Handling of Missing Data: Learnings from Saxenda® NDA

Kamilla Begtrup
Todages møde 10-11 Nov 2015
• Background

• Steps taken to address the missing data problem

• Results

• FDA evaluation

• Summary and conclusion
Background

- Saxenda® (Liraglutide 3.0 mg) intended for weight management
- New drug application (NDA) submitted to FDA December 2013
- Advisory committee meeting (AdCoM) held on September 11th 2014
- Approved December 2014

- Phase 3 program
  - 3 placebo controlled trials of 56 week duration (body weight)
    - First trial initiated ~5 years prior to NDA
  - 1 placebo controlled trial of 32 week duration (sleep apnoea)
Weight management guideline:

- Full analysis set (FAS) should be all randomised with a post baseline observation
- Primary analysis should be based on last observation carried forward (LOCF)
- A product can be considered effective if after 1 year of treatment
  
  either
  
  - difference in mean weight loss between active and placebo is at least 5 % and statistically significant

  or

  - proportion of subjects losing at least 5 % weight in active group is at least 35 %, is appr. double the proportion in placebo, and the difference between groups is statistically significant
Trial Design – Largest phase 3 trial

- Parallel group, placebo controlled, double-blind

- Co-primary endpoints:
  - percent change in body weight at week 56
  - Proportion losing at least 5% at w56
  - Proportion losing more than 10% at w56

- Tested hierarchically to preserve the family-wise type I error at 5%
Body weight change from baseline (%) – Trial 1839

Data are observed means with standard error bars for patients completing each visit (numbers [N] completing are shown at the end of the curves). The estimated mean weight loss and treatment difference for patients included in the analysis.
Addressing missing data

- Withdrawn patients invited for nominal week 56 (end-of-trial) visit to assess bodyweight (primary endpoint) and collect AE information

- Several sensitivity analyses using different imputation techniques carried out
  - Most pre-specified at protocol level
  - Some additional for the NDA

- Description of missing data patterns through plots
  - Withdrawal patterns and differences between arms have an impact on how the various missing data imputations work
# Missing data

<table>
<thead>
<tr>
<th>Trial</th>
<th>Withdrawn</th>
<th>Withdrawn due to Adverse events</th>
<th>Withdrawn due to ineffective therapy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Liraglutide 3.0 mg</td>
<td>Placebo</td>
<td>Liraglutide 3.0 mg</td>
</tr>
<tr>
<td>1839</td>
<td>28.1%</td>
<td>35.6%</td>
<td>9.6%</td>
</tr>
<tr>
<td>1923</td>
<td>25.0%</td>
<td>30.5%</td>
<td>8.5%</td>
</tr>
<tr>
<td>1922</td>
<td>23.4%</td>
<td>34.0%</td>
<td>9.2%</td>
</tr>
<tr>
<td>3970</td>
<td>25.6%</td>
<td>20.7%</td>
<td>11.1%</td>
</tr>
</tbody>
</table>
Missing data patterns

Time to discontinuation

![Graph showing the probability of withdrawal over time for Placebo and Lira 3.0 mg treatments. The graph plots time on treatment (weeks) on the x-axis and probability of withdrawal (%) on the y-axis. The Placebo group shows a higher probability of withdrawal compared to the Lira 3.0 mg group.]
Body weight change by last available on-drug measurement

![Graph showing body weight change over time for Liraglutide 3.0 mg and Placebo groups.](image)
Proportion of withdrawn patients returning for week 56 visit – Retrieved drop-outs (RD)

<table>
<thead>
<tr>
<th>Trial</th>
<th>Liraglutide 3.0 mg</th>
<th>Placebo</th>
</tr>
</thead>
<tbody>
<tr>
<td>1839</td>
<td>28.2%</td>
<td>24.5%</td>
</tr>
<tr>
<td>1923</td>
<td>37.5%</td>
<td>40.0%</td>
</tr>
<tr>
<td>1922</td>
<td>38.6%</td>
<td>32.4%</td>
</tr>
<tr>
<td>3970</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>
Imputation considerations

- Reasons for the missing data
  - Missing completely at random (MCAR)
  - Missing at random (MAR)
  - Missing not at random (MNAR)

- Imputations should preferably take the uncertainty of the imputed value into account
  - multiple imputation versus single imputation

- What is the scientific question of interest – the “estimand”. E.g.,
  - ideal treatment effect that could have been reached if all patients had fully adhered (de-jure)
  - treatment effects that occur when full adherence to treatment is lacking (de-facto)
  - The estimand concept was not used in the NDA
Analyses of % change in body weight

1. LOCF (primary)
2. “Retrieve drop-out”: Using follow-up BW if available (30%) - rest imputed by LOCF
3. MMRM (cont. endpoint)
4. MI with sequential CR-type imputation
5. Completers only
6. LOCF plus BOCF for patients without post-baseline assessment
7. BOCF for all withdrawals not increasing weight during trial otherwise the observed weight increase was used as imputed value

