The use of Historical Control Data to Assess the Benefits of New Therapies: A Case Study of Blinatumomab versus Standard Therapy of Adults relapsed/refractory Acute Lymphoblastic Leukaemia

Maurille Feudjo Tepie
Director, Observational Research, Amgen Ltd

Joint EMA – EUROPABIO Workshop
London, 22nd, November 2016
Outline

- Background

- Historical control group
  - Sources of data, analysis methods, results

- Conclusion/discussion
  - Challenges, lessons learned,
Background: ALL Disease

- Adult acute lymphoblastic leukemia (ALL):
  - Rare disease (~ 1-2/100,000 age-adjusted incidence rate among adults)
  - Large percentage of adult patients relapse after initial treatment
  - Very poor prognosis (1 year survival ~ 15% among relapsed/refractory (R/R) patients)

- Prognosis of R/R ALL is strongly impacted by:
  - Time to relapse (or duration of remission)
  - Number of previous relapses and salvage treatments
  - History of HSCT
Background: ALL Treatment Options

• No established standard treatment for R/R ALL patients

• HSCT, considered a potentially curative option, is generally not available for older patients (> 60 years)

• Palliative care often the only treatment option for many adult R/R ALL patients:
  • intolerability to aggressive chemotherapy
  • lack of curative intent if HSCT unavailable

• Promising results for blinatumomab reported in initial Phase II trial – high remission rates in R/R ALL population –
Background: Challenge for a Phase 3 RCT

- **Rare disease** – recruitment, achieving sufficient sample size are challenging

- **Unmet medical need** - poor disease prognosis

- **Limited or no treatment options** – would be unethical to allocate patients to “standard of care”

- **Clinicians unwilling to participate** in these trials

- **Other design challenges with clinical trials:**
  - Subject retention
  - Cross-over

- **New therapy** – initial promise, might offer hope

- **Some control data better than no information** – To help put results into appropriate perspective/context
Potential Data Sources/ Data Availability that could help provide some context – Historical controls

- Several studies* reported data on clinical outcomes among adult patients with R/R ALL:
  - Appeared data were available and could be assembled into a larger study relatively quickly

- Summarizing the literature was limited because of significant variation on how data were reported:
  - Differences in treatment histories (e.g. # of prior salvage therapies)
  - Differences in patient subgroup categories: time to relapse, age etc.

- Need individual patients data

**Adult R/R ALL Historical Comparator Study:**

**Study Schema**

**Investigator Databases**
- **EU**  
  N=8
- **US**  
  N=3

**Pooled Historical Comparator Database**
- **Review Data**
- **Harmonize Data**
- **Create Variables**

**Study Endpoints**
- **Primary:**
  - CR
- **Secondary:**
  - OS
  - Duration of CR
  - Rate of HSCT

**Analysis Sets/Planned analysis**
- **Ph- Difficult to Treat Analysis Set**
  - Subgroup Analysis
  - Stratum-Adjusted Analysis

**Exploratory Analyses**
- **Ph- Late First Relapse Analysis Set**
  - Subgroup Analysis
- **Ph + Analysis Set**
  - Subgroup Analysis

**Inclusion criteria:**
- Patients with Ph- B-precursor relapsed or refractory ALL
- Age ≥ 18 years at relapse
- Initial ALL diagnosis in 1990 or later
- Experienced early relapse*, were refractory to prior treatments, or were in 2\(^{nd}\) or greater salvage
Analysis approach

• Direct comparison of endpoints
  • Overall
  • By subgroups

• Weighting endpoints on key characteristics to the clinical trial population

• Propensity score analyses
Results: Complete Remission as Defined by the Study Group (CRsg)

<table>
<thead>
<tr>
<th>Stratum</th>
<th>Age at Treatment</th>
<th>Prior lines of Treatment</th>
<th>n/N</th>
<th>Stratum % Observed</th>
<th>Stratum % Observed in Trial</th>
<th>CRsg Proportion (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>&lt;35</td>
<td>alloHSCT</td>
<td>14/48</td>
<td>6.9%</td>
<td>21.2%</td>
<td>0.29 (0.17, 0.44)</td>
</tr>
<tr>
<td>2</td>
<td>&lt;35</td>
<td>In 1st salvage</td>
<td>52/119</td>
<td>17.2%</td>
<td>5.3%</td>
<td>0.44 (0.35, 0.53)</td>
</tr>
<tr>
<td>3</td>
<td>&lt;35</td>
<td>In 2nd+ salvage</td>
<td>27/150</td>
<td>21.6%</td>
<td>21.2%</td>
<td>0.18 (0.12, 0.25)</td>
</tr>
<tr>
<td>4</td>
<td>&gt;=35</td>
<td>alloHSCT</td>
<td>11/41</td>
<td>5.9%</td>
<td>12.7%</td>
<td>0.27 (0.14, 0.43)</td>
</tr>
<tr>
<td>5</td>
<td>&gt;=35</td>
<td>In 1st salvage</td>
<td>57/187</td>
<td>27.0%</td>
<td>10.1%</td>
<td>0.30 (0.24, 0.38)</td>
</tr>
<tr>
<td>6</td>
<td>&gt;=35</td>
<td>In 2nd+ salvage</td>
<td>25/149</td>
<td>21.5%</td>
<td>29.6%</td>
<td>0.17 (0.11, 0.24)</td>
</tr>
</tbody>
</table>

