

# KCOR Supplementary Information (SI)

This document provides supplementary material supporting the KCOR methodology described in the main manuscript, including extended derivations, simulation studies, robustness analyses, and additional empirical results.

## S1. Overview

This SI is organized as follows:

- **S1:** Overview
- **S2:** Extended diagnostics and failure modes
- **S3:** Positive controls (injected harm/benefit)
- **S4:** Control-test specifications and simulation parameters
- **S5:** Additional figures and diagnostics
- **S6:** Extended Czech empirical application / illustrative registry analysis

## S2. Extended diagnostics and failure modes

This section describes the **observable diagnostics and failure modes** associated with the KCOR working assumptions and the corresponding diagnostics and identifiability criteria. No additional assumptions are introduced here. KCOR is designed to **fail transparently rather than silently**: when an assumption is violated, the resulting lack of identifiability or model stress manifests through explicit diagnostic signals rather than spurious estimates.

The KCOR framework separates **working assumptions**, **empirical diagnostics**, and **identifiability criteria**; these are summarized below in Tables S1–S3.

Table S1: KCOR working assumptions.

Assumption	Description	Role in KCOR
A1 Cohort stability	Cohorts are fixed at enrollment with no post-enrollment switching or informative censoring.	Ensures cumulative hazards are comparable over follow-up
A2 Shared external hazard environment	Cohorts experience the same background hazard over the comparison window.	Prevents confounding by cohort-specific shocks
A3 Time-invariant latent frailty	Selection operates through time-invariant unobserved heterogeneity inducing depletion.	Enables geometric normalization of curvature
A4 Adequacy of gamma frailty	Gamma frailty provides a reasonable approximation to observed depletion geometry.	Allows tractable inversion and normalization
A5 Quiet-window validity	A prespecified window exists in which depletion dominates other curvature sources.	Permits identification of frailty parameters

Table S2: Empirical diagnostics associated with KCOR assumptions.

Diagnostic	Description	Observable signal
Skip-week sensitivity	Exclude early post-enrollment weeks subject to dynamic selection.	Stable fitted frailty under varying skip weeks
Post-normalization linearity	Assess curvature removal in cumulative-hazard space.	Approximate linearity after normalization

Diagnostic	Description	Observable signal
KCOR(t) stability	Inspect KCOR trajectories following anchoring.	Stabilization rather than drift
Quiet-window perturbation	Shift quiet-window boundaries by $\pm$ several weeks.	Parameter and trajectory stability
Residual structure	Examine residuals in cumulative-hazard space.	No systematic curvature or autocorrelation

Table S3: Identifiability criteria governing KCOR interpretation.

Criterion	Condition	Consequence if violated
I1 Diagnostic sufficiency	All required diagnostics pass.	KCOR interpretable
I2 Window alignment	Follow-up overlaps the hypothesized effect window.	Out-of-window effects not recoverable
I3 Stability under perturbation	Estimates robust to tuning of windows and skips.	Interpretation limited
I4 Anchoring validity	Quiet window exhibits post-normalization linearity.	Anchoring invalid
I5 Conservative failure rule	Any failure $\rightarrow$ diagnostics indicate non-identifiability.	Analysis treated as not identified; results not reported

When diagnostics indicate non-identifiability, the analysis is treated as not identified and results are not reported; this does not invalidate the KCOR estimator itself.

### S3. Positive controls

#### S3.1 Construction of injected effects

The effect window is a simulation construct used solely for positive-control validation and does not represent a real-world intervention period or biological effect window.

Positive controls are constructed by starting from a negative-control dataset and injecting a known effect into the data-generating process for one cohort, for example by multiplying the *baseline* hazard by a constant factor  $r$  over a prespecified interval:

$$h_{0,\text{treated}}(t) = r \cdot h_{0,\text{control}}(t) \quad \text{for } t \in [t_1, t_2], \quad (1)$$

with  $r > 1$  for harm and  $0 < r < 1$  for benefit.

After gamma-frailty normalization (inversion), KCOR should deviate from 1 in the correct direction and with magnitude consistent with the injected effect (up to discretization and sampling noise). Figure S1 and Table ?? confirm this behavior.

