eAppendix for:
“Survival analysis with multiple causes of death: Extending the competing risks model”
by M. Moreno-Betancur, H. Sadaoui, C. Piffaretti and G. Rey

Contents

The contents of this document are as follows:

1. Discussion on the state-specific cumulative incidence functions.
2. Details on the estimating equation methodology used to perform Cox regression on the pure hazards.
3. Details on the disease-attributed cumulative baseline hazard.
4. Details of the weight-attribution strategy assumed in the simulation study and the resulting multi-state model.
5. Additional results for the simulation study.
6. Additional results for the illustrative example.
7. R code used for simulating data.

The notation is as in the main text.

1 State-specific cumulative incidence functions

In the single-cause model (Figure 1 of main text), the cause-specific cumulative incidence function, denoted by \( \tilde{F}_k(t) \), for \( t > 0 \) and \( k \in \{0, 1\} \), represents the actual risk (probability)
of transition into the corresponding state over a period of time and is usually used when the
focus is on prognosis (Fine and Gray, 1999; Koller et al., 2012). The analogous quantities
in the the multiple-cause model (Figure 2 of main text) are the state-specific cumulative
incidence functions, denoted by $F_\pi(t)$ for $t > 0$ and $0 \leq \pi \leq 1$. The function $F_\pi(t)$ is the
probability of transition into state $\pi$ by time $t$.

Under assumption (A0) from the main text, $\lambda_1(t|X)$ and $\lambda_0(t|Z)$ completely determine the
state-specific cumulative incidence functions, which are then given by:

$$F_\pi(t|X, Z) = \int_0^t S(u|X, Z)\{\pi \lambda_1(u|X) + (1 - \pi) \lambda_0(u|Z)\}du$$

$$= \pi F_1(t|X, Z) + (1 - \pi) F_0(t|X, Z),$$

(1)

for $0 \leq \pi \leq 1$, where $S(t) = P(T \geq t)$ is the survival function and $F_1$ and $F_0$ are the pure
cumulative incidence functions.

A point of caution for future research is the interpretability and relevance of different func-
tionals. For example, the expression above would facilitate regression methods for $F_1$ and
$F_0$. However, these functionals are generally useful for prediction, and the relevant event to
be predicted, and thus the approach to be used, will depend on the context: the event could
be death strictly caused by the disease, death with any mention of the disease or death with
the disease selected as underlying cause.

## 2 Estimating equation methodology

Here we describe the methodology used to estimate the parameters of the proposed Cox
model for the pure hazards, grounded on assumptions (A0)-(A3) of the main text. In
addition to these assumptions, we require independent and identically distributed data and
independent censoring given covariates. For ease of exposition here, we assume that the log-
ratio of the baseline pure hazards is constant, i.e. $\xi(t) = \xi$, but any fully-parametric model
is possible. The estimation methodology builds on the estimating equation theory developed
by Goetghebeur and Ryan (1995) and van Rompaye et al. (2012).

If $\xi$ were known, we could estimate $\rho$ and $\phi$ based on a partial likelihood constructed following the
same arguments as Cox (1972). The contribution to this likelihood of a death with $\Pi = \pi$
is based on the probability that this event occurred given that one such event occurred in
the risk set at that time:

$$L(\rho, \phi, \xi) = \prod_{\pi \in \{\pi^{(1)}, \ldots, \pi^{(J)}\}} \prod_{i \in D_\pi} \frac{\pi \exp(\rho'X_i) + (1 - \pi) \exp(-\xi + \phi'Z_i)}{\sum_{j \in R(\tilde{t}_i)} \pi \exp(\rho'X_j) + (1 - \pi) \exp(-\xi + \phi'Z_j)}$$

where $\pi^{(1)} := 1, \pi^{(2)}, \ldots, \pi^{(J-1)}, \pi^{(J)} := 0$ denote the distinct values of $\Pi$ observed among
the finite sample of individuals under study, $D_\pi$ is the set of deaths with $\Pi = \pi$, $R(t)$
denotes the set of individuals alive just before \( t \) and the apostrophe denotes transposition. For identifiability it is necessary to observe individuals in the pure states, which is why 0 and 1 are included among the \( \pi^{(j)} \)'s.

