# Causal Inference 2: Model-based estimation of causal contrasts

Thursday, 6 June, 2024

**Esa Läärä**

Statistical Practice in Epidemiology with **R**
3 to 7 June, 2024
International Agency for Research on Cancer, Lyon, France

# Outline

- Causal questions
- Factual risks and associational contrasts
- Causal estimands: contrasts of counterfactual quantities
- Marginal and conditional contrasts, effect among treated, etc.
- Outcome regression models, standardization or g-formula
- Exposure modelling, propensity scores and weighting
- Double robust estimators and machine learning algorithms
- Time-to-event outcomes: hazards of hazard ratios and estimation of causal contrasts of cumulative risks.

# Some literature


- Austin & Stuart (2015) *Stat Med* 34(28):3661-3679.
- Jonsson Funk et al. (2011) *Am J Epidemiol* 173(7):761-767
- Hernan & Robins (2020). *Causal Inference: What if*. CRC Press.
- Luque Fernandez et al. (2018) *Stat Med* 2018;37(16):2530-2546
- Schuler & Rose (2017) *Am J Epidemiol* 185(1):65-73.
- Sjölander (2016) *Eur J Epidemiol* 31:563-574
- Smith et al. (2022) *Stat Med* 2022;41(2):407-432.
- Zhou et al. (2022) *R Journal* 2022;14(1):282-289.

# Causal question in PECOT format & Example

P **Population**: 2900 women with breast cancer (Rotterdam study)

E **Exposure**: Hormonal treatment (HT)

C **Comparator**: Placebo, no HT

O **Outcome**: Recurrence or death

T **Time frame**: 10 y from surgery to outcome

Causal questions of interest – comparisons of **counterfactuals**:

- What is the 10-year risk $\pi^1$ of the outcome, if everybody in P were exposed to HT, as compared with $\pi^0$, the risk if nobody were exposed?
- What is the 10-year risk $\pi_1^1$ of the outcome, among those in P, who are factually exposed to HT, as compared with the risk $\pi_1^0$, if they were not exposed?

# Risks by factual exposure and their associational contrasts

- Let $Y$ be a binary indicator (1/0) for the *outcome* to occur within a fixed risk period (assuming no censoring, nor competing events), and $X$ be an *exposure* variable or risk factor.
- Let $\pi_x =$ *risk* of the outcome to occur during the period in the subset of the target population factually exposed to level $X = x$:

$$\pi_x = P\{Y = 1 \mid X = x\} = E(Y|X = x).$$

- For simplicity, let $X$ be binary: exposed $(X = 1)$ vs. unexposed $(X = 0)$.
- Common **associational contrasts** of risks between exposure groups:
  - **Risk difference** $\tau = \pi_1 - \pi_0$,
  - **Risk ratio** $\phi = \pi_1/\pi_0$,
  - **Risk odds ratio** $\psi = \dfrac{\omega_1}{\omega_0} = \dfrac{\pi_1/(1-\pi_1)}{\pi_0/(1-\pi_0)}$.

# Conditional associational contrasts

- The associational quantities above were **marginal**; not conditioned on (or stratified by) any covariate – such as sex, age, etc.
- Let now $Z$ be a **covariate** (can be multivariable) and

$$\pi_{xz} = P\{Y = 1 \mid X = x, Z = z\} = E(Y|X = x, Z = z)$$

be the risk of outcome during risk period in a population group where both $X = x$ and $Z = z$, $x = 0, 1$.

- **Conditional associational contrasts** between exposed and unexposed among those with $Z = z$.
  - $\tau_z = \pi_{1z} - \pi_{0z}$ is the risk difference conditional on $Z = z$, i.e. $z$-**specific** risk difference.
  - $\phi_z = \pi_{1z}/\pi_{0z}$ and $\psi_z = \pi_{1z}(1 - \pi_{1z})/[\pi_{0z}(1 - \pi_{0z})]$ are the $z$-specific risk ratio and odds ratio, respectively.

