Efficacy</td>
<td>Effectiveness</td>
</tr>
<tr>
<td>Patients</td>
<td>Selected</td>
<td>All comers</td>
</tr>
<tr>
<td>• Protocols</td>
<td>Detailed</td>
<td>Generic</td>
</tr>
<tr>
<td>• Follow-up</td>
<td>Ad hoc</td>
<td>Standard practice</td>
</tr>
<tr>
<td>• Endpoint</td>
<td>Instr/Clin.</td>
<td>Clinical</td>
</tr>
<tr>
<td>• Centers</td>
<td>Specialized</td>
<td>Generic</td>
</tr>
<tr>
<td>• Size</td>
<td>Average</td>
<td>Large</td>
</tr>
</tbody>
</table>
## Explanatory and pragmatic trials

<table>
<thead>
<tr>
<th>Trial</th>
<th>Explanatory</th>
<th>Pragmatic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aim</td>
<td>Efficacy</td>
<td>Effectiveness</td>
</tr>
<tr>
<td>• Patients</td>
<td>Selected</td>
<td>All comers</td>
</tr>
<tr>
<td>• Size</td>
<td>Average</td>
<td>Large</td>
</tr>
</tbody>
</table>

Large numbers to address uncertainty
Evidence Based Medicine in Rare Tumors

The integration of best research evidence with clinical expertise and patient values = LARGE RCT, Systematic Review of RCT’s

Most effective therapy
Question

How should we assess the efficacy of new treatments (drugs) in rare tumors?
Rare Tumors

- Rare histologies in frequent sites (e.g. breast cancer with a squamous histology)
- Rare sites (e.g. uveal melanoma)
- Both (e.g. astrocytomas, most sarcomas)
Rarity: General problem in cancer research

- Rare Tumors
- Rare Cancer Conditions (e.g. multiple c.)
- Rare presentations (e.g. skin metastases)
- Molecular Variants (of common tumors)
How to establish therapeutic standards in very rare tumors/conditions?

a) No trials
   1. “Expert” opinion (based on what?)
   2. Indirect evidence

Standard approach since humans started treating diseases

Not acceptable in the current era dominated by Evidence Based Medicine
How to establish therapeutic standards in very rare tumors/conditions?

a) No trials
   1. “Expert” opinion (based on what?)
   2. Indirect evidence

b) “Small” trials
A study must have an adequate size

- To warrant an adequate “power” to the study (i.e. to reduce the risk of a false negative result; a false negative: an effective treatment is not recognised)

- To obtain precise estimates of the effects of the experimental therapy
Conventional Statistical Rules

• A study **must** have an adequate size

• Required Size, based on:
  – Significance level (usually 5%)
  – Power (usually 80-90%)
  – **Minimal clinically worthwhile difference**
### Minimal Clinically Worthwhile Difference (MCWD)

N. of events needed for $\alpha = 5\%$ and Power = 80%

<table>
<thead>
<tr>
<th>Risk Reduct. $(1 - HR)$</th>
<th>Required N. of Events</th>
</tr>
</thead>
<tbody>
<tr>
<td>-10%</td>
<td>2830</td>
</tr>
<tr>
<td>-20%</td>
<td>635</td>
</tr>
<tr>
<td>-30%</td>
<td>252</td>
</tr>
<tr>
<td>-40%</td>
<td>125</td>
</tr>
<tr>
<td>-50%</td>
<td>71</td>
</tr>
<tr>
<td>-60%</td>
<td>43</td>
</tr>
</tbody>
</table>
Sample Size needed in cancer trials for breakthrough drugs

In trials in early disease, cumulative mortality from 10% to 70%: 500-5000 pts

In trials in advanced disease, cumulative mortality from 50% to 90%: 300-1000 pts
Conventional Statistical Rules

• A study **must** have an adequate size (**Usually Hundreds/Thousands of patients**)

• **International cooperations**

• In many very rare cancer conditions: **NOT POSSIBLE**
  
  – Site
  – Histology
  – Stage
  – Biology
Statistical Mantra

• A study must have an adequate size

Unjustified Implication

• If an adequate size cannot be attained, (RARE CANCERS) no methodological ties

Small size  →  Poor quality
Poor Quality?

