Survival extrapolation: How to draw the lines beyond the observed data, *if you must*

Nikos Demiris, Assistant Professor, Athens University of Economics and Business, Greece

Linda Sharples, Professor in Statistics London, School of Hygiene & Tropical Medicine, UK

Alexandros Sagkriotis, Director RWE, Novartis, Basel, Switzerland

Andreas Karabis, Principal, Real World Insights, IQVIA, Amsterdam, the Netherlands

ISPOR Europe 2018

*Barcelona, 13 November 2018*
Objective

The objective of this workshop is to present some methods for the inclusion of long-term data from sources other than the RCTs, that can be used to estimate parameters of times to events and other outcomes in health economic models.
Objective

Limitations of the standard extrapolation methods

The use of long-term data from various sources

Combining sources of evidence

Illustrative example

Take-home messages & Discussion
The problem: incomplete data
Mean LYG from incomplete data

The area between the curves represents LYG

Follow up

The area between the curves represents LYG

Observed LYG

Unobserved LYG

Extrapolation

0.0

0.2

0.4

0.6

0.8

1.0

0

10

20

30

40

50

60

Time (years)

Proportion surviving

Time (years)
Who has had the need to extrapolate survival beyond the observed data?
Extrapolating state-transition models

- Estimate transition rates $\lambda(t; x)$ from short-term trial data
- Run model until all patients are in one of two death states
Use of parametric survival functions

Traditional approach: assume parametric survival

Short-term fit versus long-term plausibility
Use of parametric survival functions

Parametric approach acceptable if,
- long-term follow up,
- well fitting parametric function throughout.

NB: All assumptions about future survival and treatment effects are **untestable**
Types of external data

- **Randomised controlled trials**: Control group risk, intervention effect
- **Disease registries, cohorts**: Baseline risk, intervention effect?
- **Population life tables**: Background risk

**Extrapolate for:**
- Baseline risk
- Treatment effect
- Time horizon?

**Degree of patient selection**

- **Cohort/disease registry**: closer to patients of interest, longer follow up, but not usually lifetime
- **Population life tables**: lifetime horizon, but requires more adjustment to represent patients of interest
RWE and its role in extrapolation
What is real world evidence?

RWE is based on data collected outside of clinical trials that is used to support decision-making.

RWE aims to understand the differences and needs of real patients to improve clinical care, safety, and access to medicines.
Generating RWE from RWD

RWE can be generated by actively collecting new data (primary source) or by analyzing existing data (secondary source).

Real world data sources

- Administrative Claims Databases
- Hospital Data
- Pragmatic Clinical Trials
- Electronic Medical Records
- Surveys
- Disease & Product Registries
- Observational Studies
- Social Media

Data collection can be either:
- Prospective
- Retrospective
- Cross-sectional

External data can be used in extrapolations beyond the time horizon of RCTs.
A **patient registry** is an organized system that uses observational study methods to collect uniform data (clinical and other) to evaluate specified outcomes for a population defined by a particular disease, condition, or exposure, and that serves a predetermined scientific, clinical, or policy purpose(s).

Increasing demand for RWE along the product lifecycle
Review of NICE appraisals of pharmaceuticals from 2000 to 2016

489 individual pharmaceutical technologies assessed by NICE

22 (4%) used non-RCT data to estimate comparative clinical effectiveness

Methods for establishing external controls: published trials - 13 (59%), observational data - 6 (27%), expert opinion - 2 (9%), and responder vs non-responder analysis - 1 (5%)

Positive recommendation received from NICE
- decision based on RCT data – 83% of cases
- decision based on non-RCT data – 86% of cases

The various methods used highlight the need to establish clear guidance

Regulatory & payer shift towards RWE

There has been an increase in formal requests from European payers for more detailed real-life information to:
- prove that a treatment works effectively, efficiently and is safe
- manage uncertainty when payers are making reimbursement decisions

Call for **FDA to publish draft guidance** on:
- RWE benefit to patients, regulators and industry
- RWE availability, quality and access challenges as well as mitigating strategies
- Methods for collection, analysis and communication of RWE
- Contexts for use of RWE in regulatory decision making

The new 21st Century Cures Act requires FDA to explore the use of RWE to satisfy post-approval requirements and support indication expansion.

Use of RWE in HTA submissions/ economic models is growing

Example: Use of life tables to capture survival beyond the end of the RCT

Economic experts recommended *use of life tables in order to capture survival beyond the end of RCT* in Acute Heart Failure.

“NICE understands, from the response to clarification questions, that UK life tables based on published literature will be used to capture survival beyond day 60.

Experts note that these may not be easily available, although such data may be available from the National Heart Failure Audit (run by the Information Centre).