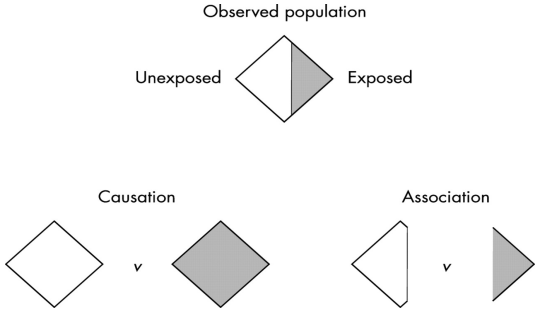
# Example: Single binary covariate $Z$

- Let the prevalence of exposure be $P\{X = 1\} = 0.45$ in the population
- Let $P\{Z = 1\} = 1 - P\{Z = 0\} = 0.40$ in the population and
$$P\{Z = 1|X = 1\} = 0.667 \quad \text{and} \quad P\{Z = 1|X = 0\} = 0.182$$
- Let also factual risks $\pi_{xz} = P\{Y = 1|X = x, Z = z\}$ $(x, z = 0, 1)$ by $X$ and $Z$ be as shown in the cells of the table below :

| | $Z = 1$ | $Z = 0$ | $\pi_x$ (obtained by formula of total probability) |
|---|---|---|---|
| $X = 1$ | 0.50 | 0.20 | $\pi_1 =$ **0.40** (0.50 × 0.667 + 0.20 × 0.333) |
| $X = 0$ | 0.25 | 0.10 | $\pi_0 =$ **0.13** (0.25 × 0.182 + 0.10 × 0.818) |
| Contrasts | $\tau_1 =$0.25 | $\tau_0 =$ 0.10 | $\tau =$ **0.27** |

- Marginal risks, $\pi_1, \pi_0$, contrast $\tau = \pi_1 - \pi_0$, and conditional contrasts $\tau_z = \pi_{1z} - \pi_{0z}$ are shown in table margins.

# Associational and causal contrasts




- **Associational**: Contrast of risks between the **subsets** of the population determined by the subjects' **factual** exposure value.
- **Causal**: Contrast of risks in the **entire population** under the alternative **potential** or **counterfactual** exposure values;
  see Hernan (2004), Hernan & Robins (2006), H&R (2020)

# Causal estimands: contrasts of counterfactual risks

- Let $Y^{X=x} = Y^x$ indicate (1/0) the event to occur within the risk period,

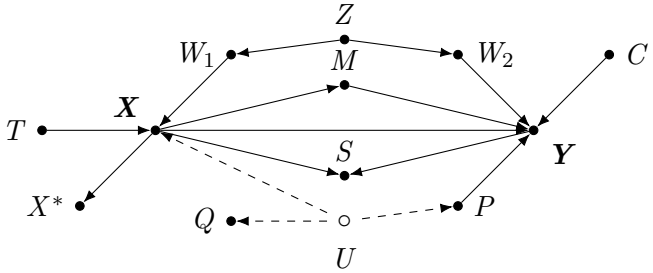
  **if** exposure $X$ were – **counterfactually** – forced to value $x$ in the whole target population.

- The **counterfactual** risk if everybody had exposure level $X = x$

$$\pi^x = P\{ Y^{X=x} = 1\} = E( Y^{X=x}), \quad x = 0, 1.$$

- **Marginal causal contrasts** of risk
  - risk difference (RD) $\tau^c = \pi^1 - \pi^0$,
  - risk ratio (RR) $\phi^c = \pi^1/\pi^0$,
  - risk odds ratio (OR) $\psi^c = [\pi^1/(1-\pi^1)]/[\pi^0/(1-\pi^0)]$,

# Identifying causal contrasts from causal diagram




- **Causal paths** $X \to Y$ and $X \to M \to Y$: Don't block!
- **Non-causal paths** between $X$ and $Y$: Block!
  – If already blocked, don't open (e.g. by conditioning on $\boxed{S}$).
- **Backdoor paths** $X \leftarrow W_1 \leftarrow Z \to W_2 \to Y$ and $X \leftarrow U \to P \to Y$: Block with minimal effort. – **Sufficient sets**: $P$ plus one from $\{Z, W_1, W_2\}$. – If $P$ unobserved, substitute by $Q$, proxy of $U$.
- No need to adjust for $T$. – Adjusting for $C$ can improve precision.