• (Study protocol)
• **Classified as Phase II trials**
• No Randomised controls
• Opaque selection of cases
• Primary endpoint: Objective response
• No statistical plan
PD-1 Blockade with...in Advanced Merkel-Cell Carcinoma

Paul T. Nghiem, M.D., Ph.D., Shailender Bhatia, M.D.,

“...phase 2, single group, Simon’s two-stage…”
Response Rate: 14/25 (56%)
(Historical RR: 50-60%)

Median PFS: 9 months
(Historical: 3 months)
in patients with chemotherapy-refractory metastatic Merkel cell carcinoma: a multicentre, single-group, open-label, phase 2 trial

Response Rate: 28/88 (32%)
Target 20%
Historical: 50-60% in 1st line

Median PFS: 2.7 months
(Historical ?)
Trials in Rare Cancers

If, despite International cooperation/prolonged accrual

• it is possible to assemble (in a reasonable time) only a limited number of patients,

• and the EXPECTED effect of the new treatment is not large...
...you are bound to conduct “Small” trials

How to design them?
4 issues

• Phase II or phase III?

• Randomised or Uncontrolled?

• Endpoints

• Conventional or unorthodox statistics?
Phase II vs Phase III

Common but False Beliefs

- If the number of patients is inadequate for a standard phase III, run a phase II trial.
- Phase II studies are uncontrolled trials without any methodological or statistical planning.
- In rare tumours you are allowed to do uncontrolled phase II of poor quality.
Phase II vs Phase III

- The phase of a trial is determined by its aim(s) not by its size or methodology

- If a trial aims at evaluating the efficacy of a treatment, it is a phase III trial

- Efficacy trials can be controlled or uncontrolled (see next slides)
4 Issues

• *Phase II or phase III?*

• Randomised or Uncontrolled?
Randomised or Uncontrolled?

False Beliefs

- Randomised trials require large numbers of patients, while uncontrolled (phase II?) trials do not

- Uncontrolled trials do not require a statistical plan

Few patients -> Uncontrolled trials
<table>
<thead>
<tr>
<th>Type of error</th>
<th>Definition</th>
<th>Errors in</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sampling error</td>
<td>Error due to chance</td>
<td>- Selection of groups</td>
</tr>
<tr>
<td>Bias</td>
<td></td>
<td>- Assessment of outcomes</td>
</tr>
</tbody>
</table>

Errors in statistical analyses distorting the evaluation of associations
# Types of Errors in scientific studies

<table>
<thead>
<tr>
<th>Type of error</th>
<th>How to prevent/reduce it</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sampling error</td>
<td>Increase sample size</td>
</tr>
<tr>
<td>Bias</td>
<td>Methodology</td>
</tr>
<tr>
<td></td>
<td>- RANDOMIZATION</td>
</tr>
<tr>
<td></td>
<td>- Masking (eg double blind)</td>
</tr>
<tr>
<td></td>
<td>- Intention to treat</td>
</tr>
</tbody>
</table>
Efficacy trial in a Rare Condition

CHOICE

Internal validity
Randomised Trial

Feasibility
Uncontrolled trial
Efficacy trial in a Rare Condition

CHOICE

Feasibility

Uncontrolled trial

Pro’s: More patients on new treatment (Activity, Toxicity) - Easier to recruit patients
Larger historical (control) group

Con’s: BIAS!?!?!
Efficacy trial in a Rare Condition

**CHOICE**

**Internal validity**

**Randomised Trial**

**Pro:** Unbiased estimates of treatment effects

**Con’s:** More difficult to enroll patients

**Less Patients on new treatment: Precision?**
Efficacy trial in a VERY Rare Condition

**CHOICE**

- **Internal validity**
  - Randomised Trial
  - Random error

- **Feasibility**
  - Uncontrolled trial
  - Bias
Note

• Sampling error and Bias are independent

• Increasing sample size is harmful in the presence of bias (more confidence in a wrong result)