Weighted estimate for historical data

0.24 (0.20, 0.27)

Clinical trial data

0.43 (0.36, 0.50) | 0.33 (0.27, 0.41)

n = number of patients achieving CRsg, N = number of patients evaluated for CRsg

- 1. CR/CRh* 2. CR
## Results: Median Overall Survival

<table>
<thead>
<tr>
<th>Stratum</th>
<th>Age at Treatment</th>
<th>Prior lines of Treatment</th>
<th>N</th>
<th>Stratum % Observed</th>
<th>Stratum % Observed in Trial</th>
<th>Median OS (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>&lt;35</td>
<td>alloHSCT</td>
<td>108</td>
<td>9.7%</td>
<td>21.2%</td>
<td>3.8 ( 2.9, 4.5)</td>
</tr>
<tr>
<td>2</td>
<td>&lt;35</td>
<td>In 1&lt;sup&gt;st&lt;/sup&gt; salvage</td>
<td>258</td>
<td>23.2%</td>
<td>5.3%</td>
<td>5.7 ( 4.9, 6.3)</td>
</tr>
<tr>
<td>3</td>
<td>&lt;35</td>
<td>In 2&lt;sup&gt;nd+&lt;/sup&gt; salvage</td>
<td>161</td>
<td>14.5%</td>
<td>21.2%</td>
<td>2.9 ( 2.3, 4.0)</td>
</tr>
<tr>
<td>4</td>
<td>&gt;=35</td>
<td>alloHSCT</td>
<td>79</td>
<td>7.1%</td>
<td>12.7%</td>
<td>4.0 ( 2.8, 4.7)</td>
</tr>
<tr>
<td>5</td>
<td>&gt;=35</td>
<td>In 1&lt;sup&gt;st&lt;/sup&gt; salvage</td>
<td>341</td>
<td>30.7%</td>
<td>10.1%</td>
<td>3.7 ( 3.2, 4.4)</td>
</tr>
<tr>
<td>6</td>
<td>&gt;=35</td>
<td>In 2&lt;sup&gt;nd+&lt;/sup&gt; salvage</td>
<td>165</td>
<td>14.8%</td>
<td>29.6%</td>
<td>2.2 ( 1.7, 2.9)</td>
</tr>
</tbody>
</table>

**Weighted estimate of historical data**

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weighted estimate of historical data</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3.3 ( 2.8, 3.6)</td>
</tr>
</tbody>
</table>

**Clinical trial data**

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clinical trial data*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6.1 (4.2, 7.5)</td>
</tr>
</tbody>
</table>

Strong evidence of higher odds of CR in the trial (treated) population compared to the ‘control’ population
Forest Plot of Hazard Ratios for Analyses of Overall Survival

IPTW=Inverse probability of treatment weighting. sIPTW=Stabilized inverse probability of treatment weighting.

Strong evidence of smaller hazard of death in the trial (treated) population compared to the ‘control’ population
In summary

• Faced with the challenge of an effective registrational phase 3 RCT, partly due to:
  • Rare disease, very poor prognosis, limited treatment options, clinician willingness to participate, etc…

• Systematically collected, carefully analyzed, historical individual R/R ALL patients data:
  • Showed strong and consistent benefit of treating R/R ALL patients with Blinatumomab compared to standard of care

• Evidence was deemed important to help accelerated approval of Blinatumomab for adults R/R ALL by the FDA
  • Among others, the robustness of the results and the importance of the effect size played key roles
  • Helped by the availability of good historical data and excellent collaboration between contributing investigators
In summary

• The TOWER study, a phase 3 randomized open label trial later confirmed these findings
  • An almost two-fold increased in median overall OS compared to SOC

• These results and their outcomes, further highlight the importance for all relevant stakeholders to continue to explore the potential role of the RWD in drug regulatory process.
  • Work was presented at a FDA symposium on how RWD can be used for faster regulatory approval

• In some situation, RWD can be used to enable faster delivery to the patients:
  • Considerable unmet medical need
  • New and very promising therapy
  • Clinicians unwilling to participate in these trials
  • Rare disease
Study Collaborators

- **Nicola Gokbuget**, Dieter Hoelzer
  - University Hospital, Goethe University, Frankfurt, Germany

- **Hagop Kantarjian, Susan O’Brien**
  - University of Texas, Houston, Texas, United States

- **Hervè Dombret**
  - Hôpital Saint-Louis, Paris, France

- **Jose-Maria Ribera**
  - ICO-Hospital Germans Trias I Pujol, Jose Carreras Research Institute, Barcelona, Spain

- **Adele K. Fielding**
  - UCL Cancer Institute, London, United Kingdom

- **Renato Bassan**
  - UOC Ematologia, Ospedale dell'Angelo, Mestre-Venezia, Italy