When \( \xi \) is unknown, as is generally the case, \( \mathcal{L} \) cannot be used to estimate it because it loses most information about this parameter. Thus, following Goetghebeur and Ryan (1995), this parameter is estimated using another likelihood originally studied by Dewanji (1992), that conditions on any event-type having occurred:

\[
\mathcal{L}^* (\rho, \phi, \xi) = \prod_{\pi \in \{\pi^{(1)}, \ldots, \pi^{(J)}\}} \prod_{i \in \mathcal{D}_{\pi}} \frac{\pi \exp(\rho'X_i) + (1 - \pi) \exp(-\xi + \phi'Z_i)}{\exp(\rho'X_j) + (J - \pi) \exp(-\xi + \phi'Z_j)}
\]

where the notation is as above, \( \tilde{\pi} := \sum_{j=1}^{J} \pi^{(j)} \) and the denominator follows from \((A0)-(A3)\) and the fact that the total hazard rate is given by

\[
\lambda(u|X_i, Z_i) = \sum_{j=1}^{J} \pi^{(j)} \lambda_1(u|X_i) + (1 - \pi^{(j)}) \lambda_0(u|Z_i) = \tilde{\pi} \lambda_1(u|X_i) + (J - \tilde{\pi}) \lambda_0(u|Z_i).
\]

Estimates of \( \theta = (\xi, \rho, \phi) \) are obtained by iterating between maximizing \( \mathcal{L} \), regarding \( \xi \) as known, and \( \mathcal{L}^* \), regarding \( \rho \) and \( \phi \) as known, until convergence (usually achieved in two or three iterations). Initial values to start the iterations are obtained from a first maximization of \( \mathcal{L} \) with respect to all parameters, as in van Rompaye et al. (2012). As described by these authors in their eAppendix, this iterative procedure is an empirical approach to solving a set of estimating equations \( T(\theta) = 0 \) where \( T(\theta) = (\frac{\partial T}{\partial \xi}, \frac{\partial T}{\partial \rho}, \frac{\partial T}{\partial \phi})' \), \( l^* = \log(\mathcal{L}^*) \) and \( l = \log(\mathcal{L}) \).

The solution, \( \hat{\theta} = (\hat{\xi}, \hat{\rho}, \hat{\phi})' \), has desirable asymptotic properties (consistency and normality).

The asymptotic results also lead to standard error estimators as detailed by van Rompaye et al. (2012). These are obtained using the fact that the asymptotic variance of \( n^{1/2} (\hat{\theta} - \theta) \) is given by \( \mathbf{V} = (\Gamma^{-1}) \Delta (\Gamma^{-1})' \) where \( \Gamma \) is the matrix of first derivatives of \( T(\theta) \), and \( \Delta \) is the matrix of covariation processes of the martingale processes contained in \( T(\theta) \). Fortunately, the following correspondence holds:

\[
\Delta = \begin{pmatrix}
-\Gamma_{\xi \xi} & -\Gamma_{\rho \xi} & -\Gamma_{\phi \xi} \\
-\Gamma_{\rho \xi} & -\Gamma_{\rho \rho} & -\Gamma_{\rho \phi} \\
-\Gamma_{\phi \xi} & -\Gamma_{\phi \rho} & -\Gamma_{\phi \phi}
\end{pmatrix}
\]

