Quantifying the impact of mortality underreporting on analyses of overall survival

Michael T Bretscher, Thibaut Sanglier
F. Hoffmann-La Roche Ltd, Basel, Switzerland

Background
- Clinical outcomes may occasionally be compared across different datasets, e.g. when a historical control group serves as comparator for a single arm trial.
- The two datasets may differ in completeness and accuracy of outcome reporting (e.g. death).
- Mortality underreporting leads to overestimation of overall survival (OS), and differential underreporting between compared datasets leads to biased estimates of mortality hazard ratios.
- Endpoint validation through collection of additional data, or linking to a “gold standard” source of mortality information is often not practical.
- A theoretical framework could help quantify and better understand the impact of mortality underreporting on median OS and mortality hazard ratios.
- Flatiron Health (FIH) developed an electronic health records (EHR)-derived database\(^1\).

Fig 1: Kaplan-Meier plots of OS, by subgroup and mortality version

Methods
- Mathematical models (described below) predict OS biases.
- Validation of theoretical predictions using real world data:
  - A cohort of incident metastatic breast cancer (mBC) patients diagnosed between January 1\(^{st}\) 2011 and July 31\(^{st}\) 2016 was identified in the FIH EHR-derived database\(^1\). This allowed for both OS1.0 and OS2.0 (from diagnosis) to be available for all patients.
  - Data beyond July 31\(^{st}\) 2016 was censored.
  - Patients were stratified by biomarker status, using well-known prognostic biomarkers.
  - Within the five biomarker-based strata, each patient was duplicated to be present with mortality information versions OS1.0 (arm 1) as well as OS2.0 (arm 2), permitting comparison of mortality endpoints.
  - Within biomarker subgroups, hazard ratios of OS1.0 (arm 1) vs OS2.0 (arm 2) were compared against theoretical predictions (equation 1 below)
  - Similarly, a ratio of Kaplan-Meier median OS1.0 (arm 1) vs median OS2.0 (arm 2) was compared against theoretical predictions (equation 2 below)

Mathematical Models

### Parameters

- \( s_1 \): proportion of reported deaths (“sensitivity”) in arms 1 and 2, respectively
- \( R_{12} \): Observed hazard ratio comparing arms 1 and 2 (biased by mortality underreporting)
- \( \rho_{12} \): True hazard ratio comparing arms 1 and 2 (unbiased, with perfect mortality data)
- \( m_{1R} \): Observed median OS in arms 1 and 2 (biased by mortality underreporting)
- \( \mu_{1R} \): True median OS in arms 1 and 2 (unbiased, with perfect mortality data)

### Impact of mortality underreporting on OS hazard ratios

1. \[
\frac{R_{12}}{\rho_{12}} = \frac{s_1}{s_2}
\]
   - “The observed hazard ratio between arms 1 and 2 compares to the true hazard ratio as the sensitivity in arm 1 compares to the sensitivity in arm 2.”

2. \[
\frac{m_1}{m_2} = \frac{\mu_1}{\mu_2} \frac{s_2}{s_1}
\]
   - “The apparent ratio of medians between arms 1 and 2 equals the true ratio of medians divided by the ratio of corresponding sensitivities”

### Impact of mortality underreporting on median OS

- Mathematical model (eq. 1)
- Division of medians divided by the ratio of corresponding sensitivities

### Results

- N=5,483 mBC patients were identified in the FIH data, with subgroups HER2+/HR+ (N=842), HER2+/HR- (N=305), HER2-/HR+ (N=3608), triple negative (N=562), missing biomarker information (N=166).
- Median follow-up: 22.6 months
- OS differed by biomarker subgroup
- Mortality data version 1.0 (OS1.0) leads to numerically longer survival due to a higher % of missed deaths

### References

1. Flatiron Health database (https://flatiron.com/real-world-evidence/), May 2018, mortality v2.0