Supplementary Materials for “Using Post Outcome Measurement Information in Censoring by Death Problems”

Fan Yang 
University of Chicago, Chicago, USA.

Dylan S. Small
University of Pennsylvania, Philadelphia, USA.

1. Bounds of the SACE

Given the value of $\pi_{1100}$, the linear programming problem (19)-(25) in the paper has a solution if and only if the set $\Phi = \{\max\{q_{110} - \pi_{1000}, \frac{q_{1111} \pi_{1100}}{\pi_{1111} + \pi_{1110} + \pi_{1010}}, \frac{q_{1101} \pi_{1100}}{\pi_{1110} + \pi_{1010}}\}\}$ is not empty, which is essentially $\frac{q_{1101}}{p_{1101}} \geq \frac{q_{1111}}{p_{1111}}$, an inequality that must be satisfied based on Assumptions 5-7. If $\Phi$ is not empty, let $T = \max\{q_{110} - \pi_{1000}, \frac{q_{1111} \pi_{1100}}{\pi_{1111} + \pi_{1110} + \pi_{1010}}\}$, $\overline{T} = \frac{q_{1101} \pi_{1100}}{\pi_{1110} + \pi_{1010}}$, the solution to the linear programming problem is,

$$\max((\pi_{1111} E(Y_i(1) | 1111) + \pi_{1110} E(Y_i(1) | 1110) + \pi_{1100} E(Y_i(1) | 1100)) | \pi_{1100})$$

$$= \frac{q_{1111} \pi_{1111} + \pi_{1110}}{\pi_{1111} + \pi_{1110} + \pi_{1010}} + \overline{T}, \quad (S.1)$$

$$\min((\pi_{1111} E(Y_i(1) | 1111) + \pi_{1110} E(Y_i(1) | 1110) + \pi_{1100} E(Y_i(1) | 1100)) | \pi_{1100})$$

$$= \begin{cases} 
T & \text{if } \frac{q_{1111} \pi_{1100}}{\pi_{1010}} \leq T \\
q_{1111} + (1 - \frac{\pi_{1010}}{\pi_{1100}}) \overline{T} & \text{if } \frac{q_{1111} \pi_{1100}}{\pi_{1010}} \geq \overline{T} \\
q_{1111} + (1 - \frac{\pi_{1010}}{\pi_{1100}}) \overline{T} & \text{if } T < \frac{q_{1111} \pi_{1100}}{\pi_{1010}} < \overline{T}. 
\end{cases} \quad (S.2)$$
where

\[ \dot{T} = \begin{cases} T & \text{if } \pi_{1010} \leq \pi_{1100} \\ \bar{T} & \text{if } \pi_{1010} > \pi_{1100} \end{cases} \]

\[ \ddot{T} = \begin{cases} T & \text{if } \pi_{1010} \leq \pi_{1100} \\ \frac{q_{1110} \pi_{1100}}{\pi_{1010}} & \text{if } \pi_{1010} > \pi_{1100} \end{cases} \]

Thus, given a fixed value of \( \pi_{1100} \), the bounds for the SACE are given by:

\[
\min(SACE | \pi_{1100}) = \min_{\pi_{1100} \in I} \left( \min\left( \pi_{1111} E(Y_i(1) | 1111) + \pi_{1110} E(Y_i(1) | 1110) + \pi_{1100} E(Y_i(1) | 1100) | \pi_{1100} \right) - (q_{1110} + q_{1100}) \right),
\]

\[
\max(SACE | \pi_{1100}) = \max_{\pi_{1100} \in I} \left( \max\left( \pi_{1111} E(Y_i(1) | 1111) + \pi_{1110} E(Y_i(1) | 1110) + \pi_{1100} E(Y_i(1) | 1100) | \pi_{1100} \right) - (q_{1110} + q_{1100}) \right).
\]

From section 3.1, we know that \( \pi_{1100} \) is not point identified, but bounded: \( \pi_{1100} \in I, I = [\max\{0, p_{1100} + p_{1000} - p_{1001}\}, \min\{p_{1000} + p_{1001}\}] \), we have,

\[
\min SACE = \min_{\pi_{1100} \in I} \left( \min\left( \pi_{1111} E(Y_i(1) | 1111) + \pi_{1110} E(Y_i(1) | 1110) + \pi_{1100} E(Y_i(1) | 1100) | \pi_{1100} \right) - (q_{1110} + q_{1100}) \right),
\]

\[
\max SACE = \max_{\pi_{1100} \in I} \left( \max\left( \pi_{1111} E(Y_i(1) | 1111) + \pi_{1110} E(Y_i(1) | 1110) + \pi_{1100} E(Y_i(1) | 1100) | \pi_{1100} \right) - (q_{1110} + q_{1100}) \right).
\]

One can prove that the expression on the right side of equation (S.2) is continuous as a function of \( \pi_{1100} \) and both the functions on the right side of equations (S.1) and (S.2) are non-decreasing as functions of \( \pi_{1100} \). Thus, the \( \max SACE \) could be achieved when \( \pi_{1100} \) is \( \min\{p_{1100} + p_{1000} - p_{1001}\} \) which is the right end point of the range for \( \pi_{1100} \), and the \( \min SACE \) could be achieved when \( \pi_{1100} \) is \( \max\{0, p_{1100} + p_{1000} - p_{1001}\} \) which is the left end point of the range for \( \pi_{1100} \). Based on this observation, we can obtain the formula for the bound of SACE which is given in (26) and (27) in the paper.