• Statistics deals (mainly) with sampling error but provides little help with bias
Study Design in rare cancers

- *Uncontrolled trial/Historical controls*
- Randomised Controls

**WHY NOT?**
RCT vs Uncontrolled trial in a rare cancer

Hypothesis: it is possible to recruit a group of patients in whom, in a reasonable time, you expect to observe 30 events

2 Options:

a) RCT (Randomization 1:1)

b) All patients receive the experimental therapy and are compared to an historical control group of similar size (+ 30 events)
RCT’s in rare cancers

• Loss of power /Precision

Available events : 30    HR = 0.5

RCT (30):          HR */÷ 2        95%  CL 0.24-1.02
Uncont. Trial     HR */÷ 1.66     95%  CL 0.3 - 0.83

(Histor. Controls = 30 x2)
RCT’s in rare cancers

- Loss of power /Precision

Available patients: 100

- RCT (50 x2): Difference +/- 15%
- Uncontrolled tr. (Histor. Controls): Difference +/- 11%
SUGGESTION

If the expected (or necessary) treatment effect is large but not outstanding, a RANDOMIZED CLINICAL TRIAL, IF ETHICALLY ACCEPTABLE, is the best way to assess it even in rare diseases.
RCT’s in rare cancers

Pro’s
- VALIDITY
- CREDIBILITY

Con’s
- Moderate loss in power
- Often no standard (untreated control group?)
  - Ethics?
  - Acceptance?
Soft Tissue Sarcomas

• Adjuvant-Neoadjuvant therapy
There is no consensus on the current role of adjuvant chemotherapy. Study results are conflicting, in the presence of negative results from the largest studies, though data are available from smaller studies suggesting that it might improve, or at least delay, distant and local recurrence in high-risk patients [13, 14]. A meta-analysis found a statistically significant limited benefit in terms of both survival- and relapse-free survival [15]. It is unknown whether adjuvant chemotherapy may be particularly beneficial in specific subgroups or even detrimental in others. Therefore, adjuvant chemotherapy is not standard treatment in adult-type STS. It can be proposed as an option to the high-risk individual patient (high-grade, deep, >5 cm tumour) for a shared decision-making with the patient [II, C] or within
A Systematic Meta-Analysis of Randomized Controlled Trials of Adjuvant Chemotherapy for Localized Resectable Soft-Tissue Sarcoma

Nabeel Pervaiz¹
Nigel Colterjohn, MD²
Forough Farrokhyar, MPH, PhD²
Richard Tozer, MD, PhD³
Alvaro Figueroedo, MD³
Michelle Ghert, MD²

¹ University of Waterloo, Waterloo, Ontario, Canada.
² BC Cancer, Vancouver, British Columbia, Canada.
³ Sunnybrook Health Sciences Centre, Hamilton, Ontario, Canada.

Absolute Risk Reductions and 95% Confidence Intervals for Local Recurrence, Distant Recurrence, Overall Recurrence, and Survival

<table>
<thead>
<tr>
<th>Local recurrence</th>
<th>Distant recurrence</th>
<th>Overall recurrence</th>
<th>Survival</th>
</tr>
</thead>
</table>

ABS indicates absolute risk reduction, 95% CI, 95% confidence interval.

5-10% absolute improvement in OS

KEYWORDS: chemotherapy, soft tissue, sarcoma, adjuvant, meta-analysis, randomized controlled trial, localized, resectable.
ISG – STS 1001

histology-tailored chemo x 3 → Surgery ± RT

MLPS: Trabectedin
LMS: GEM + DTIC
UPS: GEM + TAX
Synovial Sa: HD-IFX
MPNST: IFX + ETO

epiADM+IFX x 3 → Surgery ± RT

- high grade
- deep seated
- $\geq 5$ cm

Do not duplicate or distribute without permission from author and ESO
3rd futility analysis on Group 1 – 286 patients

70 events

Table 1. EUROSARC: Relapse free survival (RFS) by treatment arm.

