

Midterm 2 Review Session

Contents

Configuring R	1
1 Introduction	2
2 Part 1: Kaplan-Meier and Nelson-Aalen	2
2.1 The risk set shrinks by the number of subjects who <i>leave</i>	3
2.2 Nelson-Aalen estimates the <i>cumulative hazard</i> first, then exponentiates	4
2.3 Survival estimates are <i>step functions</i> : use the most recent event time	4
2.4 Median survival time is read off the curve, not modeled	5
2.5 How censored times are used: denominator yes, numerator no	5
3 Part 2: Cox proportional hazards models	6
3.1 Writing the model: name every assumption and show where it is used	6
3.2 Interpreting a hazard ratio: magnitude, reference, adjustment, <i>and</i> significance	7
3.3 A hazard ratio for a non-unit change: exponentiate	8
3.4 Comparing two coefficients: variance of a <i>difference</i>	8
3.5 A confidence interval for a hazard ratio: build it on the log scale, then exponentiate	9
4 Summary of the most common errors	10

Configuring R

Functions from these packages will be used throughout this document:

```
library(conflicted) # check for conflicting function definitions
# library(printr) # inserts help-file output into markdown output
library(rmarkdown) # Convert R Markdown documents into a variety of formats.
library(pander) # format tables for markdown
library(ggplot2) # graphics
library(ggfortify) # help with graphics
library(dplyr) # manipulate data
library(tibble) # `tibble`s extend `data.frame`s
library(magrittr) # `%>%` and other additional piping tools
library(haven) # import Stata files
library(knitr) # format R output for markdown
library(tidyr) # Tools to help to create tidy data
library(plotly) # interactive graphics
library(dobson) # datasets from Dobson and Barnett 2018
library(parameters) # format model output tables for markdown
library(haven) # import Stata files
library(latex2exp) # use LaTeX in R code (for figures and tables)
library(fs) # filesystem path manipulations
library(survival) # survival analysis
library(survminer) # survival analysis graphics
library(KMsurv) # datasets from Klein and Moeschberger
```

```
library(parameters) # format model output tables for
library(webshot2) # convert interactive content to static for pdf
library(forcats) # functions for categorical variables ("factors")
library(stringr) # functions for dealing with strings
library(lubridate) # functions for dealing with dates and times
library(broom) # Summarizes key information about statistical objects in tidy tibbles
library(broom.helpers) # Provides suite of functions to work with regression model 'broom::tidy()' t
```

Here are some R settings I use in this document:

```
rm(list = ls()) # delete any data that's already loaded into R

conflicts_prefer(dplyr::filter)
ggplot2::theme_set(
  ggplot2::theme_bw() +
    # ggplot2::labs(col = "") +
    ggplot2::theme(
      legend.position = "bottom",
      text = ggplot2::element_text(size = 12, family = "serif")))

knitr::opts_chunk$set(message = FALSE)
options('digits' = 6)

panderOptions("big.mark", ",")
pander::panderOptions("table.emphasize.rownames", FALSE)
pander::panderOptions("table.split.table", Inf)
conflicts_prefer(dplyr::filter) # use the `filter()` function from dplyr() by default
legend_text_size = 9
run_graphs = TRUE
```

1 Introduction

This chapter walks through the **most common mistakes** that students made on Midterm 2, which covered survival analysis and Cox proportional hazards models. Each section states a mistake, explains *why* it is wrong, and works through the correct approach.

The exam had two parts:

- **Part 1** (Kaplan-Meier and Nelson-Aalen estimators) used a small data set of survival times: 10, 14, 14+, 15+, 18, 21, 25 months (the + marks censored observations).
- **Part 2** (Cox proportional hazards regression) used a model fit to $n = 3,142$ men from the Western Collaborative Group Study (WCGS), relating time to incident coronary heart disease (CHD) to baseline cigarette smoking and other covariates.

Tip

The mistakes in this chapter are organized so that the most consequential and most frequent errors come first within each part. If you only have time to review a few things, start at the top of each part.