2. The ARDSNet data

861 patients were randomized to receive mechanical ventilation with either lower tidal volume or traditional tidal volume. The lower tidal volume group contained
432 patients and the traditional tidal volume group contained 429 patients. We created our variables based on the recorded answers for the study termination form and weaning form.

The first time point (day 28) survival information is obtained through the "ST2DT" variable in the study termination sub-dataset which recorded the date of death. If the date of death for subject i is below day 28, then $S_{1i}$ is 0 and the QOL is not defined; otherwise, $S_{1i}$ is 1.

For the patients who survive to day 28, the QOL that whether patient was able to breathe without assistance by day 28 was well defined. The variable "UNASSIST" in the study termination sub-dataset recorded whether the patient was able to sustain unassisted breathing for $\geq 48$ hours during the first 28 days after initiation of study procedures. However, even if the patient sustained unassisted breathing for at least 48 hours, the patient could return to assisted breathing before day 28. The variable "ASSIST" recorded this information. If the patient returned to assisted breathing from unassisted breathing for at least 48 hours, the "ASSIST" was recorded as "Yes". Thus, for patients whose "UNASSIST" was recorded as "No", we view them as the ones who were not able to breathe without assistance by day 28. For patients whose "UNASSIST" was recorded as "Yes", and "ASSIST" was recorded as "No", we view them as the ones who were able to breathe without assistance by day 28; for patients whose "UNASSIST" was recorded as "Yes" and "ASSIST" was recorded as "Yes", each patient could either (a) have had unassisted breathing at some point and then returned to assisted breathing and still be on assisted breathing at day 28 or (b) have had unassisted breathing before day 28, returned to assisted breathing before day 28 and then returned to unassisted breathing before day 28. For these patients, we further use the weaning sub-dataset which recorded in detail about each patients' breathing status to figure out whether the patient was able to breathe without assistance by day 28.

Our second time point survival indicator is whether the patient was eventually
Yang and Small

discharged home with unassisted breathing. This information was recorded in the variable "STATUS" which described patient status at study termination.

3. Extension to IV settings

The idea of using second time point survival information to sharpen inferences about the SACE in randomized trials with perfect compliance can be naturally extended to randomized trials with noncompliance or observational studies with a valid IV to obtain inference about the complier survivor average causal effect (CSACE). An instrumental variable (IV) is a variable that is (i) associated with the treatment, (ii) has no direct effect on the outcome, and (iii) is independent of unmeasured confounders conditional on measured confounders. In a randomized trial with non-compliance, the assignment of treatment can be used as an IV to assess the effects of receiving the treatment (which is not randomized when there is noncompliance) on the outcome. In observational studies, natural experiments such as a person's draft lottery number, randomly assigned federal judges or quarter of birth have been used as IVs. (Angrist, 1990; Angrist and Krueger, 1991; Kling, 1999). The basic idea of the IV method is to extract variation in the treatment that is free of the unmeasured confounders and use this confounder free variation to estimate the causal effect of the treatment on the outcome. The beauty of the IV method is that although treatment is not randomly received in observational studies, the method still allows consistent estimation of the causal effect of a treatment. In this paper, we focus on cases where the assignment of the IV is ignorable or ignorable conditional on some discrete covariates so that we can apply our method to each subgroup defined by each level of covariates. How to deal with the case of continuous measured confounders requires further research. For more literatures on IV, see Angrist, Imbens, and Rubin (1996), Abadie (2002), Baiocchi, Cheng and Small (2014), Hernan and Robins (2006), Tan (2006), Brookhart and Schneeweiss (2007), Cheng (2009), and Clarke and Windmeijer (2012).
Let $Z_i$ represent the binary IV; 1 encourages the treatment for the $i^{th}$ subject and 0 does not provide encouragement of the treatment. We use $Z$ to denote the vector of IV for all subjects. Let $D_i(z)$ be the potential binary treatment variable that would be observed under IV assignment $z$ for subject $i$; 1 being the treatment and 0 denotes the control. Let $S_{1i}(z)$ be the potential survival indicator of subject $i$ that would be observed at the first time point after which the measurement of non-mortality outcome is taken; with 0 indicating death, 1 if alive. Let $Y_i(z)$ represent the potential non-mortality binary outcome that would be observed under IV assignment $z$. Again, the non-mortality outcome would be measured after the first time point, thus if the subject would die before that time point ($S_{1i}(z) = 0$), $Y_i(z)$ is not defined; otherwise $S_{1i}(z) = 1$ and $Y_i(z) = 1$ or 0, 1 indicating a worse outcome.

We further define $S_{2i}(z)$ to be the potential indicator of survival at the second time point for subject $i$ that would be observed if under IV assignment $z$. As in section 2, if $S_{1i}(z) = 0$, then $S_{2i}(z) = 0$ by definition. We use $Z_i, D_i, S_{1i}, Y_i$ and $S_{2i}$ to denote respectively the observed IV, treatment received, observed survival indicator at the first time point, observed non-mortality outcome and observed survival indicator at the second time point for subject $i$.