<table>
<thead>
<tr>
<th>Treatment ARM</th>
<th>Total N</th>
<th>N of Events</th>
<th>N Censored</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard</td>
<td>144</td>
<td>25 (17.4%)</td>
<td>119 (82.6%)</td>
</tr>
<tr>
<td>Tailored</td>
<td>142</td>
<td>45 (37.1%)</td>
<td>97 (68.3%)</td>
</tr>
<tr>
<td>Overall</td>
<td>286</td>
<td>70 (24.5%)</td>
<td>216 (75.5%)</td>
</tr>
</tbody>
</table>

Chi-Square  
p
Log Rank (Mantel-Cox) 8.166 0.004

The Kaplan-Meier RFS probability at 46 months was 0.62 and 0.38 (log rank p=0.004) in the Standard and in the Tailored arm, respectively.
## Relapse Free Survival

![Graph showing cumulative survival over time](image)

**Table 2.** EUROSARC: RFS - Cox’s univariate HR and its 95% Confidence Intervals

<table>
<thead>
<tr>
<th>Treatment ARM</th>
<th>HR</th>
<th>95% CI</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard</td>
<td>1 (ref.)</td>
<td>-</td>
<td>0.007</td>
</tr>
<tr>
<td>Tailored</td>
<td>1.955</td>
<td>1.119-3.190</td>
<td></td>
</tr>
</tbody>
</table>

Median FU: 12.34 months (IQrange: 25.45)
3rd futility analysis on Group 1 – 286 patients

21 deaths

Table 3. EUROSARC: Overall Survival (OS) by treatment arm.

<table>
<thead>
<tr>
<th>Treatment ARM</th>
<th>Total N</th>
<th>N of Events</th>
<th>N Censored</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard</td>
<td>144</td>
<td>6 (4.2%)</td>
<td>138 (95.8%)</td>
</tr>
<tr>
<td>Tailored</td>
<td>142</td>
<td>15 (11.6%)</td>
<td>127 (89.4%)</td>
</tr>
<tr>
<td>Overall</td>
<td>286</td>
<td>21 (7.3%)</td>
<td>265 (92.7%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chi-Square</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log Rank (Mantel-Cox)</td>
<td>4.524</td>
</tr>
</tbody>
</table>

The Kaplan-Meier RFS probability at 46 months was 0.89 and 0.64 (log rank p=0.033) in the Standard and in the Tailored arm, respectively.
# Overall Survival

## Median FU: 12.34 months (IQrange: 25.45)

P = 0.033

<table>
<thead>
<tr>
<th>Treatment ARM</th>
<th>HR</th>
<th>95% CI</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard</td>
<td>1 (ref.)</td>
<td>-</td>
<td>0.034</td>
</tr>
<tr>
<td>Tailored</td>
<td>2.687</td>
<td>1.104-6.937</td>
<td></td>
</tr>
<tr>
<td>Hazard Ratio</td>
<td>HR (95% CL)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>----------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.25</td>
<td>1.03 (0.24 - 4.39)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>2.28 (0.27 - 12.66)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2.38 (0.69 - 8.19)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1.85 (0.65 - 5.22)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>2.17 (0.98 - 4.80)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.95 (1.12 - 3.19)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Histology:
- Malignant peripheral nerve sheath tumor
- Synovial sarcoma
- Undifferentiated pleomorphic sarcoma
- All
Myxoid Liposarcoma

Phase II clinical trial of neoadjuvant trabectedin in patients with advanced localized myxoid liposarcoma


Background: To evaluate neoadjuvant trabectedin (1.5 mg/m² 34 h i.v. infusion every 3 weeks) in patients with locally advanced myxoid liposarcoma (ML) previously untreated with chemotherapy or radiation.