**These tables must accurately reflect the patient population of interest** (specifically, those patients who have survived the initial 60 days), as expert opinion is that survival past 60 days is very dependent on the baseline characteristics of the patients.”
Assumptions of proposed methodology

**Assumption of extrapolating survival from randomized trials using external data is that the RCT and RW populations are comparable.**

However, is that true in real life?
In a Novartis non-interventional study an attempt was made to identify a trial-like population in the real world using criteria from the confirmatory program (pivotal trials).

**Overall Study Cohort**
N = 98,821

**Cases included**
N = 46,091

**Cases excluded**
N = 14,915

**Trial-like population in the real world**
N = 31,177
32% of overall cohort

**Major limitation:** very few of the In-/Exclusion criteria of the RCT could be mapped, since information that is usually collected in RWD sources is very minimalistic.
Difference between efficacy and effectiveness

EFFICACY
Does it work in a clinical trial / under ideal conditions?

EFFECTIVENESS
Does it work in real life?
Factors that Drive Effectiveness

**Health system care delivery factors**

**Drug use factors**
- Patterns of use, dose, treatment duration
- Adherence of prescribers to label recommendations
- Adherence of patients to prescriptions
- Past history of exposure

**Patient population factors**
- Patient physical and behavioral characteristics: age, gender, weight, ethnicity, smoking/eating/exercise habits, etc.
- Co-morbidities
- Disease stage/severity
- Co-prescriptions
- Other baseline risk factors and genetics relevant to the disease/drug

**Efficacy**

**Effectiveness**

**Interaction**
Have you ever used external data to extrapolate survival beyond the observed data?
Combining evidence sources
Types of external data

- **Cohort/disease registry**: closer to patients of interest, longer follow up, but not usually lifetime
- **Population life tables**: lifetime horizon, but requires more adjustment to represent patients of interest
Mean LYG from incomplete data
Extrapolating from incomplete data
Extrapolating from incomplete data
Extrapolating from incomplete data

Suppose that we have short-term information on the survivor functions $S_{new}(t)$ and $S_{control}(t)$ (e.g. from an RCT).

Further we have short and long-term information on $S_{external}(t)$ (e.g. from registries or population statistics).

Then use external data to estimate long-term effects by making assumptions about:

1. how survival differs between the RCT and external populations ($\beta$)
2. and how observed comparison of $S_{new}(t)$ and $S_{control}(t)$ will continue in the long term ($\theta$)
Hazard options

- Proportional hazards
- Additive hazards
- Converging hazards
- Proportional cause-specific hazards
General framework reviewed

- Assumed relation between disease and external population mortality ($\beta$)
- Assumed relation between treated and control population mortality ($\theta$)
- Options for systematic long-term difference between external and disease population
- Choice of functional form for hazard or survival
Minimum requirements

1. **Robust treatment effects**: a treatment effect size that is precise (large trial, meta-analysis), from a relevant population, with at least medium-term follow up

2. **External evidence**: long-term, complete and high quality (trial standard?) records from a relevant population (disease registry, population?)

3. **Synthesis parameters**: a minimum set of parameters that model the relationships between data sources ($\theta, \beta$)

4. **Model assumptions**: well-fitting survival functions and ‘plausible’ (but non-verifiable) extrapolation

5. **Additionally**: if individual patient data are available, matching or standardisation to decision-problem target population

6. **Sensitivity analysis**
Application to Implantable Cardioverter Debrillators (secondary prevention)
ICDs for secondary prevention of SCD
ICDs for secondary prevention

ICD implanted

ICD repair/battery change admission

ICD replacement admission

Arrhythmia admission

Arrhythmia-related death

Non-arrhythmia-related death
ICD data sources

ICD implant compared to anti-arrhythmia drugs (AAD) for secondary prevention of sudden cardiac death

Data:
1. Meta-analysis of three (non-UK) RCTs (published HRs).
   - Overall HR(ICD:AAD) = 0.72 (0.60, 0.87)
   - Arrhythmia-related deaths HR = 0.50 (0.37, 0.67)
2. Individual patient data from cohort of 535 UK cardiac arrhythmia patients implanted with ICDs.
   - Relatively short-term follow up: approximately 75% patients followed for less than 5 years, maximum 10 years
3. UK population mortality statistics

Aim:
- Estimate cost-effectiveness over the lifetime of ICD and AAD patients in the UK.
Application to ICD

Hazards for three groups:

\[ h_{UK}(t) \]  
Estimated from UK Government audit data

\[ h_{ICD}(t) = e^\beta h_{UK}(t) \]  
UK Government and registry data

\[ h_{AAD}(t) = e^{\beta+\theta} h_{UK}(t) \]  
UK Government, registry and meta-analysis results

The (UK population) baseline hazard can take non-parametric or parametric form (e.g. Weibull)
ICDs graphical model

\[ \Lambda, \beta \]

\[ \gamma(t) \]

\[ S(t) \]

\[ c \]

\[ LYG \]

\[ \theta \]

Registry \( \{t_i, \delta_i\} \)