1-5: FAS defined as all patients with post baseline measurements,
6-7: All randomised
LOCF: last observation carried forward, BOCF: baseline observation carried forward, BW: Body weight
MMRM: mixed model repeated measurements, MI: Multiple imputation, CR: Copy reference
Repeated measures - MMRM

- Utilises all post baseline observations in the estimation of treatment difference at end-of-trial
- Assumes data are MAR - response trajectories for patients withdrawing from treatment are comparable to those for similar patients that complete treatment
- Estimate what would have been the result at the end of the trial had all patients remained in the trial and on treatment (a de-jure estimand addressing efficacy)
Missing at random (MAR)

- Example: Likely imputations for a patient in the Liraglutide 3.0 mg arm who withdraws after visit 2.

Source: Clinical trials with missing data: a guide for practitioners (2014), M. O’Kelly and B. Ratitch
1. Intermittent missing values imputed using a MCMC method to obtain a monotone missing data pattern. 100 copies of dataset generated.

2. For each copy, ANOVA model with the same factors and covariates as the primary model was fitted to the second post-baseline visit value for the completing placebo patients. Estimated parameters, and their variances, from this model were used to impute missing values at the second post-baseline visit for WD patients in both treatment groups.

3. Missing values at the next planned visit imputed in the same manner, but also included the body weight value from the previous visit as a covariate in the model.
   - Repeated stepwise for all available planned visits.

4. Estimates and SDs for the 100 data sets were pooled to one estimate and associated SD using Rubin’s rule.
Multiple Imputation - sequential CR-type imputation

- Estimate what would have been the result at the end of the trial had all patients remained in trial and on diet an exercise after treatment discontinuation
Copy reference

• ITT-like approach. The treatment effect gradually disappear after it has been discontinued.

• The patient’s missing values when imputed will look like a rather successful patient in the reference group.

Source: Clinical trials with missing data: a guide for practitioners (2014), M. O’Kelly and B. Ratitch
Sensitivity Analyses of Change in Body Weight (%) – Trial 1839

Estimated treatment difference (%)

<table>
<thead>
<tr>
<th>Analysis Type</th>
<th>Change from baseline (%)</th>
<th>Liraglutide 3.0 mg LSMean (N)</th>
<th>Placebo LSMean (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary analysis</td>
<td>-7.99 (2432)</td>
<td>-2.60 (1220)</td>
<td></td>
</tr>
<tr>
<td>All measurements*</td>
<td>-7.80 (2437)</td>
<td>-2.63 (1225)</td>
<td></td>
</tr>
<tr>
<td>Repeated measures analysis</td>
<td>-8.52 (2432)</td>
<td>-2.69 (1220)</td>
<td></td>
</tr>
<tr>
<td>Multiple imputation</td>
<td>-8.26 (2437)</td>
<td>-2.74 (1225)</td>
<td></td>
</tr>
<tr>
<td>Completers</td>
<td>-9.22 (1781)</td>
<td>-3.53 (798)</td>
<td></td>
</tr>
<tr>
<td>LOCF with BOCF patients w/o valid post baseline#</td>
<td>-7.83 (2481)</td>
<td>-2.57 (1239)</td>
<td></td>
</tr>
<tr>
<td>BOCF or weight increase#</td>
<td>-6.67 (2481)</td>
<td>-2.04 (1239)</td>
<td></td>
</tr>
</tbody>
</table>

Favors liraglutide 3.0 mg
Favors placebo

NDA; FAS at end-of-treatment. *All randomized. N, number of patients contributing to analysis. *All measurements included fasting and non-fasting weight measurements, off drug measurements, and follow-up measurements after 56 weeks of randomization for patients who discontinued.
Sensitivity Analysis for Achieving ≥5% Weight Loss – Trial 1839

Estimated odds ratio

- **Primary analysis**: 63.53 (2432) vs. 26.61 (1220)
- **All measurements***: 63.16 (2437) vs. 27.74 (1225)
- **Multiple imputation**: 67.78 (2437) vs. 32.46 (1225)
- **Completers**: 73.33 (1781) vs. 35.68 (798)
- **LOCF with BOCF patients w/o valid post baseline#**: 62.25 (2481) vs. 26.22 (1239)
- **Early withdrawals counted as non-resp. #**: 53.22 (2481) vs. 23.05 (1239)

NDA; FAS at end-of-treatment. #All randomized. N, number of patients contributing to analysis. *All measurements included fasting and non-fasting weight measurements, off drug measurements, and follow-up measurements after 56 weeks of randomization for patients who discontinued.
Feedback - FDA

- LOCF not acceptable
- FDA considered FAS as being *all randomised*
- NN sensitivity analyses inadequate as they generally reflects a *de-jure* analysis rather than a *de-facto* analysis
  - Multiple Imputation method applied to Per-Protocol like
  - Imputed values do not agree with observed values for retrieved drop-outs in Liraglutide 3.0 mg arm
- FDA notes that for retrieved drop-outs, placebo patients on average loose weight from WD visit to FU visit, whereas the opposite is the case for Liraglutide 3.0 mg patients
Comparison of weight change at last visit before withdrawal and at week 56 for RD patients