Patients and methods: Primary efficacy end point was pathological complete response (pCR) or tumor regression rate. Objective response according to RECIST v1.1 was a secondary end point.

Results: Three of 23 assessable patients had pCR (13%; 95% confidence interval 8.9% to 23.4%). Furthermore, very good and moderate histological responses were observed in another 2 and 10 patients, respectively. Histological deciliation in the cellular and vascular tumor component and maturation of the tumor cells to low-grade fibrous tissue were observed in both myxoid and myxoid/fibrous cell variants. Seven patients had partial responses according to RECIST objective response rate of 24%; 95% CI, 10% to 44%; no disease progression was reported. Neoadjuvant treatment was usually well tolerated, with a safety profile similar to that described in patients with soft tissue sarcomas of other tumor types.

Conclusion: Trabectedin 1.5 mg/m² given as a 24 h i.v. infusion every 3 weeks is a promising option in the neoadjuvant setting of ML.

Key words: chemotherapy, liposarcoma, neoadjuvant, soft tissue sarcoma, surgery.

Introduction

Myxoid liposarcoma (ML) accounts for one-third of liposarcomas [1] and represents a morphological continuum encompassing myxoid and myxofibroid cellularity, the latter characterized by an extent of sound cell composition >50%. Molecular hallmark of ML is the presence of FUS/CHOP fusion proteins.

Surgery and radiation therapy are the standard therapy for ML, but few data are available on neoadjuvant treatment. In retrospective studies on neoadjuvant treatment of patients with high-grade primary liposarcoma [3] or ML [4], oligometastases disease-specific survival was provided. Despite palliative and pathologically documented response was not given. Favourous results of preoperative trabectedin in a cohort of advanced MAs, in which two patients with multifocal and local tumor masses, respectively, showed regression of 50% of their tumors, with adjuventic differentiation observed in one of them [5], prompted the conclusions of the present clinical trial.

Further molecular characterization showed type I, II and IV FUS-CHOP fusion proteins in patients responding to trabectedin treatment.

The aim of this exploratory Phase II clinical trial was to evaluate the efficacy and safety of neoadjuvant trabectedin in patients with locally advanced ML previously untreated with chemotherapy or radiation. The primary end point was the achievement of pathological complete remission. Response rate by the Response Evaluation Criteria in Solid Tumors (RECIST, v1.1) and changes in tumor density were also analysed.

Patients and methods

Patients were recruited at eight investigational sites of France (n = 3), Germany (n = 1), Italy (n = 2), and the United States (n = 2). The study protocol was approved by the Independent Local Ethics Committee of each participating center and was conducted in accordance with the Declaration of Helsinki, Good Clinical Practice guidelines and local regulations on clinical trials. Signed informed consent was obtained prior to all patients to any study-specific procedure.
What to do?

1. Stop the study? YES!
2. Publish the results? YES!
3. Use routinely A+I as adjuvant tx in STS? YES!
3. And in Mixoid Liposarcomas?
   - A+I (ignore subgroup analyses)
   - Trabectidin (less toxic, ≈ efficacy?)
   - RANDOM (A+I vs T): Noninferiority, undersized trial with a Bayesian design

Offer randomization until you konw better!
QUESTION

• Are there situations in which we CAN forego (avoid) the Randomized Controlled Trial?

Good Reasons?

- Sample Size?
- Organization?
- Patient Consent?
- Ethical Reasons?
BETTER QUESTION

• Are there situations in which we **MUST** forego (avoid) the Randomized Controlled Trial?

Good Reasons?

- *Sample Size*?
- *Organization*?
- *Patient Consent*?
- *Ethical Reasons*?
Randomization: acceptable?