# Identifying causal contrasts from causal diagram

- Let $Z'$ be a set of observed covariates that are **non-descendants** of $X$
- If $Z \subset Z'$ were sufficient to **block** all **open non-causal paths** btw $X$ and $Y$, then counterfactuals are identified by **standardization** – or **g-formula**:

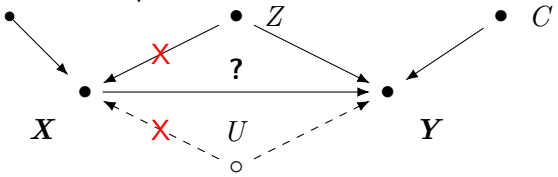
$$\pi^x = E(Y^{X=x}) = E_Z[E_Y(Y|X=x,Z)]$$
$$= \sum_z P\{Y=1 \mid X=x, Z=z\}P\{Z=z\}, \quad \text{for discrete } Z \ \& \ x=0,1.$$

- Causal contrasts $\tau^c$, $\phi^c$, $\psi^c$ are obtained from $\pi^1$ and $\pi^0$ thus derived.
- If there are open paths btw $X$ and $Y$, e.g. via unmeasured confounders $U$, the causal contrasts are not identified $\Leftrightarrow$ **residual confounding**.
- If $X$ is **randomized**, then $X \perp\!\!\!\perp Z \cup U$, and it holds simply

$$\pi^x = P\{Y^{X=x}=1\} = P\{Y=1 \mid X=x\} = \pi_x, \quad x=0,1.$$

# Randomized study and causal diagram

$R$ = **Randomization** of exposure




- When $X \equiv R$, no arrow points to $X$, and $X$ is independent of $Z, U, \ldots,$ measured and unmeasured.

$\Rightarrow$ No confounding!

$\Rightarrow$ Estimation of causal effect: unadjusted, crude comparison is enough.

- Precision may be improved by including $Z$ and $C$ as covariates.
- Often realized exposure $X$ is affected by $Z$ and $U$, thus differing from $R$. Then, $R$ may be utilized as an **instrumental variable**.

# Example (cont'd): Single binary common cause $Z$

- Causal diagram $X \to Y$, $X \leftarrow Z \to Y$; classical confounding triangle.
- Counterfactual risks (from items on slide 6) are obtained by g-formula $\pi^x = \sum_z \pi_{xz} P\{Z = z\}$, $x = 0, 1$, with weights from total population:

$$\pi^1 = 0.50 \times 0.4 + 0.20 \times 0.6 = \mathbf{0.32},$$
$$\pi^0 = 0.25 \times 0.4 + 0.10 \times 0.6 = \mathbf{0.16}$$

- Marginal causal contrasts (vs. associational ones)

$$\tau^c = 0.32 - 0.16 = \mathbf{0.16} \neq 0.27 = 0.40 - 0.13 = \tau,$$
$$\phi^c = 0.32/0.16 = \mathbf{2.00} \neq 3.14 = 0.40/0.13 = \phi,$$
$$\psi^c = \frac{0.32/(1-0.32)}{0.16/(1-0.16)} = \mathbf{2.47} \neq 4.57 = \psi.$$

- Associational contrasts were clearly confounded by $Z$.

# Conditional causal contrasts

- With covariate $Z$, counterfactual $z$-specific risks are defined

$$\pi_z^x = P\{\, Y^{X=x} = 1 \mid Z = z\}, \quad \text{for all } z \text{ and } x = 0, 1.$$

- These have their own identifiability conditions.
- **Conditional** or $z$**-specific causal contrasts** of risks are

$$\tau_z^c = \pi_z^1 - \pi_z^0, \qquad \phi_z^c = \pi_z^1/\pi_z^0, \qquad \psi_z^c = [\pi_z^1/(1-\pi_z^1)]/[\pi_z^0/(1-\pi_z^0)].$$

- If $\tau_z^c$ has the same value for all $z$, the risk difference is **homogenous**. Otherwise it is **heterogenous** or **modified** by $Z$. – These concepts are defined similarly for risk ratio and odds ratio.
- Homogeneity of one type of contrast implies heterogeneity of other types.
- There is an issue of **non-collapsibility** with the odds ratio. (Greenland et al. 1999, Sjölander et al. 2016, Didelez & Stensrud 2021)