Meta-analysis of ICD:AAD trials

Prior parameters

UK survival

\[ S(t) \]

\[ \theta \]
ICDs graphical model

\[
\Lambda, \beta, \gamma(t), c, \delta_i \quad S(t) \quad \theta
\]

UK survival

Registry \{t_i, \delta_i\}

Meta-analysis of ICD:AAD trials

Prior parameters

LYG
ICDs graphical model

\[ \Lambda, \beta, t_i, \delta_i, \theta, \gamma(t), S(t), \text{ registry, ICD:AAD trials, UK survival, LYG, } c, \text{ prior parameters} \]
ICDs graphical model

ICDs graphical model
ICDs graphical model

\[ \Lambda, \beta \]

\[ \gamma(t) \]

\[ S(t) \]

\[ \theta \]

\[ c \]

\[ \text{UK survival} \]

\[ \text{Registry \ \{t_i, \delta_i\}} \]

\[ \text{Meta-analysis of ICD:AAD trials} \]

\[ \text{Prior parameters} \]

- \( c \) is a ‘tuning parameter’
- governs relative importance of trial and external evidence

\[ S(t) \]

\[ \text{LYG} \]
ICDs mean LYG by distribution

<table>
<thead>
<tr>
<th>Model</th>
<th>LYG (SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weibull</td>
<td>1.8 (0.5)</td>
</tr>
<tr>
<td>Cox</td>
<td>2.0 (0.5)</td>
</tr>
<tr>
<td>Additive (c=100)</td>
<td>1.7 (0.4)</td>
</tr>
</tbody>
</table>

![Graph showing LYG means for different models](image-url)
Does cause-specific hazard change as a proportion of total hazards?

Use cause-specific hazards for three groups:

\[ h_{UK.A}(t) + h_{UK.NA}(t) \]  
UK Government audit data

\[ h_{ICD}(t) = e^{\beta'} h_{UK.A}(t) + h_{UK.NA}(t) \]  
UK Government and registry data

\[ h_{AAD}(t) = e^{\beta'+\theta'} h_{UK.A}(t) + h_{UK.NA}(t) \]  
UK Government, registry and meta-analysis of cause-specific hazards
ICDs mean LYG by distribution

<table>
<thead>
<tr>
<th>Model</th>
<th>Women LYG (SE)</th>
<th>Men LYG (SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weibull</td>
<td>1.9 (0.6)</td>
<td>1.7 (0.5)</td>
</tr>
<tr>
<td>Cox-like</td>
<td>2.0 (0.7)</td>
<td>1.9 (0.5)</td>
</tr>
<tr>
<td>Poly-Weibull</td>
<td>3.1 (0.8)</td>
<td>2.9 (0.6)</td>
</tr>
</tbody>
</table>
Minimum requirements

1. **Robust treatment effects**: a treatment effect size that is precise (large trial, meta-analysis), from a relevant population, with at least medium-term follow up

2. **External evidence**: long-term, complete and high quality (trial standard?) records from a relevant population (disease registry, population?)

3. **Synthesis parameters**: a minimum set of parameters that model the relationships between data sources \((\theta, \beta)\)

4. **Model assumptions**: well-fitting survival functions and ‘plausible’ (but non-verifiable) extrapolation

5. **Additionally**: if individual patient data are available, matching or standardisation to decision-problem target population

6. **Sensitivity analysis**
Which data sources are you thinking to use in order to leverage on RWD?
How organized the RWD sources are to be used for extrapolation purposes?
In absence of systematically captured longitudinal RWD, which are the actions you take?
Have you ever considered running pivotal trials as more pragmatic in order to generate clinical evidence that is closer to the real population?
Thank you!

- **Nikolaos Demiriris**, Assistant Professor, Athens University of Economics and Business, Greece
- **Linda Sharples**, Professor in Statistics London, School of Hygiene & Tropical Medicine, UK
- **Alexandros Sagkriotis**, Director RWE, Novartis, Basel, Switzerland
- **Andreas Karabis**, Principal, Real World Insights, IQVIA, Amsterdam, the Netherlands
Back-up slides
FDA Real World Evidence meeting- Key highlights

FDA is open to RWE studies (data generated from sources that are outside of RCTs that are randomized or observational)

The Agency finds RWE to be valuable in order to close the gap on the external validity from RCTs

FDA encourages recruitment of patients from within the healthcare system to increase generalizability

The Agency does not see a major distinction between RCT and pragmatic/randomized trials

Some of the key caveats that FDA believes should be addressed in RWE:

- Loss of power or diminished effect size in RWE trials and identifying ways to counteract this with enrichment factors (after randomization)
- Control for bias and unmeasured confounding
- Data Quality – sources, collection, analysis and validation
- Additional methods to establish causal inference in RWE
- Loss to follow up when patients switch insurance plans

FDA is open to exploring RWE designs and methodologies to ensure key aspects are incorporated and encourages inclusion of RWE into clinical development programs