- Prognosis with standard therapy (no therapy)
- *Toxicity of experimental therapy*
- Plausible Efficacy of Experimental Therapy
  - Uncontrolled Trials in same cancer
  - RCT in different stages same cancer
  - RCT’s in other cancers with similar biology
  - *(dramatic effects in other cancers)*
If a new drug…
- with a well-identified molecular target
- present in different tumors
- shows strong beneficial clinical effects in one of these (usually the most frequent one)

...is it always necessary/acceptable/possible to conduct a RANDOMISED TRIAL OF ADEQUATE SIZE in each one of the other tumors?
Imatinib

CML -> Large RCT

GIST -> Large uncontrolled trial

Other rare indications -> Case Series
(dermatofibrosarcoma protuberans, plexiform neurofibromas, chordomas)
Anti-B-RAF + Anti-MEK in BRAF- Positive tumors

Melanoma -> Large RCT’s

NSCLC BRAF+?
JUNE 22, 2017

- Combination of dabrafenib and trametinib gained approval from the Food and Drug Administration (FDA) for the treatment of pts BRAF V600+ metastatic NSCLC.

- That approval was based on results from a three-cohort, multicenter, nonrandomized, open-label study of patients with stage IV NSCLC.

- In this phase 2 study, 36 untreated pts and 57 pre-treated pts were assigned to ..... Investigator-assessed objective response rate was the primary endpoint.

- At a median follow-up of nine months, the ORR was 61.1 percent in the treatment-naïve group, and 68 percent of patients did not show progression.
Anti-B-RAF + Anti-MEK in BRAF- Positive tumors

Melanoma -> Large RCT’s

NSCLC BRAF+ -> FDA approved
(1-arm trial?)

Patient with a BRAF+ c. in another site?

Do not duplicate or distribute without permission from author and ESO.
New paths to drug use

Large RCT in a frequent cancer with the target - Proof of principle – Toxicity

Uncontrolled (but formal) trial(s) in other (rare) cancers with the target

Off label use in individual cases with the target

?
New paths to drug use

Large RCT in a frequent cancer with the target - Proof of principle – Toxicity

Uncontrolled (but formal) trial(s) in other cancers with the target

Off label use in individual cases with the target

Acceptable? Methodology?
Uncontrolled trials
Historical Controls

– Well Known Biases
– May be sufficient if outstanding benefit
– Necessary if control group unethical

Careful and transparent methodology
Need of guidelines/research
PRUDENCE!!!
4 issues

• Phase II or phase III?

• Randomised or Uncontrolled?

• Endpoints
Endpoints in cancer trials

- “True” endpoints
  - OS
  - QoL scores

- Surrogate endpoints
  - Response rate
  - Progression free survival
  - Others
Surrogate Endpoints in trials in rare cancers

- None has been validated (impossible)
- Response
  - Objective-Reproducible
  - Consistently associated with a clinical benefit in solid tumors (even without control group)
- PFS
  - Sensitive to type/timing of assessments
  - Meaningless without a control group (historical?)
Surrogate Endpoints in trials in rare cancers

Acceptable only if the new treatment is associated with dramatic changes in (long term?) prognosis
4 issues

• *Phase II or phase III?*

• *Randomised or Uncontrolled?*

• *Endpoints*

• *Conventional or unorthodox statistics?*
FREQUENTIST PROBABILITY

The expected frequency of the observation given a hypothesis is the probability (P) of the observed trial results if the experimental therapy does not work.
Statistical Foundations of RCT

1. Starting Hypothesis (H0): The 2 treatments being compared are “identical”
2. Randomizations ensures that differences between groups are due to chance
3. Double blind prevents bias in assessment
4. Are observed results compatible with H0?
FREQUENTIST TEST OF HYPOTHESIS

If it is not a rare event

$P = \text{Expected frequency of the observation given a hypothesis}$

If it is too rare

REJECT THE HYPOTHESIS
Advancement of knowledge in Medicine (conventional statistics)

• Dominant theory is true (=standard therapy is better) until sufficient evidence becomes available against it

• To this purpose, only evidence collected within one or more trials aimed at falsifying it can be used

• No use of
  – External evidence
  – Evidence in favor of…
The 2 Reasons why large numbers of patients are needed in clinical trials

- Outstanding efficacy seldom observed
- Any knowledge outside the primary analysis of the clinical trial is ignored in the design and analysis of the trial
What can be done?