# Causal contrasts in factual exposure groups

- Causal risk difference **among exposed** is defined
$$\tau_1^c = P\{Y^{X=1} = 1 \mid X = 1\} - P\{Y^{X=0} = 1 \mid X = 1\},$$
also known as **average treatment effect among treated** (ATT).
– The contrast **among unexposed** (ATU) is analogously defined.
- The effect often heterogenous, and groups noncomparable.
- If $Z$ is a sufficient set, g-formulas for identifying these are
$$\text{ATT} = \pi_1 - \sum \pi_{0z} P\{Z = z|X = 1\} = \text{“observed} - \text{expected”},$$
$$\text{ATU} = \sum \pi_{1z} P\{Z = z|X = 0\} - \pi_0 = \text{“expected} - \text{observed”}.$$

- Different standard populations for ATT, ATU, and for marginal contrast, a.k.a. **average treatment effect in the whole population**:
$$\text{ATE} = \tau^c = \pi^{X=1} - \pi^{X=0} = \sum_z \pi_{1z} P\{Z = z\} - \sum_z \pi_{0z} P\{Z = z\}.$$

# Example: Single binary $Z$ (cont'd)

- $z$-specific risks, marginal assoc. & causal contrasts are on slides 6 & 12.
- For ATT, we have the observed risk $\pi_1^1 = \pi_1 = \mathbf{0.40}$, and the expected risk is $\pi_1^0 = \sum_z \pi_{0z} P\{Z = z|X = 1\} = 0.25{\times}0.667 + 0.10{\times}0.333 = \mathbf{0.20}$, so ATT $= 0.40 - 0.20 = \mathbf{0.20}$.
- For ATU, the expected risk is $\pi_0^1 = \sum_z \pi_{1z} P\{Z = z|X = 0\} = 0.50{\times}0.182 + 0.20{\times}0.818 = \mathbf{0.26}$, the observed risk is $\pi_0^0 = \pi_0 = \mathbf{0.13}$, and ATU $= 0.26 - 0.13 = \mathbf{0.13}$.
- Here, the causal risk difference is bigger among exposed. – Being exposed seems to be a modifier of the effect of exposure on this scale!
- Interestingly, the causal risk ratio $= 2$ is homogenous.

**NB** Popular design for estimating ATT: **matched cohort study**.

# Outcome regression modelling (see lecture on Monday)

Modelling how expected values, risks, hazards, etc. depend on exposure $X$ and covariates $Z$ (modifiers, and/or confounders). – Common elements:

- Each subject $i$ $(i = 1, \ldots, n)$ has an own **profile**, i.e. vector $(x_i, z_i^{\mathsf{T}})$ of values of $X$ and covariates $Z$.
- In the spirit of **generalized linear models**, let vector $(\alpha, \beta, \gamma^{\mathsf{T}})$ contain regression coefficients, and specify the **linear predictor** – assuming so far no **interactions**, nor **effect modifications**

$$\eta_i = \alpha + \beta x_i + \gamma^{\mathsf{T}} z_i$$

- **Product terms** can be added for interactions and modifications if needed, and **splines** may be used for continuous covariates.
- Further model specification depends on the type of outcome variable, causal contrasts of interest, and importance and choice of time scale(s).

# Binary outcome model and classical causal estimation

- Basic outcome regression model for risks $\pi$ in fixed risk periods:

$$g\{\pi(x_i)\} = \alpha + \beta x_i + \gamma^{\mathsf{T}} z_i, \quad i = 1, \ldots, n.$$

- Link $g(\cdot)$ and causal interpretation of $\beta$, assuming the validity of model (including homogeneity or non-modification of the contrast in question) and that $Z$ blocks all backdoor paths:
  - id $\Rightarrow \beta =$ risk difference (RD) $\tau^c$ for $X = 1$ vs. $X = 0$, adjusted for $Z$
  - log $\Rightarrow \beta =$ log of risk ratio (RR) $\phi^c$ – " –
  - logit $\Rightarrow \beta =$ log of conditional risk odds ratio (OR), $\psi_z^c$, – " –
    **NB.** This is different from marginal OR due to **non-collapsibility**.