- **Sebastian Giebel**
  - Maria Sklodowska Curie Memorial Cancer Center and Institute of Oncology, Gliwice, Poland

- **Anjali Advani**
  - Cleveland Clinic, Cleveland, Ohio, United States

- **Michael Doubek**
  - University Hospital, Brno, Czech Republic

- **Giovanni Martinelli**
  - Policlinico S Orsola Istituto Seragnoli, Italy

- **Martha Wadleigh**
  - Dana Farber Cancer Institute, Boston, Massachusetts

- **Norbert Ifrah**
  - Center Hospitalier Universitaire, Angers, France

- **Mireia Morgades**
  - H. Germans Trias I Pujol, Barcelona, Spain

- **Jacob M Rowe**
  - Rambam Medical Center, Haifa, Israel

- **Victoria Chia, Aaron Katz, Michael Kelsh, Julia Steiglmaier**
  - Amgen

* Principal Investigator
Thank You!
Back UP
Particular Efforts to Minimize Bias

- At data collection stage – requested sites to provide all patients with R/R ALL – rather than having sites apply selection criteria
- Inclusion/exclusion criteria applied centrally across all data sets
- Study sites reflected centers of excellence for treatment of ALL
- Weighting, stratified, and propensity score analyses to make endpoints more comparable
- Variety of sensitivity analyses conducted in order to address assumptions
Strength/Limitations of the approach

- Availability of and access to external control data
- Data definitions – outcomes, exposure, covariates
- Study biases:
  - Selection
  - Confounding
  - Immortal Time
- Treatment differences: across time, geographic regions
- Heterogeneity
Propensity Score Analysis – Methods

• Propensity scores derived from logistic regression models considering available covariates

• Odds ratio (OR) for complete remission estimated from logistic regression models, using stabilized inverse probability treatment weighting (sIPTW)

• Hazard ratio (HR) for death estimated from Cox models, using inverse probability treatment weighting (IPTW)

• Sensitivity analysis conducted by:
  • Alternating weighting factors
  • Time period
  • Further model adjustments
Propensity Score Analysis

- Aim to create balance in baseline covariates between patients treated with blinatumomab and patients treated with standard of care (historical comparator)

- Covariates:
  - Age (years)
  - Sex (male, female)
  - Duration between most recent treatment and initial diagnosis
  - Region (USA, Europe)
  - Prior HSCT (yes, no)
  - Number of salvage therapies (1, 2, 3, and 4+)
  - Primary refractory and in/entering first salvage (yes, no)
  - Refractory to last salvage therapy (yes, no)
# Covariate balance before and after propensity score (PS) adjustments

<table>
<thead>
<tr>
<th></th>
<th>Before PS adjustments</th>
<th></th>
<th>p-value</th>
<th></th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Historical comparator</td>
<td>Blinatumomab</td>
<td></td>
<td>Historical comparator</td>
<td>Blinatumomab</td>
</tr>
<tr>
<td>Age, Mean (SD)</td>
<td>38 (14)</td>
<td>41 (17)</td>
<td>0.0018</td>
<td>38 (14)</td>
<td>36 (16)</td>
</tr>
<tr>
<td>Female, %</td>
<td>44%</td>
<td>37%</td>
<td>0.09</td>
<td>44%</td>
<td>38%</td>
</tr>
<tr>
<td>Duration since initial diagnosis in months, mean (SD)</td>
<td>11 (12)</td>
<td>24 (23)</td>
<td>&lt;0.0001</td>
<td>14 (17)</td>
<td>17 (17)</td>
</tr>
<tr>
<td>Region – Europe, %</td>
<td>83%</td>
<td>50%</td>
<td>&lt;0.0001</td>
<td>77%</td>
<td>77%</td>
</tr>
<tr>
<td>Prior alloHSCT, %</td>
<td>21%</td>
<td>34%</td>
<td>0.0003</td>
<td>23%</td>
<td>21%</td>
</tr>
<tr>
<td>Number of salvage therapies, mean (SD)</td>
<td>1.5 (0.8)</td>
<td>2.3 (1.0)</td>
<td>&lt;0.0001</td>
<td>1.6 (0.9)</td>
<td>1.7 (0.9)</td>
</tr>
<tr>
<td>Primary refractory, %</td>
<td>6%</td>
<td>2%</td>
<td>0.0395</td>
<td>5%</td>
<td>11%</td>
</tr>
<tr>
<td>Refractory to last salvage, %</td>
<td>21%</td>
<td>52%</td>
<td>&lt;0.0001</td>
<td>27%</td>
<td>25%</td>
</tr>
</tbody>
</table>
Particular Efforts to Minimize Bias

• At data collection stage – requested sites to provide all patients with R/R ALL – rather than having sites apply selection criteria

• Inclusion/exclusion criteria applied centrally across all data sets

• Study sites reflected centers of excellence for treatment of ALL

• Weighting, stratified, and propensity score analyses to make endpoints more comparable

• Variety of sensitivity analyses conducted in order to address assumptions