Recent developments
- *Surrogate endpoints*
- *New types of systematic reviews*
- *Adaptive trials*
- *Bayesian Statistics*
Common beliefs

Frequentist probability
- Objective
- «Hard»
- Useful to analyse experiments
- Scientific

Bayesian Probability
- Subjective
- «Soft»
- Inappropriate to analyse experiments
- Not scientific
Differences between Conventional (Frequentist) and Bayesian Statistics

- Meaning of probability
- Use of prior evidence
Frequentist Probability

Probability of an observation
(given a hypothesis)

Bayesian Probability

Probability that a hypothesis is true
(given observation and prior knowledge)
Frequentist Probability

Probability of the observed difference (if the experimental therapy does not work)

Bayesian Probability

Probability that the experimental therapy works/doesn’t work (given observed difference and prior knowledge)
Hypothetical Example

• As a statistician, I’m asked to design 2 separate trials in the same rare disease, squamous gastric cancer (no standard treat.)

• Study A:
  Experimental therapy: Radiochemotherapy
  – Effective in squamous cancers of other sites

• Study B

• Experimental therapy: Intercessory prayer
Squamous gastric cancer

Planning a trial of RT+CTX

Same Numbers, Same Interpretation of the results (p value)

Analysing its results

Intercessory prayer
Squamous gastric cancer

Results of the trials

RT+CTX

20% reduction in deaths \( P = 0.06 \)

Intercessory prayer

20% reduction in deaths \( = 0.06 \)

Treatment x next patient with SGC?
Frequentist P

Probability of the observed difference if either therapy does not work = 6%

Bayesian Probability

Probability that either therapy works a lot/works a little/does not work?

Is it the same for the two treatments?
Differences between Conventional and Bayesian Approaches

- *Meaning of probability*

- *Use of prior evidence*
Conventional P

Probability of the observed difference (if the experimental therapy does not work)

Bayesian Probability

Probability that the experimental therapy works/doesn’t work (given observed difference and prior knowledge)
Example

Tumor X Nil vs A 15% vs 10%
N=2000 P = 0.0001

H0 Rejected: A is effective in X
Example

**Mortality**

<table>
<thead>
<tr>
<th>Tumor</th>
<th>Nil vs A</th>
<th>15% vs 10%</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>N=2000</td>
<td>( P = 0.0001 )</td>
</tr>
<tr>
<td>Y</td>
<td>N=240</td>
<td>( P=0.066 )</td>
</tr>
</tbody>
</table>

H0 not rejected: A not shown effective in y
Prior Information:
X and Y are BRAF+
A = Anti BRAF

Mortality

<table>
<thead>
<tr>
<th>Tumor</th>
<th>Treatment</th>
<th>Mortality</th>
<th>N</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tumor X</td>
<td>Nil vs A</td>
<td>15% vs 10%</td>
<td>2000</td>
<td>0.0001</td>
</tr>
<tr>
<td>Tumor Y</td>
<td>Nil vs A</td>
<td>15% vs 7.5%</td>
<td>240</td>
<td>0.066</td>
</tr>
</tbody>
</table>

**INTERPRETATION?**
Interpretation of the two trials

CONVENTIONAL
Tumor X: $P = 0.0001$
Tumor Y: $P = 0.066$
Efficacy of treatment A proven in X
undemonstrated in Y
Interpretation of the two trials

CONVENTIONAL

Efficacy of treatment A is proven in X, undemonstrated in Y

BAYESIAN

(Posterior) Probability that treatment A significantly (HR<0.8) lowers mortality in tumor X: 90%
in tumor Y: 90%
Disadvantages of Bayesian Statistics