We assume the following assumptions hold for the IV setup. These assumptions combine those of Angrist, Imbens and Rubin (1996) for the IV setup and the ranked average score assumptions with two time points survival information in section 2 of the paper.

**Assumption IV-1.** Stable unit treatment value assumption (SUTVA).

- Let $z$ and $z'$ be any two possible IV assignments. If $z_i = z'_i$, then $D_i(z) = D_i(z')$, $S_{1i}(z) = S_{1i}(z')$, $S_{2i}(z) = S_{2i}(z')$, and $Y_i(z) = Y_i(z')$.

SUTVA means that a subject's potential treatments and outcomes are not affected by other individuals' IV status and means that we can write $D_i(z)$ as $D_i(z_i)$, $S_{1i}(z)$ as $S_{1i}(z_i)$, $S_{2i}(z)$ as $S_{2i}(z_i)$ and $Y_i(z)$ as $Y_i(z_i)$.

**Assumption IV-2.** Nonzero average causal effect of $Z$ on $D$. The average causal
effect of $Z$ on $D$, $E[D_i(1) - D_i(0)]$, is not equal to zero.

Assumption IV-3. Independence of the instrument from unmeasured confounders: the random vector $(D(1), D(0), S_1(1), S_1(0), S_2(1), S_2(0), Y(1), Y(0))$ is independent of $Z$.

Based on subjects’ compliance behavior, we can first partition the population into four groups:

$$U_i = \begin{cases} 
00, & \text{if } D_i(1) = D_i(0) = 0 \\
10, & \text{if } D_i(1) = 1, D_i(0) = 0 \\
11, & \text{if } D_i(1) = D_i(0) = 1 \\
01, & \text{if } D_i(1) = 0, D_i(0) = 1 
\end{cases} \quad (S.3)$$

where 00, 10, 11, and 01 represent never taker, complier, always taker and defier, respectively. Because $D_i(1)$ and $D_i(0)$ are never observed jointly, the compliance behavior of a subject is unknown.

Assumption IV-4. Monotonicity of effect of IV on treatment: $D_i(1) \geq D_i(0)$. There is no $U=01$ group.

Assumption IV-5. Monotonicity of effect of IV on survival: $S_{1i}(1) \geq S_{1i}(0), S_{2i}(1) \geq S_{2i}(0)$.

The monotonicity of the effect of the IV on the survival will hold if the treatment never causes death and Assumption IV-4 holds if the IV has a monotone effect on treatment.

Assumption IV-6. Exclusion restrictions among never-takers and always-takers: $S_{1i}(1) = S_{1i}(0), S_{2i}(1) = S_{2i}(0), Y_i(1) = Y_i(0)$, for $U_i = 00$ or 11.

This means that the IV only affects the outcomes through treatment and has no direct effect on outcomes.

Based on the possible joint combinations of $(D_i(1), D_i(0), S_{1i}(1), S_{1i}(0))$ under the above assumptions, we can define principal strata as shown in Table S.1.

Different from the case of randomized experiments with perfect compliance, the
Table S.1. Principal Strata

<table>
<thead>
<tr>
<th>$D_i(1)$</th>
<th>$D_i(0)$</th>
<th>$S_{1i}(1)$</th>
<th>$S_{1i}(0)$</th>
<th>Principal Strata</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>Compliers, always survivors</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>Compliers, protected</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>Compliers, never survivors</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>Never takers, always survivors</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>Never takers, never survivors</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>Always takers, always survivors</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>Always takers, never survivors</td>
</tr>
</tbody>
</table>

Principal strata in the IV setup are defined with respect to IV levels, for example, the “compliers, always survivors” are compliers (comply with their IV encouragement of treatment) who would survive under both IV levels. Among all the principal strata, the “compliers, always survivors” (1011) group is the only group that we can observe the outcome under treatment if IV is 1, as well as the outcome under control if IV is 0, and that would survive under both treatment such that the non-mortality outcome $Y$ is well defined in both cases. Thus, it is the only group for which variation in the IV can identify the causal effect of the treatment on the non-mortality outcome:

$$CSACE = E(Y_i(1) - Y_i(0) | 1011).$$

Similar to the case of randomized experiments with perfect compliance (section 2 in the paper), we can further incorporate the information of a second time survival indicator to create finer strata as shown in Table S.2.