where \( \Gamma_{ab} \) denotes the derivative of the a-component of \( T(\theta) \) with respect to \( b \). For instance, \( \Gamma_{\phi \xi} = \frac{\partial T}{\partial \phi \xi} \) and \( \Gamma_{\xi \xi} = \frac{\partial^2 T}{\partial \xi^2} \). Note that \( \Gamma \) is not symmetric because \( \Gamma_{\rho \rho} \neq \Gamma_{\xi \rho} \) and \( \Gamma_{\phi \xi} \neq \Gamma_{\rho \phi} \). Hence, \( \Delta \neq -\Gamma \) due to the two upper left terms in (2), and \( \Delta \) is symmetric. The consequence is that the ‘sandwich’ expression of \( \mathbf{V} \) above does not simplify as in usual likelihood theory. However, expression (2) implies that to estimate \( \mathbf{V} \) it suffices to obtain an
estimate of $\Gamma$, which is easily achieved by plugging in the parameter estimates i.e. $\hat{\Gamma} = \Theta(\hat{\theta})$. This estimate and (2) can then be used to obtain an estimate of $\Delta$, $\Delta$. Finally, standard errors for the parameter estimates are obtained by taking the square-root of the diagonal elements of $V = (\hat{\Gamma}^{-1})\Delta(\hat{\Gamma}^{-1})'$.

The consistency of parameter and standard error estimates obtained using the approach described were verified in additional simulations following the same set-up as that described in the main text with varying $n$ and parameter values (results omitted). An important finding was that asymptotic conditions for this approach are attained at a slower pace than for the single-cause and any-mention approaches (i.e. needs a larger sample size). The reason is that asymptotics for this approach depend on the number of pure events (i.e. those with $\Pi = 1$), while for the single-cause and any-mention approaches these depend on the number of events with, respectively, $\Pi = 1$ and $\Pi = 0$, and $\Pi = 0$. This explains why, in the simulation study results provided in the main text, the distribution of the estimator is not as symmetric in scenarios where the number of pure events in either category of $X$ is expected to be lower (e.g. when $\rho = -2$). Thus augmenting $n$ results in a more symmetric distribution. Still, even at $n = 1000$, the coverage probability of confidence intervals assuming normality was satisfactory.

The estimation procedure described yields standard solutions for the single-cause model if we ignore multiple causes of death (i.e. attribute each death 100% to one cause). An alternative estimation approach would be to use only $L^*$ to estimate all parameters (Lu and Tsiatis, 2005; Nicolaie et al., 2011). This approach may be more efficient and easier to implement, but does not yield standard Cox estimators when there are no multiple causes of death (i.e. when for all $i$, $\Pi_i = 0$ or $\Pi_i = 1$) and it tends to be more sensitive to misspecification of the parametric model for $\xi(t)$ (Goetghebeur and Ryan, 1995).

3 Disease-attributed cumulative baseline hazard

A quantity reflecting the total absolute burden of the disease of interest, that can be seen as a cumulative/time-dependent version of (1) in the main text for the baseline population, is given by:

$$\tilde{\Lambda}_{10}(t) = \sum_{j=1}^{J} \pi^{(j)} \hat{\Lambda}_{10}(t) = \tilde{\pi} \hat{\Lambda}_{10}(t),$$

(3)

where $\hat{\Lambda}_{10}(t)$ is a Breslow-type estimator of the cumulative baseline pure hazard of the cause of interest, $\Lambda_{10}(t) = \int_{0}^{t} \lambda_{10}(u) du$, derived from the estimated Cox model as described below. We refer to (3) as the disease-attributed cumulative baseline hazard. For the burden of other diseases, the expression is $\Lambda_{00}(t) = \sum_{j=1}^{J} \{1 - \pi^{(j)}\} \hat{\Lambda}_{00}(t) = \{J - \tilde{\pi}\} \hat{\Lambda}_{00}(t)$. As shown below, the sum $\hat{\Lambda}_{10}(t) + \hat{\Lambda}_{00}(t)$ yields a consistent estimator for the total cumulative baseline hazard.