2 Part 1: Kaplan-Meier and Nelson-Aalen

```
library(survival)
library(dplyr)

surv_data <- tibble(
```

```
time = c(10, 14, 14, 15, 18, 21, 25),
death = c(1, 1, 0, 0, 1, 1, 1),
surv = Surv(time, death)
)
```

Here is the full table that everyone was trying to reproduce; we will refer back to it throughout this part.

```
KM_est <- survfit(surv ~ 1, type = "kaplan-meier", data = surv_data)

surv_table <-
  KM_est |>
  summary(censored = TRUE, data.frame = TRUE) |>
  as_tibble() |>
  mutate(
    hazard      = n.event / n.risk,
    nonhazard    = 1 - hazard,
    cusum_hazard = cumsum(hazard),
    surv_KM      = cumprod(nonhazard),
    surv_NA      = exp(-cusum_hazard)
  ) |>
  select(time, n.risk, n.event, hazard, surv_KM, cusum_hazard, surv_NA)

surv_table |> pander::pander()
```

Table 1: Kaplan-Meier and Nelson-Aalen calculations for the Part 1 data

time	n.risk	n.event	hazard	surv_KM	cusum_hazard	surv_NA
10	7	1	0.1429	0.8571	0.1429	0.8669
14	6	1	0.1667	0.7143	0.3095	0.7338
15	4	0	0	0.7143	0.3095	0.7338
18	3	1	0.3333	0.4762	0.6429	0.5258
21	2	1	0.5	0.2381	1.143	0.3189
25	1	1	1	0	2.143	0.1173

2.1 The risk set shrinks by the number of subjects who *leave*

Exam reference: **Exercise 1.1** (6 points) — compute the KM and Nelson-Aalen estimates of $S(t)$.

Common mistake

Reducing the number at risk by **one** at every row, regardless of how many subjects actually left the study. With this mistake, the number at risk after the two censored subjects (at $t = 14^+$ and $t = 15^+$) was computed as 5 instead of 4.

Solution

Solution. The number at risk n_j is the number of subjects still under observation *just before* time t_j — that is, everyone who has neither had the event nor been censored yet.

Track the cohort of 7:

time	what happens	number at risk <i>before</i> this time
10	1 death	7
14	1 death and 1 censored	6
15	1 censored	4

18	1 death	3
21	1 death	2
25	1 death	1

At $t = 14$, one subject dies *and* a different subject is censored (the 14+), so **two** subjects leave the risk set; the count drops from 6 to 4, not from 6 to 5. Every subject who has an event **or** is censored is removed from all later risk sets.

2.2 Nelson-Aalen estimates the *cumulative hazard* first, then exponentiates

Exam reference: **Exercise 1.1** (6 points) — compute the KM and Nelson-Aalen estimates of $S(t)$.

⚠ Common mistake

Two versions of this error were common:

1. Computing a running sum of arbitrary quantities (e.g. summing survival probabilities, or summing $1 - \hat{\lambda}_j$) instead of summing the hazards $\hat{\lambda}_j = d_j/n_j$.
2. Computing the cumulative hazard correctly but then forgetting the final step, $\hat{S}_{\text{NA}}(t) = \exp\{-\hat{\Lambda}(t)\}$, and reporting the cumulative hazard itself (or the KM product) as the Nelson-Aalen survival estimate.

Solution

Solution. The Nelson-Aalen estimator builds the **cumulative hazard** by adding up the instantaneous hazard contributions at each event time:

$$\hat{\Lambda}(t) = \sum_{j:t_j \leq t} \frac{d_j}{n_j}$$

and then converts it to a survival estimate using $S(t) = \exp\{-\Lambda(t)\}$:

$$\hat{S}_{\text{NA}}(t) = \exp\{-\hat{\Lambda}(t)\}$$

For example, at $t = 10$:

$$\hat{\Lambda}(10) = \frac{1}{7} = 0.143, \quad \hat{S}_{\text{NA}}(10) = \exp\{-0.143\} = 0.867$$

Contrast this with Kaplan-Meier, which multiplies conditional survival probabilities:

$$\hat{S}_{\text{KM}}(t) = \prod_{j:t_j \leq t} \left(1 - \frac{d_j}{n_j}\right), \quad \hat{S}_{\text{KM}}(10) = 1 - \frac{1}{7} = 0.857$$

The two estimators are close but **not** identical, and $\hat{S}_{\text{NA}}(t) \geq \hat{S}_{\text{KM}}(t)$ always. Reporting the same numbers for both, or reporting $\hat{\Lambda}$ where $\hat{S}_{\text{NA}}$ was asked for, loses credit.