In terms of the fine strata in Table S.2, the CSACE is expressed as:

$$CSACE = E(Y_i(1) - Y_i(0) | 1011)$$

$$= \frac{(\pi_{101111}E(Y_i(1) | 101111) + \pi_{101110}E(Y_i(1) | 101110) + \pi_{101100}E(Y_i(1) | 101100))}{\pi_{101111} + \pi_{101110} + \pi_{101100}}$$

$$- \frac{(\pi_{101111}E(Y_i(0) | 101111) + \pi_{101110}E(Y_i(0) | 101110) + \pi_{101100}E(Y_i(0) | 101100))}{\pi_{101111} + \pi_{101110} + \pi_{101100}}.$$

The same assumptions are made for compliers as we made for subjects under randomized trials with perfect compliance (Assumptions 5-7 in section 2 of the main
Table S.2. Fine Strata

<table>
<thead>
<tr>
<th>Probability</th>
<th>$D_i(1)$</th>
<th>$D_i(0)$</th>
<th>$S_{1i}(1)$</th>
<th>$S_{1i}(0)$</th>
<th>$S_{2i}(1)$</th>
<th>$S_{2i}(0)$</th>
<th>Principal Strata at Time Point 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\pi_{101111}$</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>Compliers, always survivors</td>
</tr>
<tr>
<td>$\pi_{101110}$</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>Compliers, always survivors</td>
</tr>
<tr>
<td>$\pi_{101100}$</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>Compliers, always survivors</td>
</tr>
<tr>
<td>$\pi_{101010}$</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>Compliers, protected</td>
</tr>
<tr>
<td>$\pi_{101000}$</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>Compliers, protected</td>
</tr>
<tr>
<td>$\pi_{100000}$</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>Compliers, never survivors</td>
</tr>
<tr>
<td>$\pi_{111111}$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>Always takers, always survivors</td>
</tr>
<tr>
<td>$\pi_{111100}$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>Always takers, always survivors</td>
</tr>
<tr>
<td>$\pi_{110000}$</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>Always takers, never survivors</td>
</tr>
<tr>
<td>$\pi_{001111}$</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>Never takers, always survivors</td>
</tr>
<tr>
<td>$\pi_{001100}$</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>Never takers, always survivors</td>
</tr>
<tr>
<td>$\pi_{000000}$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>Never takers, never survivors</td>
</tr>
</tbody>
</table>

Assumption IV-7. Among "compliers, always survivors", the probability of worse outcome for group 101111 is the lowest, whereas the probability of worse outcome for group 101100 is the highest under both IV assignments:

\[
P(Y_i(1) = 1 \mid 101111) \leq P(Y_i(1) = 1 \mid 101110) \leq P(Y_i(1) = 1 \mid 101100),
\]

\[
P(Y_i(0) = 1 \mid 101111) \leq P(Y_i(0) = 1 \mid 101110) \leq P(Y_i(0) = 1 \mid 101100).
\]

Assumption IV-8. Among "compliers, protected", under treatment assignment, the probability of worse outcome for group 101010 is not higher than that for group 101000:

\[
P(Y_i(1) = 1 \mid 101010) \leq P(Y_i(1) = 1 \mid 101000).
\]

Assumption IV-9. Among compliers, under treatment assignment, the probability of worse outcome for group 101100 is not lower than that for group 101010, but
not higher than that for group 101000, and the probability of worse outcome for
group 101110 is not higher than that for group 101010:

\[ P(Y_i(1) = 1 \mid 101110) \leq P(Y_i(1) = 1 \mid 101010) \leq P(Y_i(1) = 1 \mid 101100) \leq P(Y_i(1) = 1 \mid 101000). \]

3.1. Derivation of Bounds of CSACE

As for the SACE in randomized experiments setup, the CSACE is not point iden-
tified without further assumptions based on the observable joint distribution of
\((Z_i, D_i, S_{1i}, S_{2i}, Y_i)\), but can be bounded. We will again adopt the two step method
we used in section 3 of the paper to obtain the bound.

3.1.1. Bounds for proportions of each stratum

The observable strata of \((Z_i, D_i, S_{1i}, S_{2i})\) are mixtures of fine strata, if we use \(p_{s_1s_2d|z}\)
to denote \(P(S_{1i} = s_1, S_{2i} = s_2, D_i = d \mid Z_i = z)\), we have the following identities:

\[ p_{111|1} = \pi_{101111} + \pi_{101110} + \pi_{101010} + \pi_{111111}, \quad (S.3) \]
\[ p_{101|1} = \pi_{101100} + \pi_{101000} + \pi_{111100}, \quad (S.4) \]
\[ p_{001|1} = \pi_{100000} + \pi_{110000}, \quad (S.5) \]
\[ p_{110|1} = \pi_{001111}, \quad (S.6) \]
\[ p_{100|1} = \pi_{001100}, \quad (S.7) \]
\[ p_{000|1} = \pi_{000000}, \quad (S.8) \]
\[ p_{110|0} = \pi_{101111} + \pi_{001111}, \quad (S.9) \]
\[ p_{100|0} = \pi_{101100} + \pi_{001100} + \pi_{101110}, \quad (S.10) \]
\[ p_{000|0} = \pi_{100000} + \pi_{000000} + \pi_{101010} + \pi_{101000}, \quad (S.11) \]
\[ p_{111|0} = \pi_{111111}, \quad (S.12) \]
\[ p_{101|0} = \pi_{111100}, \quad (S.13) \]
and the constraint

\[ 0 \leq \pi_{101111}, \pi_{101110}, \pi_{101100}, \pi_{101000}, \pi_{110000}, \pi_{111111}, \pi_{111110}, \pi_{111000}, \pi_{001111}, \pi_{001100}, \pi_{000000} \leq 1. \]  