A Breslow-type estimator of the cumulative baseline pure hazard of the disease of interest,
\( \Lambda_{10}(t) = \int_{0}^{t} \lambda_{10}(u)du \), can be derived from the estimated Cox model following arguments similar to those of van Rompaye et al. (2012). This estimator is given by:

\[
\hat{\Lambda}_{10}(t) = \sum_{i=1}^{n} \frac{1(U_i = 0 \land \tilde{T}_i \leq t)}{\sum_{j \in R(\tilde{T}_i)} \tilde{\pi} \exp(\hat{\rho}'X_j) + (J - \tilde{\pi}) \exp(-\hat{\xi} + \hat{\phi}'Z_j)}.
\] (4)

The cumulative baseline pure hazard of other diseases can be estimated by:

\[
\hat{\Lambda}_{00}(t) = \sum_{i=1}^{n} \frac{1(U_i = 0 \land \tilde{T}_i \leq t)}{\sum_{j \in R(\tilde{T}_i)} \tilde{\pi} \exp(\hat{\xi} + \hat{\rho}'X_j) + (J - \tilde{\pi}) \exp(\hat{\phi}'Z_j)}.
\] (5)

In the simulation set-up described in the main text, we compared the estimator \( \hat{\Lambda}_{10}(t) + \hat{\Lambda}_{00}(t) \) to the theoretical total cumulative baseline hazard. Table 1 shows the percentage bias, confirming that this is a consistent estimator for the total cumulative baseline hazard.

<table>
<thead>
<tr>
<th>( \rho )</th>
<th>Percentile of event-times</th>
<th>25th</th>
<th>50th</th>
<th>75th</th>
<th>100th</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2</td>
<td></td>
<td>0.3</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>-1</td>
<td></td>
<td>0.2</td>
<td>0.1</td>
<td>0.1</td>
<td>-0.3</td>
</tr>
<tr>
<td>0</td>
<td></td>
<td>0.3</td>
<td>0.2</td>
<td>0.1</td>
<td>-0.1</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>0.1</td>
<td>0.2</td>
<td>0.2</td>
<td>-0.1</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>0.0</td>
<td>-0.1</td>
<td>0.0</td>
<td>-0.4</td>
</tr>
</tbody>
</table>

Table 1: Percentage bias (%) of \( \hat{\Lambda}_{10}(t) + \hat{\Lambda}_{00}(t) \) as an estimator of the total cumulative baseline hazard.

Of note, \( \exp(-\hat{\Lambda}_{10}(t)) \) and \( \exp(-\tilde{\Lambda}_{10}(t)) \) do not have interpretations as probabilities, nor does the equivalent transformation for the estimated cause-specific baseline cumulative hazard in the single-cause model. However, under the unidentifiable assumption of independence of the latent failure times (see Discussion in main text), it would be the quantity \( \exp(-\hat{\Lambda}_{10}(t)) \) that would have a useful interpretation as the counterfactual baseline survival probability. Indeed, in that case the baseline pure hazard would coincide with the baseline marginal hazard i.e. the baseline hazard of death in the world where other causes have been removed.

### 4 Weight-attribution strategy and multi-state model for the simulation study

To generate data in the simulation study, we assumed the following strategy to attribute weights to each of the six possible detailed event-types considered. The resulting multi-state model is shown in Figure 1.
Rule 1 $\Pi = 1$ if the disease is the underlying cause and no other diseases are mentioned in Part II of the certificate.

Rule 2 $\Pi = 0.75$ when the disease is the underlying cause and there are other diseases appearing in Part II of the certificate.

Rule 3 $\Pi = 0.25$ if the disease is not the underlying cause but is mentioned in Part II of the certificate by itself.

Rule 4 $\Pi = 0.25/2 = 0.125$ if the disease is not the underlying cause but is mentioned in Part II of the certificate along with one other disease (different from the underlying cause).

Rule 5 $\Pi = 0.25/3 \approx 0.083$ if the disease is not the underlying cause but is mentioned in Part II of the certificate along with two other diseases (different from the underlying cause).

Rule 6 $\Pi = 0$ if the disease is not mentioned anywhere in the certificate or it is mentioned in Part I but it is not the underlying cause.