2.3 Survival estimates are *step functions*: use the most recent event time

Exam reference: **Exercise 1.3** (1 point) — the KM and NA survival estimates for $t = 17$ months.

⚠ Common mistake

For $\hat{S}(17)$, looking at the row for $t = 18$ (the next event *after* 17), or interpolating between rows. Some answers gave only a verbal description (“it’s flat there”) without the numeric value.

Solution

Solution. $\hat{S}(t)$ is a **right-continuous step function**: it only changes at observed event times and stays flat in between. To evaluate it at $t = 17$, find the **most recent event time at or before 17**, which is $t = 14$ (the subsequent $t = 15^+$ is a censoring, not an event, so the curve does not step there). So:

$$\hat{S}_{\text{KM}}(17) = \hat{S}_{\text{KM}}(14) = 0.714, \quad \hat{S}_{\text{NA}}(17) = \hat{S}_{\text{NA}}(14) = 0.734$$

Give the actual numbers, not just “it’s flat.”

2.4 Median survival time is read off the curve, not modeled

*Exam reference: **Exercise 1.4** (1 point) — the KM and NA estimates of median survival time.*

⚠ Common mistake

Estimating the median by fitting an exponential model ($\hat{\lambda} = \text{events}/\text{total follow-up}$, then $\hat{E}[T] = 1/\hat{\lambda}$). That computes an exponential-model **mean**, not the nonparametric **median**, and answers a different question. A second error was reading the median off the wrong curve or the wrong row.

Solution

Solution. The median survival time is the smallest t at which the estimated survival curve drops to (or below) 0.5:

$$\hat{t}_{\text{median}} = \min\{t : \hat{S}(t) \leq 0.5\}$$

Reading from Table 1:

- $\hat{S}_{\text{KM}}(t)$ first reaches ≤ 0.5 at $t = 18$ (where it drops to 0.476), so the KM median is **18 months**.
- $\hat{S}_{\text{NA}}(t)$ first reaches ≤ 0.5 at $t = 21$ (where it drops to 0.319), so the NA median is **21 months**.

We can confirm this in R:

```
quantile(KM_est, p = 0.5)$quantile
#> 50
#> 18
```

No distributional assumption (exponential or otherwise) is used: the Kaplan–Meier estimator is a nonparametric read-off of the estimated curve.

2.5 How censored times are used: denominator yes, numerator no

*Exam reference: **Exercise 1.5** (2 points) — describe how the censored time(s) are utilized in the Kaplan–Meier estimate.*

⚠ Common mistake

Describing only the *symptom* (“the curve doesn’t drop at a censoring time”) without the *mechanism*, or implying that censored subjects are simply dropped from the analysis entirely.

Solution

Solution. A censored subject contributes to the **risk-set denominators** (n_j) for every event time up to and including their censoring time, because we know they survived event-free until

then. They **never** contribute to an event **numerator** (d_j), because no event was observed for them.

In this example, subjects #3 and #4 (censored at 14^+ and 15^+) are counted in the number at risk at $t = 10$ and $t = 14$, which is why those denominators are 7 and 6. After their censoring times they drop out of all later risk sets. This mechanism is exactly what lets Kaplan-Meier use partial information from censored subjects instead of discarding them.

3 Part 2: Cox proportional hazards models

For reference, here is the fitted model from the exam. The primary exposure is baseline smoking category (nonsmoker = reference, light smoker $L = 1$, heavy smoker $H = 1$), adjusting for age in decades (A), behavior pattern (P), overweight (B), and high cholesterol (C).

Table 2: Cox PH regression for the WCGS CHD data

Characteristic	$\log(\widehat{HR})$	$\widehat{HR}$	95% CI for HR	p
nonsmoker (ref.)	—	—	—	—
light smoker (L)	0.36	1.43	(1.05, 1.95)	0.023
heavy smoker (H)	0.78	2.18	(1.63, 2.91)	< 0.001
age per decade (A)	0.66	1.94	(1.56, 2.40)	< 0.001
behavior type A (P)	0.73	2.06	(1.58, 2.70)	< 0.001
overweight (B)	0.30	1.34	(1.05, 1.72)	0.019
high cholesterol (C)	0.75	2.12	(1.65, 2.71)	< 0.001