(S.15)

Given (S.3)-(S.14), we can express each \( \pi \) in terms of \( p_{s_1s_2d|z} \) and \( \pi_{101100} \):

\[
\begin{align*}
\pi_{000000} &= p_{000|1}, \\
\pi_{001111} &= p_{110|1}, \\
\pi_{001100} &= p_{100|1}, \\
\pi_{111111} &= p_{111|0}, \\
\pi_{110000} &= p_{001|0}, \\
\pi_{111100} &= p_{101|0}, \\
\pi_{100000} &= p_{001|1} - p_{001|0}, \\
\pi_{101111} &= p_{110|0} - p_{110|1}, \\
\pi_{101000} &= p_{101|1} - p_{101|0} - \pi_{101100}, \\
\pi_{101110} &= p_{100|0} - p_{100|1} - \pi_{101100}, \\
\pi_{101010} &= p_{111|1} + p_{110|1} + p_{100|1} - p_{110|0} - p_{100|0} - p_{111|0} + \pi_{101100},
\end{align*}
\]

and subject to the constraint of (S.15), we have,

\[
\max\{0, p_{110|0} + p_{100|0} + p_{111|0} - p_{111|1} - p_{110|1} - p_{100|0} - p_{111|0} + \pi_{101100}\} \leq \pi_{101100} \leq \min\{p_{101|1} - p_{101|0}, p_{100|0} - p_{100|1}\}.
\]
3.1.2. Bounds for the CSACE

For fixed $\pi'$s, let $q_{yS_1S_2D|z}$ denote $P(Y_i = y; S_{1i} = s_1, S_{2i} = s_2, D_i = d \mid Z_i = z)$. We have the following identities based upon the observable strata of $(Y_i, S_{1i}, S_{2i}, D_i, Z_i)$:

$$q_{11111} = \pi_{101111}E(Y_i(1) \mid 101111) + \pi_{101110}E(Y_i(1) \mid 101110) + \pi_{101010}E(Y_i(1) \mid 101010) + \pi_{111111}E(Y_i(1) \mid 111111),$$

$$q_{11011} = \pi_{101100}E(Y_i(1) \mid 101100) + \pi_{101000}E(Y_i(1) \mid 101000) + \pi_{111100}E(Y_i(1) \mid 111100),$$

$$q_{11101} = \pi_{100111}E(Y_i(1) \mid 001111),$$

$$q_{11100} = \pi_{001000}E(Y_i(1) \mid 001100),$$

$$q_{11110} = \pi_{111111}E(Y_i(0) \mid 111111),$$

$$q_{11010} = \pi_{111100}E(Y_i(0) \mid 111100),$$

$$q_{11000} = \pi_{101111}E(Y_i(0) \mid 101111) + \pi_{001111}E(Y_i(0) \mid 001111),$$

$$q_{10100} = \pi_{101110}E(Y_i(0) \mid 101110) + \pi_{101100}E(Y_i(0) \mid 101100) + \pi_{001100}E(Y_i(0) \mid 001100).$$

Recall that

$$CSACE = \frac{(\pi_{101111}E(Y_i(1) \mid 101111) + \pi_{101110}E(Y_i(1) \mid 101110) + \pi_{101010}E(Y_i(1) \mid 101010))}{\pi_{101111} + \pi_{101110} + \pi_{101010}} - \frac{(\pi_{101111}E(Y_i(0) \mid 101111) + \pi_{101110}E(Y_i(0) \mid 101110) + \pi_{101010}E(Y_i(0) \mid 101010))}{\pi_{101111} + \pi_{101110} + \pi_{101010}},$$

which is point identified. Thus to bound the CSACE, we only need to bound $\pi_{101111}E(Y_i(1) \mid 101111) + \pi_{101110}E(Y_i(1) \mid 101110) + \pi_{101100}E(Y_i(1) \mid 101100)$, which defines a linear programming problem:

$$\min / \max \quad (\pi_{101111}E(Y_i(1) \mid 101111) + \pi_{101110}E(Y_i(1) \mid 101110) + \pi_{101100}E(Y_i(1) \mid 101100)) \mid \pi_{101100}$$

Subject to:

$$q_{11111} - q_{11110} = \pi_{101111}E(Y_i(1) \mid 101111) + \pi_{101110}E(Y_i(1) \mid 101110) + \pi_{101010}E(Y_i(1) \mid 101010),$$

$$q_{11011} - q_{11010} = \pi_{101100}E(Y_i(1) \mid 101100) + \pi_{101000}E(Y_i(1) \mid 101000),$$

$$E(Y_i(1) \mid 101111) \leq E(Y_i(1) \mid 101110) \leq E(Y_i(1) \mid 101100), \quad \text{(S.16)}$$

$$E(Y_i(1) \mid 101010) \leq E(Y_i(1) \mid 101000), \quad \text{(S.17)}$$

$$E(Y_i(1) \mid 101110) \leq E(Y_i(1) \mid 101010) \leq E(Y_i(1) \mid 101100) \leq E(Y_i(1) \mid 101000), \quad \text{(S.18)}$$

$$0 \leq E(Y_i(1) \mid 101111), E(Y_i(1) \mid 101110), E(Y_i(1) \mid 101010), E(Y_i(1) \mid 101000), E(Y_i(1) \mid 101010), E(Y_i(1) \mid 101000) \leq 1. \quad \text{(1)}$$
where constraints (S.16)-(S.18) are imposed by assumptions (IV-7) - (IV-9).