![Multi-state model assumed for the simulation study.](image)

**Figure 1: Multi-state model assumed for the simulation study.**

### 5 Additional results for the simulation study

Figure 2 shows further results for the simulation study described in the main text, corresponding to other parameter values, with $\xi \in \{-1, 0, 1\}$ and $\rho, \phi \in \{-1, 0, 1\}$. Each plot
shows, for different values of $\rho$, the distribution of the bias $\hat{\rho} - \rho$ across the 1,000 datasets generated, for each approach and given values of $\phi$ and $\xi$.

Figure 3 shows the mean squared error results for the simulation study described in the main text for all the scenarios and approaches considered, and the additional ‘Only pure’ approach, considering only deaths with $\Pi_i = 1$ as events in standard Cox regression. This last approach is seen to perform reasonably well for Goal 2 in our simplified simulation setting, with only one covariate and the model being correctly specified. However, in reality the true effects of covariates may be different across pure and mixed events, so that the average effect across these strata, which is the actual target for Goal 2 and what assumption (A0) and the fitting procedure aim at recovering, may be quite different from what is estimated from the pure events alone. The incorporation of deaths partially attributed to the disease, i.e. requirement (i), is therefore essential for this Goal.

Overall, these additional results confirm the main conclusions of the simulation study provided in the main text.

6 Additional results for the illustrative example

Table 2 shows the estimates of the inequality index associated with the pure hazard of other diseases ($\text{RII}_0$). It is seen that results are fairly similar across all approaches, with $\text{RII}_0 \approx 2$, which as expected approximately equal to the all-cause relative index of inequality (Moreno-Betancur et al., 2015). For all weight-attribution strategies, the log-ratio of the baseline pure hazards $\xi(t)$ was negative for all diseases across the age-scale, except for the dominant groups at some ages: Neoplasms in the range 40-89 years, Cardiovascular diseases at ages 70 years or older and Other diseases in the age-group 30-49 years (results not shown).

Figure 4 shows the estimated cumulative baseline rate of mortality for each disease group according to strategy (b) ($\omega = 0.75$) and the single-cause model, where the baseline population consists of those with higher socioeconomic status.

Table 3 shows the ranking of disease groups by their burden on mortality for ages 50-89 based on estimates of the cumulative baseline rate of mortality across all approaches. All multiple-cause approaches portray a similar picture of the burden of each disease group relative to others. The single-cause approach ($\omega = 1$) yielded slightly different results across these ages, ranking diseases of the Nervous system and the sense organs and Infectious diseases lower, and Mental diseases higher than the multiple-cause approaches. The ranks of Respiratory and Endocrine diseases shift slightly as well, and Neoplasms are ranked higher by the multiple-cause approaches than the single-cause model for the age-group 50-59 years.

Table 4 shows the resulting ranking of diseases by their burden on mortality according to several approaches for the youngest and eldest age-groups. Results are relatively stable across
approaches for these groups. Some shifts are seen in the 40-49 age-group, but this should be interpreted with care as it is based on small numbers of events.

7 R code for simulating data

In the following we provide the code in R (R Foundation for Statistical Computing, Vienna, Austria) used for generating data in the simulation study. R code for analyzing the data using all the compared approaches can be obtained from the corresponding author.

```R
# Number of individuals:
n<-4000

# Weights corresponding to each of the 6 states in the multi-state model:
p_gen<-c(1,0,0.75,0.25,0.125,0.083)

# Pure baseline hazard of disease of interest assumed to follow a Weibull model with parameters:
lambda<-0.001
v<-2

# Log-ratio of the pure baseline hazards assumed constant:
xi<-(-1)  #=1,0,1

# Effects of binary exposure X1(=Z1) on each pure hazard:
rho<-(-2)  #=2,-1,0,1,2
phi<0  #=-1,0,1

# Data generation
datagen<-function(n,xi,rho,phi,p_gen,lambda,v){
  # Generate balanced binary exposure
  X1<-rbinom(n,1,0.5)
  Z1<-X1

  # Generate failure times
  u<-runif(n,0,1)
  times<-cumhazinv(u,X1,Z1,xi,rho,phi,p_gen,lambda,v)

  # Generate censoring times (results in around 30% censoring)
cens<-runif(n,5,15)
sun=(1*(cens<=times)+0)/n

  # Observed right-censored time-to-event
tt<-pmin(times, cens)
}
```
# Determine event-type
Qgen <- length(pgen)
h1 <- lambda*v*times^v*exp(rho*X1)
h0 <- lambda*v*times^v*exp(-xi+phi*Z1)
allh <- sum(pgen)*h1+(Qgen-sum(pgen))*h0