- It is (felt as)
  - Subjective
  - Arbitrary
  - Amenable to manipulations
    (pharma companies?)
Conceptual Advantages of Bayesian Statistics

- Reflects human reasoning ("common sense")
- It is focused on estimates of effect
- Provides a conceptual framework for medical decision making
- **IT IS TRANSPARENT**
Practical Advantages of Bayesian Statistics in rare tumors

1. No need to set the sample size in advance
   Adaptive designs: enrol patients until sufficient evidence in favour or against efficacy

2. When strong a priori evidence is available and trial results are in agreement with it
   Smaller sample size is necessary – You can stop any time
Prior evidence in Bayesian statistics

**Note:** The difference between Bayesian and conventional statistics decreases with increasing strength of the empirical evidence

**Rare Tumors!**
Sources of prior evidence

- *Randomised Trials*
- Biological & Preclinical Studies
- Case-reports
- Uncontrolled studies
- Studies with surrogate endpoints
- Studies on other similar cancers
- Studies on the same cancer in different stages
- Others?
Differences between the present and the proposed approach

• Present:
  – Rational but informal integration of the available knowledge

• Proposed (Bayesian):
  – Formal, explicit and quantitative integration of the available knowledge
    • Verifiable quantitative methods
    • Sensitivity analyses
    • Focus on summary effect estimates
Final Considerations

• Medical decision is ALWAYS Bayesian (based on perceived probability of benefit)
• Perceived probability of benefit = experimental evidence + prior knowledge
• In rare tumors, less experimental evidence - -> more uncertainty -> Prior knowledge more important
Final Considerations

NEW EBM

Direct evidence
+ Formal use of external evidence
+ Transparency

ACCEPTANCE OF UNCERTAINTY!
Special issue of *The American Statistician* dedicated to criticising the use of P-value as a research tool (in all fields)
Contents

How to build a world without “P”?
1. “Don’t” is not enough
2. Don’t Say “Statistically Significant”
3. There Are Many Do’s
4. Editorial, Educational and Other Institutional Practices Will Have to Change
5. It Is Going to Take Work, and It Is Going to Take Time
6. Why Will Change Finally Happen Now?
3. There Are Many Do’s

3.1. Accept Uncertainty

Uncertainty exists everywhere in research. Significance tests and dichotomized $p$-values have turned many researchers into scientific snowbirds, trying to avoid dealing with uncertainty by escaping to a “happy place” where results are either statistically significant or not. In the real world, data provide a noisy signal. Variation, one of the causes of uncertainty, is everywhere. Exact replication is difficult to achieve. So it is time to …“move toward a greater acceptance of uncertainty and embracing of variation”
3. There Are Many Do’s

3.1. Accept Uncertainty

Statistical methods do not rid data of their uncertainty. “Statistics is often sold as a sort of alchemy that transmutes randomness into certainty, an ‘uncertainty laundering’ that begins with data and concludes with success as measured by statistical significance.”
“...will precision medicine usher in an age of diagnostic and prognostic certainty?”

“The new tools for tailoring treatment will demand a greater tolerance of uncertainty and greater facility for calculating and interpreting probabilities than we have been used to as physicians and patients”
Rare Cancers Europe (RCE) methodological recommendations for clinical studies in rare cancers: a European consensus position paper

P. G. Casali¹*, P. Bruzzi², J. Bogaerts³ & J.-Y. Blay⁴ on behalf of the Rare Cancers Europe (RCE) Consensus Panel
Methodological recommendations for clinical studies in rare cancers

‘alternative ways to conceive study design, analysis of data and combination of results would be exceedingly important.

It is possible that some innovative solutions may imply a price to pay in terms of a higher uncertainty.’
Evidence Based Medicine in **Rare Tumors**

The integration of **best research evidence** with clinical expertise and patient values = **Available knowledge**

**Decision Theory** (Probability x Utility)

**What is available**