The above linear programming problem has a solution if and only if \( \frac{q_{11111} - q_{11110}}{p_{1011} - p_{1010}} \geq \frac{q_{11111} - q_{11110}}{p_{1111} - p_{1110}} \).

For each possible value of \( \pi_{101100} \), we can solve the above linear programming problem; then, combining this result with the bound for \( \pi_{101100} \) derived, let \( L = p_{110|0} + p_{100|0} + p_{111|0} - p_{111|1} - p_{110|1} - p_{100|1} \), \( U = \min\{p_{101|1} - p_{101|0}, p_{100|0} - p_{100|1}\} \), then \( \pi_{101100} \in I \), where \( I = \{\max\{0, L\}, U\} \), we obtain,

\[
\min \text{CSACE} = \begin{cases} 
\max\left\{ \frac{q_{11111} - q_{11110}}{p_{1111} - p_{1110}} - \frac{q_{11010} - q_{11010}}{p_{1101} - p_{1100}} + \frac{q_{11100} - q_{11100}}{p_{1110} - p_{1100}} + \frac{q_{11000} - q_{11000}}{p_{1100} - p_{1100}} \right\}, & \text{if } L \geq 0 \\
\max\left\{ -\frac{q_{11111} - q_{11110}}{p_{1111} - p_{1110}} + \frac{q_{11010} - q_{11010}}{p_{1101} - p_{1100}} + \frac{q_{11100} - q_{11100}}{p_{1110} - p_{1100}} + \frac{q_{11000} - q_{11000}}{p_{1100} - p_{1100}} \right\}, & \text{if } L < 0 
\end{cases}
\]

\[
\max \text{CSACE} = \begin{cases} 
\max\left\{ -\frac{q_{11111} - q_{11110}}{p_{1111} - p_{1110}} - \frac{q_{11010} - q_{11010}}{p_{1101} - p_{1100}} + \frac{q_{11100} - q_{11100}}{p_{1110} - p_{1100}} + \frac{q_{11000} - q_{11000}}{p_{1100} - p_{1100}} \right\}, & \text{if } L \geq 0 \\
\max\left\{ -\frac{q_{11111} - q_{11110}}{p_{1111} - p_{1110}} + \frac{q_{11010} - q_{11010}}{p_{1101} - p_{1100}} + \frac{q_{11100} - q_{11100}}{p_{1110} - p_{1100}} + \frac{q_{11000} - q_{11000}}{p_{1100} - p_{1100}} \right\}, & \text{if } L < 0 
\end{cases}
\]

3.2. Comparisons with Bounds Under Different Sets of Assumptions

In this section, we provide the bounds on the CSACE under (i) Assumptions IV-1 to IV-6 and the ranked average score assumption with one time point survival information: when assigned to treatment, on average, the outcome for "compliers, always survivors" is not worse than "compliers, protected"; (ii) only Assumptions IV-1 to IV-6. For (i), the bounds are:

\[
\min \text{CSACE} = \max\left\{0, \frac{q_{11111} - q_{11110}}{p_{1110} - p_{1101}} + \frac{q_{11011} - q_{11010}}{p_{1110} - p_{1101}} + \frac{q_{11100} - q_{11100}}{p_{1110} - p_{1101}} + \frac{q_{11000} - q_{11000}}{p_{1110} - p_{1101}} \right\}
\]

\[
\max \text{CSACE} = \frac{q_{11111} - q_{11110} + q_{11011} - q_{11010} - (p_{1111} - p_{1110}) + (p_{1101} - p_{1100}) + p_{1110} - p_{1000}}{p_{1110} - p_{1101} + p_{1100} - p_{1100}}
\]

\[
\min \text{CSACE} = \frac{q_{11111} - q_{11110} + q_{11011} - q_{11010} - (p_{1111} - p_{1110}) + (p_{1101} - p_{1100}) + p_{1110} - p_{1000}}{p_{1110} - p_{1101} + p_{1100} - p_{1100}}
\]

\[
\max \text{CSACE} = \frac{q_{11111} - q_{11110} + q_{11011} - q_{11010} - (p_{1111} - p_{1110}) + (p_{1101} - p_{1100}) + p_{1110} - p_{1000}}{p_{1110} - p_{1101} + p_{1100} - p_{1100}}
\]
Table S.3. Setup

<table>
<thead>
<tr>
<th>% of compliers</th>
<th>Fine Strata</th>
<th>% of $Y_i(1) = 1$</th>
<th>% of $Y_i(0) = 1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>35</td>
<td>1111</td>
<td>20</td>
<td>5</td>
</tr>
<tr>
<td>10</td>
<td>1110</td>
<td>25</td>
<td>10</td>
</tr>
<tr>
<td>20</td>
<td>1100</td>
<td>40</td>
<td>25</td>
</tr>
<tr>
<td>5</td>
<td>1010</td>
<td>35</td>
<td>-</td>
</tr>
<tr>
<td>25</td>
<td>1000</td>
<td>60</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>0000</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