pr <- NULL
for(ff in 1:Qgen)
    pr <- cbind(pr,(pgen[ff]*h1+(1-pgen[ff])*h0)/allh)

event <- rep(NA,n)
for( m in 1:n) {
    mtn <- rmultinom(1,size=1,prob =pr[m,])
    event[m]<-which(mtn==1)
}
case <- (times<=cens)*event+0
Pi <- rep(NA,n)
Pi[case!=0]<-pgen[case[case!=0]]
CensStatus <- 1*(case==0)+0
#data
dat <- data.frame(Z1,X1,rep(0,n),tt,CensStatus,Pi)
dat <- dat[order(dat[,"tt"]),]
names(dat)[3:4]<-c("TimeEntry","TimeExit")
return(dat)}

# Inverse of the total cumulative hazard
cumhazinv <- function(u,X1,Z1,xi,rho,phi,pgen,lambda,v){
    Qgen <- length(pgen)
    cumhinv <- (-log(u)/(lambda*(sum(pgen)*exp(rho*X1)+
(Qgen-sum(pgen))*exp(-xi+phi*Z1))))^(1/v)
    return(cumhinv)
}

# DATA GENERATION EXAMPLE

dat <- datagen(n,xi,rho,phi,pgen,lambda,v)
head(dat)
table(dat[,"CensStatus")
table(dat[,"Pi"],useNA="always")