The exam also provided the estimated covariance matrix of the coefficient estimates (on the $\log(HR)$ scale). The relevant entries are:

$$\text{Var}(\hat{\beta}_L) = 0.024840, \quad \text{Var}(\hat{\beta}_H) = 0.021905, \quad \text{Cov}(\hat{\beta}_L, \hat{\beta}_H) = 0.010504$$

3.1 Writing the model: name every assumption and show where it is used

*Exam reference: **Exercise 2.1** (10 points) — write the mathematical form of the proportional hazards model corresponding to Table 1.*

Common mistake

Several patterns lost the most points here:

- Treating the question as “interpret the coefficients” and writing only the linear predictor, omitting the **likelihood**, the **distribution functions**, and the **assumptions**.
- Listing assumption *names* without **showing where they are used**, or writing the math without naming the assumption.
- Adding a spurious “**baseline hazard is exponential**” assumption. The Cox model leaves the baseline hazard $\lambda_0(t)$ **unspecified**; assuming a parametric form is wrong.
- Omitting one or more of the standalone distribution functions (survival, density, hazard, log-hazard, cumulative hazard).

Solution

Solution. A full-credit answer connects the likelihood to the linear predictor through a chain of named components and assumptions. Using $\tilde{X} = (L, H, A, P, B, C)$:

Joint likelihood of the data set:

$$\mathcal{L} \stackrel{\text{def}}{=} p(\tilde{Y} = \tilde{y}, \tilde{D} = \tilde{d} \mid \mathbf{X} = \mathbf{x})$$

Marginal likelihood contribution of observation i :

$$\mathcal{L}_i \stackrel{\text{def}}{=} p(Y_i = y_i, D_i = d_i \mid \tilde{X}_i = \tilde{x}_i)$$

Independent-observations assumption (used to factor the likelihood):

$$\mathcal{L} = \prod_{i=1}^n \mathcal{L}_i$$

Non-informative censoring assumption (used so each contribution reduces to survival and hazard terms): $T_i \perp\!\!\!\perp C_i \mid \tilde{X}_i$, giving

$$\mathcal{L}_i \propto [f(y_i \mid \tilde{x}_i)]^{d_i} [S(y_i \mid \tilde{x}_i)]^{1-d_i} = S(y_i \mid \tilde{x}_i) \cdot [\lambda(y_i \mid \tilde{x}_i)]^{d_i}$$

Distribution functions (define each one):

$$S(t \mid \tilde{x}) \stackrel{\text{def}}{=} P(T > t \mid \tilde{X} = \tilde{x}) = \exp\{-\Lambda(t \mid \tilde{x})\}$$

$$f(t \mid \tilde{x}) \stackrel{\text{def}}{=} \lambda(t \mid \tilde{x}) S(t \mid \tilde{x})$$

$$\lambda(t \mid \tilde{x}) \stackrel{\text{def}}{=} p(T = t \mid T \geq t, \tilde{X} = \tilde{x}) = \frac{f(t \mid \tilde{x})}{S(t \mid \tilde{x})}$$

$$\Lambda(t \mid \tilde{x}) \stackrel{\text{def}}{=} \int_0^t \lambda(u \mid \tilde{x}) du = -\log S(t \mid \tilde{x})$$

$$\eta(t \mid \tilde{x}) \stackrel{\text{def}}{=} \log \lambda(t \mid \tilde{x})$$

Proportional-hazards assumption (used to split the hazard into a baseline that depends only on time and a factor that depends only on covariates):

$$\lambda(t \mid \tilde{x}) = \lambda_0(t) \cdot \theta(\tilde{x})$$

where $\lambda_0(t)$ is the **unspecified** baseline hazard.

Logarithmic-link assumption (used to make the covariate factor a function of a linear predictor):

$$\eta(t \mid \tilde{x}) = \eta_0(t) + \Delta\eta(\tilde{x}), \quad \theta(\tilde{x}) = \exp\{\Delta\eta(\tilde{x})\}$$

Linear functional-form assumption (used to write the covariate term as a linear combination):

$$\Delta\eta(\tilde{x}) = \tilde{x} \cdot \tilde{\beta} = \beta_L l + \beta_H h + \beta_A a + \beta_P p + \beta_B b + \beta_C c$$

Notice that the baseline hazard $\lambda_0(t)$ is carried along symbolically the whole time — we never assume a shape for it.