- Random component: Binomial family – Model fitting: `glm()`.

- Problems with id & log in keeping predicted $\widehat{\pi}(\cdot)$ between $0$ and $1$.

# Modern approach: Causal contrasts by g-formula

- Assuming that $Z$ is sufficient to block non-causal paths, a logistic model is fitted, which may even contain product terms allowing modification

$$\text{logit}(\pi_i) = \log[\pi_i/(1-\pi_i)] = \alpha + \beta x_i + \gamma^{\mathsf{T}} z_i + \delta^{\mathsf{T}}(x_i z_i), \quad i = 1, \ldots, n.$$

- In R, for instance, with two covariates

```
> mY <- glm(y ~ x + z1 + z2 + x:z1 + x:z2, family=binomial,data=dd)
```

- For each individual $i$, **predicted risks** are computed for both possibilities of exposure: $X_i = 1$ and $X_i = 0$, but keeping $Z_i = z_i$ as it is

$$\widetilde{\pi}_i^{X_i=x} = \text{expit}\{\widehat{\alpha} + \widehat{\beta} x + \widehat{\gamma}^{\mathsf{T}} z_i + \widehat{\delta}^{\mathsf{T}}(x z_i)\}, \quad x = 0, 1.$$

```
> dd$ypred1 <- predict( mY, type="response",   #  set x=1
                newdata=data.frame(x=rep(1,n), dd[ , c("z1", "z2")] ) )
> dd$ypred0 <- predict( mY, type="response",   #  set x=0
                newdata=data.frame(x=rep(0,n), dd[ , c("z1", "z2")] ) )
```

# Modern approach: Causal contrasts by g-formula (cont'd)

- Marginal potential or counterfactual risks for $x = 1, 0$ are estimated applying the principle of **standardization** or **g-formula**:

$$\widehat{\pi}^{X=x} = \widehat{E}_Z[E(Y|X = x, Z)] = \frac{1}{n}\sum_{i=1}^{n} \widetilde{E}(Y_i|X_i = x, Z = z_i) = \frac{1}{n}\sum_{i=1}^{n} \widetilde{\pi}_i^{X_i=x}$$

  as the data provide a non-parametric estimate of the joint distribution of $Z$.

```
> EYpot.1 <- mean(dd$ypred1)
> EYpot.0 <- mean(dd$ypred0)
```

- **Marginal causal contrasts** of risks are now estimated, e.g.

$$\widehat{\tau}^c = \widehat{\pi}^{X=1} - \widehat{\pi}^{X=0}, \qquad \widehat{\psi}^c = [\widehat{\pi}^1/(1-\widehat{\pi}^1)]/[\widehat{\pi}^0/(1-\widehat{\pi}^0)]$$

```
> tau.c <- EYpot.1 - EYpot.0
> phi.c <- EYpot.1 / EYpot.0
> psi.c <- ( EYpot.1/(1-EYpot.1) ) / ( EYpot.0/(1-EYpot.0) )
```

# Exposure modelling, propensity scores and weighting

Let $X$ be a binary exposure variable. Assume again that $Z$ is a sufficient set

- **Exposure model** predicting individual $X_i$:s by confounders is fitted

$$\text{logit}[P\{X_i = 1|Z = z_i\}] = \alpha^* + z_i^{\mathsf{T}}\gamma^*, \quad i = 1, \ldots, n.$$

```
> mX <- glm(x ~ z1 + z2 + z1:z2, family=binomial, data=dd)
```

- **Propensity scores** $\text{PS}_i$, or fitted probabilities of exposure are obtained

$$\text{PS}_i = \widehat{P}\{X_i = 1|Z = z_i\} = \text{expit}(\widehat{\alpha}^* + z_i^{\mathsf{T}}\widehat{\gamma}^*).$$

```
> dd$PS <- predict(mX, type="response")
```

- Compute **inverse probability weights** (IPW), i.e. inverses of fitted probabilities of belonging to the realized exposure group.