References


Figure 2: Additional results of the simulation study: Bias distribution for each approach across 1,000 datasets of \( n = 4,000 \) individuals, with varying values of the model parameters (\( \xi, \phi \) and \( \rho \)).
Figure 3: Mean squared error results for simulation study: Mean squared error for each approach across 1,000 datasets of \( n = 4,000 \) individuals, with varying values of the model parameters (\( \xi, \phi \) and \( \rho \)).
<table>
<thead>
<tr>
<th>Disease group</th>
<th>Any-mention</th>
<th>Single-cause $\omega = 1$</th>
<th>Multiple-cause $\omega = 0.75$</th>
<th>Multiple-cause $\omega = 0.5$</th>
<th>Equal weights</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neoplasms</td>
<td>2.2 (2.0; 2.4)</td>
<td>2.1 (2.0; 2.3)</td>
<td>2.2 (2.0; 2.3)</td>
<td>2.2 (2.0; 2.4)</td>
<td>2.2 (2.0; 2.4)</td>
</tr>
<tr>
<td>Cardiovascular</td>
<td>2.0 (1.8; 2.2)</td>
<td>2.0 (1.8; 2.1)</td>
<td>2.0 (1.9; 2.1)</td>
<td>2.0 (1.9; 2.1)</td>
<td>2.0 (1.9; 2.1)</td>
</tr>
<tr>
<td>Digestive</td>
<td>1.9 (1.8; 2.1)</td>
<td>2.0 (1.8; 2.1)</td>
<td>1.9 (1.8; 2.1)</td>
<td>1.9 (1.8; 2.1)</td>
<td>2.0 (1.8; 2.1)</td>
</tr>
<tr>
<td>Nervous/sense</td>
<td>2.1 (2.0; 2.2)</td>
<td>2.1 (1.9; 2.2)</td>
<td>2.1 (1.9; 2.2)</td>
<td>2.0 (1.9; 2.2)</td>
<td>2.0 (1.9; 2.2)</td>
</tr>
<tr>
<td>Musculoskeletal</td>
<td>2.0 (1.9; 2.1)</td>
<td>2.0 (1.9; 2.1)</td>
<td>2.0 (1.9; 2.1)</td>
<td>2.0 (1.9; 2.1)</td>
<td>2.0 (1.9; 2.1)</td>
</tr>
<tr>
<td>Respiratory</td>
<td>1.9 (1.8; 2.0)</td>
<td>2.0 (1.8; 2.1)</td>
<td>2.0 (1.8; 2.1)</td>
<td>2.0 (1.8; 2.1)</td>
<td>2.0 (1.8; 2.1)</td>
</tr>
<tr>
<td>Endocrine/nutritional</td>
<td>2.0 (1.9; 2.1)</td>
<td>2.0 (1.9; 2.1)</td>
<td>2.0 (1.9; 2.1)</td>
<td>2.0 (1.9; 2.1)</td>
<td>2.0 (1.9; 2.1)</td>
</tr>
<tr>
<td>Mental</td>
<td>1.9 (1.8; 2.0)</td>
<td>2.0 (1.9; 2.1)</td>
<td>2.0 (1.9; 2.1)</td>
<td>2.0 (1.9; 2.1)</td>
<td>2.0 (1.9; 2.1)</td>
</tr>
<tr>
<td>Infectious</td>
<td>2.0 (1.9; 2.1)</td>
<td>2.0 (1.9; 2.1)</td>
<td>2.0 (1.9; 2.1)</td>
<td>2.0 (1.9; 2.1)</td>
<td>2.0 (1.9; 2.1)</td>
</tr>
<tr>
<td>Genitourinary</td>
<td>2.0 (1.9; 2.1)</td>
<td>2.0 (1.9; 2.1)</td>
<td>2.0 (1.9; 2.1)</td>
<td>2.0 (1.9; 2.1)</td>
<td>2.0 (1.9; 2.1)</td>
</tr>
<tr>
<td>Skin</td>
<td>2.0 (1.9; 2.1)</td>
<td>2.0 (1.9; 2.1)</td>
<td>2.0 (1.9; 2.1)</td>
<td>2.0 (1.9; 2.1)</td>
<td>2.0 (1.9; 2.1)</td>
</tr>
<tr>
<td>Blood</td>
<td>2.0 (1.9; 2.1)</td>
<td>2.0 (1.9; 2.1)</td>
<td>2.0 (1.9; 2.1)</td>
<td>2.0 (1.9; 2.1)</td>
<td>2.0 (1.9; 2.1)</td>
</tr>
<tr>
<td>Other</td>
<td>2.0 (1.8; 2.2)</td>
<td>2.0 (1.9; 2.2)</td>
<td>2.0 (1.9; 2.2)</td>
<td>2.0 (1.9; 2.2)</td>
<td>2.0 (1.9; 2.2)</td>
</tr>
</tbody>
</table>

*Disease groups sorted as in Table 1 of the main text

Table 2: For each disease group, estimates (and 95% confidence intervals) of the relative index of inequality in other-disease-related mortality (RII₀) obtained using various approaches
Figure 4: Cumulative baseline rate of mortality for each disease group according to the multiple-cause model with the weight-attribution strategy where $\omega = 0.75$, and the single-cause model.
<table>
<thead>
<tr>
<th>Disease group</th>
<th>Age of death and value of $\omega$ or weight-attribution strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>50-59 yrs</td>
</tr>
<tr>
<td></td>
<td>1(^b)</td>
</tr>
<tr>
<td>Neoplasms</td>
<td>2 1 1 1</td>
</tr>
<tr>
<td>Cardiovascular</td>
<td>3 3 3 3</td>
</tr>
<tr>
<td>Other</td>
<td>1 2 2 2</td>
</tr>
<tr>
<td>Nervous/sense</td>
<td>6 4 4 4</td>
</tr>
<tr>
<td>Digestive</td>
<td>5 5 5 5</td>
</tr>
<tr>
<td>Respiratory</td>
<td>7 9 9 9</td>
</tr>
<tr>
<td>Endocrine/nutritional</td>
<td>8 7 7 7</td>
</tr>
<tr>
<td>Mental</td>
<td>4 6 6 6</td>
</tr>
<tr>
<td>Infectious</td>
<td>9 8 8 8</td>
</tr>
<tr>
<td>Genitourinary</td>
<td>10 10 10 10</td>
</tr>
<tr>
<td>Blood</td>
<td>12 12 12 12</td>
</tr>
<tr>
<td>Skin</td>
<td>13 13 13 13</td>
</tr>
</tbody>
</table>