3.2 Interpreting a hazard ratio: magnitude, reference, adjustment, and significance

*Exam reference: **Exercise 2.2** (4 points) — summarize how baseline smoking category is associated with hazard of incident CHD.*

 Common mistake

- Reporting the direction of the effect but omitting the **magnitude** (“X% higher/lower hazard”) or the **reference group**.
- Forgetting to say the estimate is **adjusted for** / **holding constant** the other covariates.
- Omitting **statistical significance** (whether the 95% CI excludes 1, or $p < 0.05$).
- Calling the hazard ratio a **risk** or an **odds ratio**. It is a ratio of *hazards*, not of risks or odds.

Solution

Solution. A complete interpretation has four parts: magnitude, reference group, adjustment, and significance. For the smoking effect in Table 2:

Adjusting for age, behavior pattern, overweight, and cholesterol, **light smokers** had an estimated hazard of incident CHD about **43% higher** than **nonsmokers** ($\widehat{HR} = 1.43$); this hazard ratio is statistically significant at the 0.05 level, since the 95% CI (1.05, 1.95) excludes 1 ($p = 0.023$).

Heavy smokers had an estimated hazard about **118% higher** (roughly double; $\widehat{HR} = 2.18$) than nonsmokers, also adjusting for the other covariates; this hazard ratio is highly significant ($p < 0.001$, CI (1.63, 2.91)).

Note “43% higher” comes from $1.43 - 1 = 0.43$, and “118% higher” from $2.18 - 1 = 1.18$. Interpret each smoking level **separately** against the reference, rather than lumping them together.

3.3 A hazard ratio for a non-unit change: exponentiate

Exam reference: Exercise 2.3 (2 points) — the hazard ratio associated with a 7.5-year increase in age.

 Common mistake

For the hazard ratio associated with a **7.5-year** increase in age (with the table reporting the HR per **decade**), the errors were:

- Multiplying: $1.94 \times 0.75 = 1.46$. **Wrong** — the HR for a multi-unit change is *not* linear in the HR.
- Arithmetic slips that produced answers like 2.08.

Solution

Solution. On the **log** scale, the effect is linear: a change of c units multiplies the log-hazard by $c \cdot \beta$. So the hazard ratio for a c -unit change is

$$HR(c) = \exp\{c\beta\} = (\exp\{\beta\})^c = HR^c$$

A 7.5-year increase is $c = 0.75$ decades, and the per-decade $\widehat{HR} = 1.94$ (i.e. $\hat{\beta}_A = \log 1.94 = 0.66$):

$$\widehat{HR}(0.75) = 1.94^{0.75} = \exp\{0.75 \times 0.66\} = \exp\{0.495\} = 1.64$$

So a 7.5-year-old man has about **64% higher** estimated hazard of CHD, all else equal. The operation is **exponentiation** ($HR^{0.75}$), not multiplication.

3.4 Comparing two coefficients: variance of a difference

Exam reference: Exercise 2.4 (2 points) — test whether the hazard of incident CHD differs between heavy and light smokers; compute the z -statistic and two-sided p -value.

⚠ Common mistake

This mistake was the **single most consequential Part 2 error**. To test $H_0 : \beta_H = \beta_L$, the standard error of $\hat{\beta}_H - \hat{\beta}_L$ was computed incorrectly in several ways:

- Using a **single covariance entry** $\text{Cov}(\hat{\beta}_L, \hat{\beta}_H)$ as the variance.
- **Adding** 2 Cov instead of **subtracting** it.
- Plugging in the **hazard ratios** (2.18, 1.43) instead of the **coefficients** (0.78, 0.36) in the numerator.
- Guessing a value for z because the variance-of-a-difference formula was not on the formula sheet.

Solution

Solution. Work on the $\log(HR)$ (coefficient) scale, where the estimates are approximately normal. The point estimate of the difference is

$$\hat{\Delta} = \hat{\beta}_H - \hat{\beta}_L = 0.78 - 0.36 = 0.42$$

The variance of a **difference** of two estimates uses **all three** relevant entries of the covariance matrix:

$$\text{Var}(\hat{\beta}_H - \hat{\beta}_L) = \text{Var}(\hat{\beta}_H) + \text{Var}(\hat{\beta}_L) - 2 \text{Cov}(\hat{\beta}_H, \hat{\beta}_L)$$