$$W_i = w(\text{PS}_i, X_i) = \frac{\mathbf{1}_{\{X_i=1\}}}{\text{PS}_i} + \frac{\mathbf{1}_{\{X_i=0\}}}{1 - \text{PS}_i}. \quad i = 1, \ldots, n,$$

```
> dd$w <- 1*(dd$x==1)/dd$PS + 1*(dd$x==0)/(1-dd$PS)
```

# Exposure modelling, propensity scores and weighting (cont'd)

- With IPW, counterfactual risks in the whole population are estimated as weighted averages of the outcome in the two exposure groups

$$\widehat{\pi}^{X=x} = \frac{\sum_{i=1}^{n} \mathbf{1}_{\{X_i=x\}} W_i Y_i}{\sum_{i=1}^{n} \mathbf{1}_{\{X_i=x\}} W_i} = \frac{\sum_{X_i=x} W_i Y_i}{\sum_{X_i=x} W_i}, \quad x = 0, 1$$

```
> EYpot.1 <- sum( 1*(dd$x==1)*dd$w*dd$y ) / sum( 1*(dd$x==1)*dd$w )
> EYpot.0 <- sum( 1*(dd$x==0)*dd$w*dd$y ) / sum( 1*(dd$x==0)*dd$w )
```

and from these, marginal causal contrasts are estimated as before.

- For causal contrasts **among the treated** (ATT), use **treated weights**:

$$W_i = 1 \text{ for } X_i = 1, \text{ and } W_i = \mathsf{PS}_i/(1 - \mathsf{PS}_i) \text{ for } X_i = 0.$$

- Other: **overlap weights**, **matching weights**, **entropy weights**.
- The goodness-of-fit of the exposure model needs to be assessed.
  For that purpose, various measures of **covariate balance** are developed.

# Double robust (DR) estimators and machine learning methods

- The validity of estimation utilizing g-formula or PS-based weighting depends on, how accurately the outcome model or exposure model is specified.
- **Double robust** (DR) estimation of causal contrasts:
  Combination of g-formula and IPW. – Alternatives
  - **Augmented IPW** (AIPW); see Jonsson Funk et al. (2011),
  - **Targeted maximum likelihood estimation** (TMLE);
    see Schuler & Rose (2017), Luque-Fernandez et al. (2018)

  Validity of a DR estimator requires that either the exposure model or the outcome model (or both) is correctly specified.

- Algorithms developed for **supervised learning** increase flexibility in modelling both outcome and exposure (see Bi et al. 2019, Blakely et al. 2020).

# Time-to-event outcomes: associational hazard quantities

- Let $T$ = time to outcome event from a defined zero time, and $Y(t) = \mathbf{1}_{\{T \leq t\}}$ indicator (1/0) for the outcome to occur by $t$.
- The **hazard** of outcome at $t$, $\lambda_x(t)$, and the **risk** of outcome during $(0, t]$, $\pi_x(t)$, for those factually exposed to level $X = x$, $x = 1, 0$:

$$\lambda_x(t) = \lambda(t \mid X = x) \approx P\{Y(t+h) = 1 \mid Y(t) = 0, X = x\}/h,$$

$$\pi_x(t) = P\{Y(t) = 1 | X = x\} = 1 - \exp\left(-\int_0^t \lambda_x(v)dv\right)$$

- Common associational contrasts of hazards:
  - **Hazard difference** $\delta(t) = \lambda_1(t) - \lambda_0(t)$,
  - **Hazard ratio** $\rho(t) = \lambda_1(t)/\lambda_0(t)$.
    This is often assumed constant $\rho$ – as in **Cox regression**.