For (ii) (only Assumptions IV-1 to IV-6), the lower bound on SACE is still (S.19), and the upper bound on SACE is given by

$$\max SACE = \min\left\{ \frac{q_{1111} - q_{1111} + q_{1101} - q_{1101}}{p_{1100} - p_{1101} + p_{1000} - p_{1001}}, 1 \right\} - \frac{q_{1110} - q_{1110} + q_{1100} - q_{1100}}{p_{1100} - p_{1101} + p_{1000} - p_{1001}}. $$

3.3. A numerical example

Let’s modify the setup described in section 3.4.1 (Example 1) in the paper so that the link between SACE and CSACE can be clearly seen. Instead of describing the underlying truth about the whole population, Table S.3 describes the underlying truth about the compliers only. The CSACE equals 0.15.

Given each set of values of the proportions of compliers, always takers and never takers in the whole population, the proportions of fine strata not described in Table S.3 (see Table 2 in the paper), and the chances of bad quality of life under both treatment and control assignment of fine strata not described in Table S.3 (see Table 2 in the paper) when well-defined, one can obtain $q'$s and $p'$s when we have an infinite sample. One can easily show that no matter what those values are, as long as the underlying truth about the compliers is given by Table S.3, under Assumptions IV-1 to IV-6 as well as the ranked average score assumptions with two time points survival information (Assumptions IV-7 to IV-9), the bounds for the CSACE are [0.106, 0.238];
under Assumptions IV-1 to IV-6 and the ranked average score assumption with one
time point survival information, the bounds on the CSACE are \([-0.054, 0.241]\); under
Assumptions IV-1 to IV-6 only, the bounds on the CSACE are \([-0.054, 0.408]\).

3.4. Checking the plausibility of ranked average score assumptions with two
time points survival information and exclusion restriction assumptions
Based on the same argument in section 4 of the paper, the corresponding necessary
conditions that the probability distribution of the data must satisfy under Assump-
tions (IV-1)-(IV-9) are as follows:,

\[ q_{11011} - q_{11010} \geq 0, \quad p_{10111} - p_{10101} \geq 0, \quad q_{11111} - q_{11110} \geq 0, \quad p_{11111} - p_{11110} \geq 0, \]

\[ \frac{q_{11011} - q_{11010}}{p_{10111} - p_{10101}} \geq \frac{q_{11111} - q_{11110}}{p_{11111} - p_{11110}}. \]

Pearl (1995) provides a necessary condition on the joint probability distribution
of the outcome, treatment and IV when the exclusion restriction holds. Extending
Pearl’s result to our case where exclusion restrictions are assumed on both survival
at the first time point and the second time point as well as a non-mortality out-
come which may be censored, a necessary condition is that the following inequali-
ties hold:

\[ p_{00d|z_1} + q_{010d|z_2} + q_{110d|z_3} + q_{011d|z_4} + q_{111d|z_5} \leq 1, \]

where \(d \in \{0, 1\}, z_i \in \{0, 1\} \) for \(i = 1, 2, 3, 4, 5\).

4. Additional simulation study
This section describes the results on a simulation study that are not presented in
section 6 of the paper.
We conducted a simplified simulation study to examine the performance of the CLR half-median unbiased estimators. We simulated 2000 data sets of size 2000 based on the following mechanism: each subject $i$ has two binary variables, namely, $X_i$ and $Y_i$ following $P(X_i = 1) = P(X_i = 0) = 0.5$, $P(Y_i = 1 \mid X_i = 1) = 1 - P(Y_i = 0 \mid X_i = 1) = 0.58$, $P(Y_i = 1 \mid X_i = 0) = 1 - P(Y_i = 0 \mid X_i = 0) = 0.54$. Suppose a quantity of interest could be partially identified with lower bound being $\max\{P(Y_i = 1 \mid X_i = 1), P(Y_i = 1 \mid X_i = 0)\}$, and with upper bound being $\min\{P(Y_i = 1 \mid X_i = 1) + 1, P(Y_i = 1 \mid X_i = 0) + 1\}$, which corresponds to $[0.58, 1.54]$ for the above setup. In this very simple setup, the probability that the CLR half-median unbiased estimates of the upper bound exceeds the true upper bound is estimated to be 72.40%, and the probability that the CLR half-median unbiased estimates of the lower bound falls below the true lower bound is estimated to be 70.35%. Based on these results, we see that the half-median unbiased estimators can be pretty conservative even in a very simple setup.

5. Extension to ordinal QOL

This section describes our proposed extension to ordinal QOL using post outcome measurement survival information. Suppose that the QOL can take $k$ values, namely, $a_1, a_2, \ldots, a_k$ with larger values representing worse outcome. With some modifications, we can still use the two steps procedures to obtain the bounds on the SACE under Assumptions 1-3 and our ranked average score assumptions with two time points survival information.