<table>
<thead>
<tr>
<th>Treatment Group</th>
<th>N</th>
<th>LAO-OT Mean (SE)</th>
<th>Week 56 (Actual) Mean (SE)</th>
<th>Mean Difference; LAO-OT – Week 56 (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Trial 1839</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liraglutide 3.0 mg</td>
<td>171</td>
<td>-4.9% (0.4)</td>
<td>-3.0% (0.6)</td>
<td>-1.8% (-2.7, -1.0)</td>
</tr>
<tr>
<td>Placebo</td>
<td>100</td>
<td>-0.4% (0.4)</td>
<td>-1.3% (0.7)</td>
<td>0.9% (-0.4, 2.1)</td>
</tr>
<tr>
<td><strong>Trial 1922</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liraglutide 3.0 mg</td>
<td>33</td>
<td>-4.4% (0.7)</td>
<td>-2.5% (0.8)</td>
<td>-1.8% (-3.2, -0.5)</td>
</tr>
<tr>
<td>Liraglutide 1.8 mg</td>
<td>8</td>
<td>-4.3% (1.3)</td>
<td>-2.4% (1.8)</td>
<td>-1.9% (-5.1, 1.3)</td>
</tr>
<tr>
<td>Placebo</td>
<td>23</td>
<td>-1.4% (0.4)</td>
<td>-1.7% (0.7)</td>
<td>0.3% (-1.5, 2.0)</td>
</tr>
<tr>
<td><strong>Trial 1923</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liraglutide 3.0 mg</td>
<td>12</td>
<td>-6.4% (1.0)</td>
<td>-1.1% (1.9)</td>
<td>-5.3% (-7.8, -2.8)</td>
</tr>
<tr>
<td>Placebo</td>
<td>18</td>
<td>-0.5% (1.0)</td>
<td>-1.1% (2.0)</td>
<td>0.5% (-2.8, 3.8)</td>
</tr>
</tbody>
</table>

Source: FDA statistical reviewer

LAO-OT: Last observation on treatment
Smoothed histograms of actual and imputed weight changes for RD patients using NN MI method

FDA comments to NN MI:
- Imputation for Liraglutide arm appear over-optimistic
- This will lead to biased results

Source: FDA statistical reviewer
FDA sensitivity analyses

- FDA performed two sensitivity analyses of their own:
  - MI-RD (stated as their preferred approach):
    - Combined follow-up measurements from returning drop-outs with a MI approach:
      - MI based on observed follow-up measurements from returning drop-outs.
      - Grouping patients by treatment and time of last on-treatment measurement
    - Imputes under a MAR assumption
  - RD-Weighted:
    - Continuous endpoint analysed using weighted ANCOVA:
      - Completers assigned a weight of 1
      - Withdrawals not returning assigned a weight of zero
      - Returning drop-outs weighted depending on time of last on-treatment value
    - Imputes under a MCAR assumption
### Most Conservative Sensitivity Analyses - NN versus FDA analyses – Trial 1839

<table>
<thead>
<tr>
<th>Change in body weight (%)</th>
<th>Difference Liraglutide to Placebo [95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>NN: BOCF or weight increase</td>
<td>-4.6 [-5.1; -4.2]</td>
</tr>
<tr>
<td>FDA: MI-RD</td>
<td>-4.8 [-5.3; -4.3]</td>
</tr>
<tr>
<td>FDA: RD-Weighted</td>
<td>-4.6 [-5.4, -3.9]</td>
</tr>
</tbody>
</table>

All randomized
Most Conservative Sensitivity Analyses - NN versus FDA analyses – Trial 1839

<table>
<thead>
<tr>
<th>Achieving ≥5% weight loss</th>
<th>LSMean Liraglutide 3.0 mg</th>
<th>LSMean Placebo</th>
</tr>
</thead>
<tbody>
<tr>
<td>NN: non-responders</td>
<td>53%</td>
<td>23%</td>
</tr>
<tr>
<td>FDA: MI-RD</td>
<td>62%</td>
<td>34%</td>
</tr>
<tr>
<td>FDA: RD-Weighted</td>
<td>62%</td>
<td>31%</td>
</tr>
</tbody>
</table>

All randomized

Difference between treatments were statistical significant for all
Limitations of the FDA approach

- The RD patients may not be representative of the non-RD patients
  - Only ~30% of withdrawn patients were RD
  - RD placebo patients continued to lose weight after withdrawal
Summary and conclusion

• LOCF unacceptable

• No unique imputation method will address the missing data

• It is likely that additional post-hoc sensitivity analyses are needed

• Retrieval of follow-up information for withdrawn patients is considered important
  • FDA did not agree with NN definition of missing data

• The efficacy of Liraglutide 3.0 mg not questioned – only the magnitude of the effect