# Causal contrasts of hazards

- Let $T^{X=x} = T^x$, be time to event, and $Y^{X=x}(t) = Y^x(t) = \mathbf{1}_{\{T^x \leq t\}}$ indicate the event occurring during risk period $(0, t]$, if exposure $X$ were forced to value $x$ in the whole target population.
- The **counterfactual hazard**, if everybody were exposed to $X = x$:

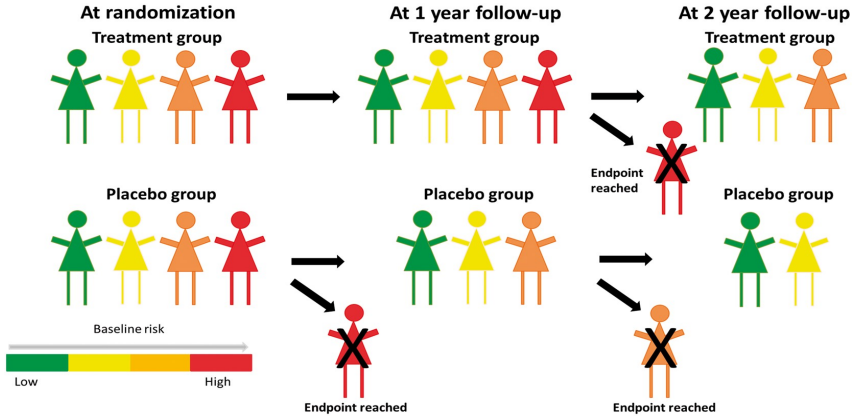
$$\lambda^x(t) \approx P\{Y^x(t+h) = 1 | Y^x(t) = 0\}/h, \quad x = 1, 0.$$

- Marginal contrasts of these counterfactuals can be defined: hazard difference (HD) $\lambda^1(t) - \lambda^0(t)$, and hazard ratio (HR) $\lambda^1(t)/\lambda^0(t)$.
- If $X$ is randomized, these are identified by corresp. assoc. contrasts.
- Yet, hazard at any $t$ is conditional on survival by $t$. If $X$ has any effect, $Y^1(t) = 0$ and $Y^0(t) = 0$ imply different populations at risk for $t > 0$.

$\Rightarrow$ Even if exposure groups were comparable at $t = 0$, after that they aren't.

- Causal interpretation of HR problematic even in a randomized study.

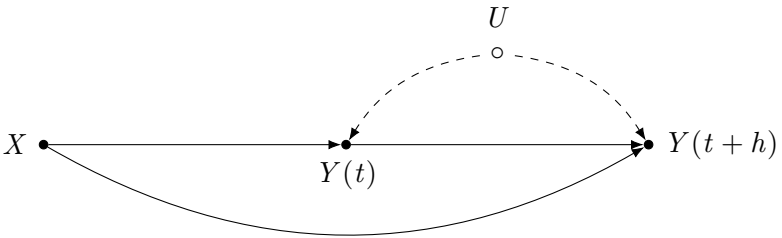
# Example: The untreated have a higher hazard (Stensrud et al 2019)




- In the course of time, the prognostic profile of the remaining active treatment group will be worse than that in the remaining placebo group.

# Hazard of hazard ratios (Hernan 2010, Aalen et al. 2015 )

- The hazard at any time $t > 0$ is affected by known and unknown causes of the outcome $\Rightarrow$ individual **frailty** $U$ varies in the population.




- $Y(t)$ is a **collider** on the path from $X$ to $Y(t+h)$ via $U$.
  Conditioning on $Y(t) = 0$ opens this non-causal path $\Rightarrow$ **selection bias**.
- The observable hazards may behave strangely over time and lead to conclusions like "HR$> 1$ before $t^*$ but HR$< 1$ after that".

# Example: WHI Trial on MHT and CHD (Manson et al. 2003)

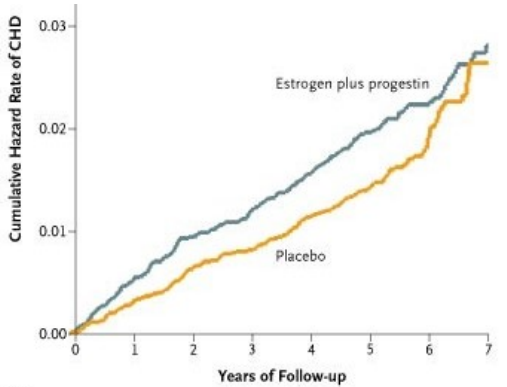
**Table 2.** Estrogen plus Progestin and the Risk of CHD, According to Year of Follow-up.*