When the QOL is ordinal with more than two possible values, we extend our ranked average score assumptions (Assumptions 5-7) to the following assumptions which we denote as Assumptions 5' to 7':

\textit{Assumption 5'}. Among always survivors at time point 1, the average outcome for group 1111 is the best, whereas the average outcome for group 1100 is the worst
under both treatment arms:

\[ E(Y_i(1) | 1111) \leq E(Y_i(1) | 1110) \leq E(Y_i(1) | 1100), \]

\[ E(Y_i(0) | 1111) \leq E(Y_i(0) | 1110) \leq E(Y_i(0) | 1100). \]

Assumption 6’. Among protected at time point 1, the average outcome for group 1010 is not worse than that for group 1000 under treatment:

\[ E(Y_i(1) | 1010) \leq E(Y_i(1) | 1000). \]

Assumption 7’. Under treatment, the average outcome for group 1100 is not better than that for group 1010, but not worse than that for group 1000, and the average outcome for group 1110 is not worse than that for group 1010:

\[ E(Y_i(1) | 1110) \leq E(Y_i(1) | 1010) \leq E(Y_i(1) | 1100) \leq E(Y_i(1) | 1000). \]

To obtain the bounds on SACE, the first step given in the Section 3.1 in the paper to derive bounds for proportions of each stratum remains the same, let’s now describe the method to derive the bounds for the SACE with known proportions of each fine stratum.

As the binary case, the observable strata of \((Y_i, S_{1i}, S_{2i} | D_i)\) are mixtures of potential outcomes from the fine strata. Letting \(q_{Y|S_1S_2|D} \) denote \( P(Y_i = y, S_{1i} = s_1, S_{2i} = s_2 | D_i = d) \), we have the following identities: \( \forall j = 1, 2, ..., k \)

\[ q_{a_j11|1} = \pi_{1111}P(Y_i(1) = a_j | 1111) + \pi_{1110}P(Y_i(1) = a_j | 1110) + \pi_{1010}P(Y_i(1) = a_j | 1010), \]

\[ q_{a_j10|1} = \pi_{1100}P(Y_i(1) = a_j | 1100) + \pi_{1000}P(Y_i(1) = a_j | 1000), \]

\[ q_{a_j11|0} = \pi_{1111}P(Y_i(0) = a_j | 1111), \]

\[ q_{a_j10|0} = \pi_{1110}P(Y_i(0) = a_j | 1110) + \pi_{1100}E(Y_i(0) = a_j | 1100). \]

Recall that
Given \( \pi' \)'s,

\[
(\pi_{1111} \sum_{j=1}^{j=k} a_j P(Y_j(1) = a_j \mid 1111) + \pi_{1110} \sum_{j=1}^{j=k} a_j P(Y_j(1) = a_j \mid 1110) + \pi_{1100} \sum_{j=1}^{j=k} a_j P(Y_j(0) = a_j \mid 1100))
\]

which is point identified. Thus to bound the SACE, we only need to bound the following expression:

\[
(\pi_{1111} \sum_{j=1}^{j=k} a_j P(Y_j(1) = a_j \mid 1111) + \pi_{1110} \sum_{j=1}^{j=k} a_j P(Y_j(1) = a_j \mid 1110) + \pi_{1100} \sum_{j=1}^{j=k} a_j P(Y_j(1) = a_j \mid 1100)),
\]

which defines a linear programming problem:

\[
\begin{align*}
\min & / \max (\pi_{1111} \sum_{j=1}^{j=k} a_j P(Y_j(1) = a_j \mid 1111) + \pi_{1110} \sum_{j=1}^{j=k} a_j P(Y_j(1) = a_j \mid 1110) + \pi_{1100} \sum_{j=1}^{j=k} a_j P(Y_j(1) = a_j \mid 1100)) \\
\text{Subject to:} \\
q_{11111} & = \pi_{1111} P(Y_j(1) = a_j \mid 1111) + \pi_{1110} P(Y_j(1) = a_j \mid 1110) + \pi_{1100} P(Y_j(1) = a_j \mid 1010), \forall j = 1, 2, \ldots, k - 1, \\
q_{11001} & = \pi_{1100} P(Y_j(1) = a_j \mid 1100) + \pi_{1000} P(Y_j(1) = a_j \mid 1000), \forall j = 1, 2, \ldots, k - 1,
\end{align*}
\]
\[
\sum_{j=1}^{k} a_j P(Y_i(1) = a_j \mid 1010) \leq \sum_{j=1}^{k} a_j P(Y_i(1) = a_j \mid 1000),
\]

(S.31)

\[
\sum_{j=1}^{k} a_j P(Y_i(1) = a_j \mid 1110) \leq \sum_{j=1}^{k} a_j P(Y_i(1) = a_j \mid 1100) \leq \sum_{j=1}^{k} a_j P(Y_i(1) = a_j \mid 1000) \leq \sum_{j=1}^{k} a_j P(Y_i(1) = a_j \mid 1010).
\]

(S.32)

Where constraints (S.30)-(S.32) are imposed by Assumptions 5’-7’. There are essentially \(3k - 3\) free parameters, with fixed value of \(\pi_{1100}\), one can obtain the minimum and maximum values for the SACE through solving the above linear programming problem, and the final bounds can be obtained through varying the value of \(\pi_{1100}\) in its bounded interval.

References