\(^a\)Disease groups sorted by increasing rank based on the analysis with $\omega = 0.75$ at age-group 70-79.

\(^b\)Single-cause model

\(^c\)Equal='Equal weights' strategy

Table 3: Ranking of disease groups by burden on mortality for ages 50-89, estimated by the cumulative baseline mortality rate attributable to the disease group according to various approaches at the mid-point of each age-group.
<table>
<thead>
<tr>
<th>Disease group</th>
<th>Age of death and value of $\omega$ or weight-attribution strategy</th>
<th>30-39 yrs</th>
<th>40-49 yrs</th>
<th>≥ 90 yrs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1(^b) 0.75 0.5 Equal(^c)</td>
<td>1(^b) 0.75 0.5 Equal(^c)</td>
<td>1(^b) 0.75 0.5 Equal(^c)</td>
<td></td>
</tr>
<tr>
<td>Neoplasms</td>
<td>2 2 2 2</td>
<td>2 2 2 2</td>
<td>2 2 2 2</td>
<td></td>
</tr>
<tr>
<td>Cardiovascular</td>
<td>3 3 3 3</td>
<td>3 3 3 3</td>
<td>1 1 1 1</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>1 1 1 1</td>
<td>1 1 1 1</td>
<td>3 3 3 3</td>
<td></td>
</tr>
<tr>
<td>Nervous/sense</td>
<td>4 4 4 4</td>
<td>4 4 4 4</td>
<td>4 4 4 4</td>
<td></td>
</tr>
<tr>
<td>Digestive</td>
<td>5 5 5 5</td>
<td>5 5 5 5</td>
<td>8 8 8 8</td>
<td></td>
</tr>
<tr>
<td>Respiratory</td>
<td>6 6 6 6</td>
<td>8 7 7 7</td>
<td>5 5 5 5</td>
<td></td>
</tr>
<tr>
<td>Endocrine/nutritional</td>
<td>7 7 7 7</td>
<td>9 8 8 8</td>
<td>6 6 6 6</td>
<td></td>
</tr>
<tr>
<td>Mental</td>
<td>8 8 8 8</td>
<td>6 6 6 6</td>
<td>7 7 7 7</td>
<td></td>
</tr>
<tr>
<td>Infectious</td>
<td>9 9 9 9</td>
<td>7 9 9 9</td>
<td>10 10 10 10</td>
<td></td>
</tr>
<tr>
<td>Genitourinary</td>
<td>10 10 10 10</td>
<td>10 10 10 10</td>
<td>9 9 9 9</td>
<td></td>
</tr>
<tr>
<td>Blood</td>
<td>12 12 12 12</td>
<td>12 12 12 12</td>
<td>12 12 12 12</td>
<td></td>
</tr>
<tr>
<td>Skin</td>
<td>13 13 13 13</td>
<td>13 13 13 13</td>
<td>13 13 12 12</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\)Disease groups sorted as in Table 3 of eAppendix.
\(^b\)Single-cause model
\(^c\)Equal='Equal weights' strategy

Table 4: Ranking of disease groups by burden on mortality for the youngest and eldest age-groups, estimated by the cumulative baseline mortality rate attributable to the disease group according to various approaches at the mid-point of each age-group.