| Year of Follow-up | CHD | | Hazard Ratio for CHD (95% CI) |
|---|---|---|---|
| | Estrogen-plus-Progestin Group | Placebo Group | |
| | *no. of cases (annualized percentage)* | | |
| 1 | 42 (0.50) | 23 (0.29) | 1.81 (1.09–3.01) |
| 2 | 38 (0.45) | 28 (0.35) | 1.34 (0.82–2.18) |
| 3 | 19 (0.23) | 15 (0.19) | 1.27 (0.64–2.50) |
| 4 | 32 (0.39) | 25 (0.32) | 1.25 (0.74–2.12) |
| 5 | 29 (0.41) | 19 (0.28) | 1.45 (0.81–2.59) |
| ≥6 | 28 (0.37) | 37 (0.56) | 0.70 (0.42–1.14) |

* CHD includes acute myocardial infarction (MI) necessitating hospitalization, silent myocardial infarction as determined by serial electrocardiography, and death due to CHD. There were nine silent myocardial infarctions (four in the estrogen-plus-progestin group and five in the placebo group). Hazard ratios are stratified according to age, presence or absence of a previous coronary event, and randomly assigned diet-modification group and are adjusted for previous coronary-artery bypass grafting or percutaneous transluminal coro-

- Women, 50-79 y; MHT: $N_1 = 8506$, placebo: $N_0 = 8102$
- Follow-up max 8.6 y, mean 5.6 y.
- Cases & rates/$10^4$ y
  $D_1 = 188$, $I_1 = 39$,
  $D_0 = 147$, $I_0 = 33$.
- Crude IR $= 1.20$,
  adjusted 1.24 (1.00–1.54)
- Effect of MHT? Increases the risk at start, but after 6 years reduces the risk? Overall effect still harmful?

# Example: WHI Trial (cont'd)




- Curves of cumulative hazard approximate the cumulative risks over time.
- In early years, the curve for MHT runs on top, reflecting higher hazard then in that group.
- After some years there are less frail subjects on MHT than on placebo in the risk set.
- Increased hazard after 6 y in the remaining placebo group has lifted its cumulative risk curve to the same level as in MHT group.

# Estimation of causal contrasts of risks

- Counterfactual risks $\pi^{X=x}(t)$ and their contrasts are causally interpretable. Various methods to estimate them exist (see Denz et al. 2023 ) – e.g.

(a) Fit a hazard model, e.g. $\lambda(t|x_i, z_i) = \lambda_0(t)\exp(\beta x_i + \gamma^T z_i)$, extract $\widehat{\beta}$, $\widehat{\gamma}$, and baseline cumulative hazard $\widehat{\Lambda}_0(t)$ and plug in them:

$$\widetilde{\pi}_i^{X_i=x}(t) = 1 - \exp\{-\widehat{\Lambda}_0(t)\exp(\widehat{\beta}x + \widehat{\gamma}^{\mathsf{T}} z_i)\}.$$

Counterfactuals $\pi^{X=x}(t)$ and contrasts are then estimated by g-formula.

(b) Get weights $W_i$ from PS$_i$s of an exposure model, fit Cox with "intercept only" specifying $X$ as a `strata()` variable and $W_i$:s as `weights`, and estimate $\widehat{\pi}^{X=x}(t)$ using `survfit()`, etc.

- Other: IPW Kaplan-Meier, use of pseudo-values, DR methods, ...
- Competing event setting: additional complexities in defining and analysing causal contrasts (see Rudolph et al. 2020, Young et al. 2020).

# Conclusions

- Careful specification of causal question and estimands needed.
- Adjustment for confounding via efficient blocking of backdoor paths.
- Basic estimation methods: outcome regression & g-formula, exposure modelling & PS-weighting, double robust estimators.
- Sufficiently flexible models desirable to reduce misspecification bias.
- Statistical inference (ignored here): robust covariance matrix & delta method, bootstrapping, efficient influence curve, etc.
- Extensions exist for polytomous or time-varying exposure and confounders.
- Time-to-event outcomes: contrasts of risks or survival preferable as causal estimands; there are hazards with hazard ratios.
- There can still remain open non-causal paths between $X$ and $Y$ inducing residual confounding and/or selection bias.