The -2 Cov term is essential — its sign is **minus** for a difference. Plugging in:

```
var_H <- 0.021905
var_L <- 0.024840
cov_HL <- 0.010504

diff <- 0.78 - 0.36
var_diff <- var_H + var_L - 2 * cov_HL
se_diff <- sqrt(var_diff)

z <- diff / se_diff
p_value <- 2 * pnorm(-abs(z))

c(diff = diff, se = se_diff, z = z, p_value = p_value) |> round(4)
#>      diff      se      z p_value
#> 0.4200 0.1604 2.6180 0.0088
```

So $z = 0.42/0.160 = 2.62$, two-sided $p = 0.0088$. We **reject** H_0 : the hazard of CHD differs significantly between heavy and light smokers, holding the other covariates constant.

Note

Compare with the **naive** standard error that ignores the covariance, $\sqrt{\text{Var}(\hat{\beta}_H) + \text{Var}(\hat{\beta}_L)} = 0.216$. Because $\hat{\beta}_H$ and $\hat{\beta}_L$ are *positively* correlated, subtracting 2 Cov makes the correct SE (0.160) **smaller**, which sharpens the test. Using a single covariance cell as the variance gave a wildly wrong SE and an absurd z .

3.5 A confidence interval for a hazard ratio: build it on the log scale, then exponentiate

Exam reference: Exercise 2.5 (2 points) — a 95% confidence interval for the hazard ratio comparing heavy to light smokers.

⚠ Common mistake

- Forming the interval directly on the HR scale, $\widehat{HR} \pm 1.96 \cdot \text{SE}$ (e.g. 1.52 ± 0.31), instead of on the log scale.
- Building the log-scale interval correctly but **never exponentiating**, then reporting a log-scale interval as if it were the HR (sometimes even declaring “significant” while the reported interval contained 1).
- Reusing the wrong SE from the Wald test in Section 3.4.

Solution

Solution. A hazard ratio is positive and its sampling distribution is skewed, so the symmetric normal approximation is applied to $\log(HR)$, and the endpoints are **then** exponentiated. For the heavy-vs-light comparison ($\widehat{\Delta} = 0.42$, $\text{SE} = 0.160$):

$$\text{log-scale CI: } 0.42 \pm 1.96 \times 0.160 = (0.106, 0.734)$$

$$\text{HR CI: } (e^{0.106}, e^{0.734}) = (1.11, 2.08)$$

```
est <- 0.42
se <- sqrt(0.021905 + 0.024840 - 2 * 0.010504)
exp(est + c(-1, 1) * 1.96 * se) |> round(2)
#> [1] 1.11 2.08
```

The point estimate is $\widehat{HR} = e^{0.42} = 1.52$, and the 95% CI (1.11, 2.08) **excludes 1**, consistent with the significant Wald test in Section 3.4. Always exponentiate the endpoints, and check that your CI and your hypothesis test agree.

4 Summary of the most common errors

#	Exam Q	Topic	The fix
1	1.1	Risk set after censoring (Section 2.1)	Remove everyone who has an event <i>or</i> is censored from later risk sets.
2	1.1	Nelson-Aalen (Section 2.2)	Sum hazards d_j/n_j , then $\hat{S}_{\text{NA}} = \exp\{-\hat{\Lambda}\}$.
3	1.3	Reading $\hat{S}(t)$ (Section 2.3)	Step function: use the most recent event time $\leq t$; give the number.
4	1.4	Median survival (Section 2.4)	First t with $\hat{S}(t) \leq 0.5$ — not a mean, not an exponential fit.
5	1.5	Role of censoring (Section 2.5)	Denominator until censoring; never the event numerator.
6	2.1	Writing the model (Section 3.1)	Name and apply every assumption; baseline hazard stays unspecified.
7	2.2	Interpreting an HR (Section 3.2)	Magnitude + reference + adjustment + significance; HR $\neq$ risk/odds.
8	2.3	HR for a non-unit change (Section 3.3)	$HR^c = \exp\{c\beta\}$ — exponentiate, don't multiply.
9	2.4	Comparing two coefficients (Section 3.4)	$\text{Var}(\hat{\beta}_H - \hat{\beta}_L) = \text{Var}_H + \text{Var}_L - 2 \text{Cov}$.

#	Exam Q	Topic	The fix
10	2.5	CI for an HR (Section 3.5)	Build on the log scale, then exponentiate the endpoints.