### S4. Control-test specifications and simulation parameters

#### S4.0 Summary tables for control-test and simulation parameters

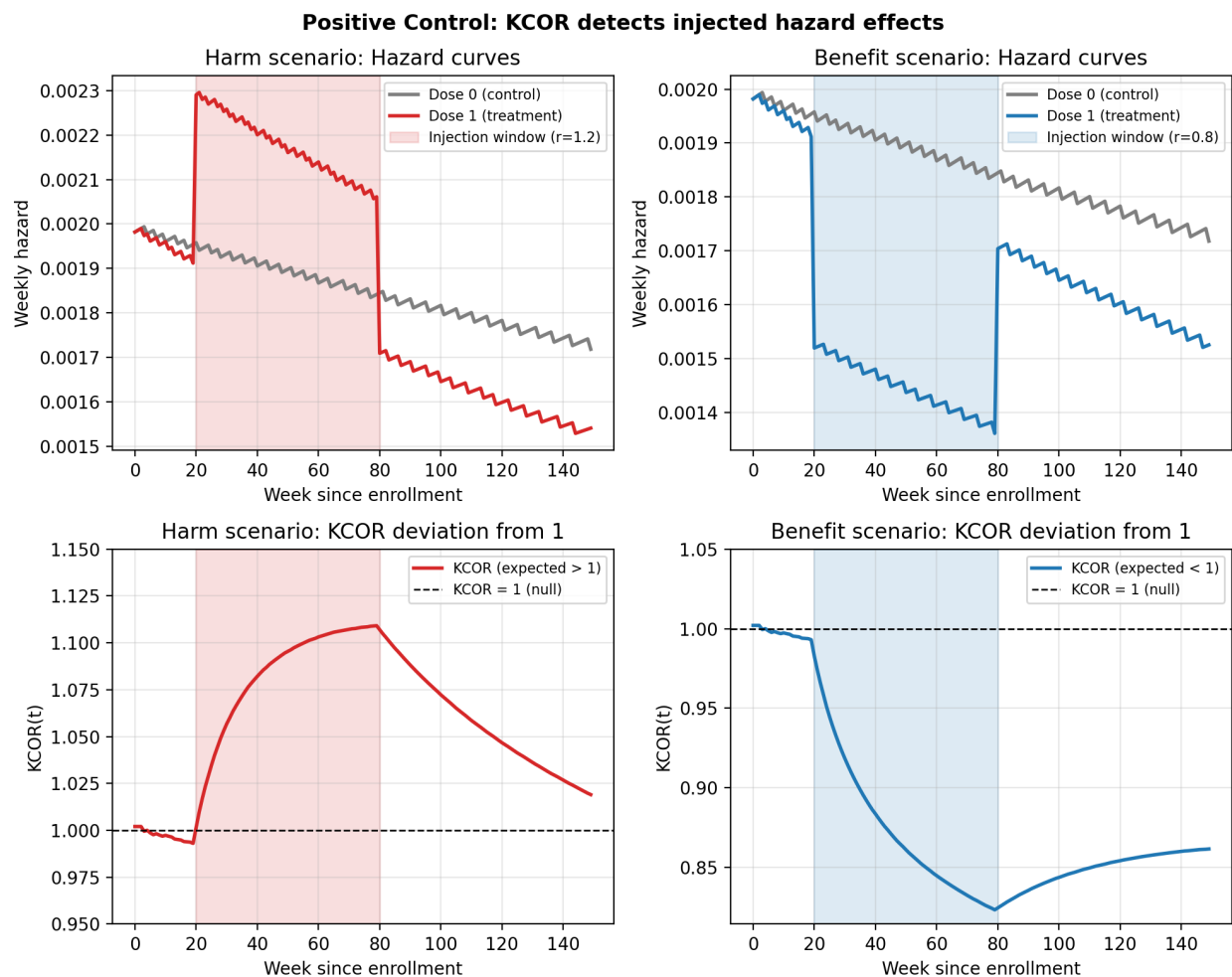


Figure S1: Positive control validation: KCOR correctly detects injected effects. Left panels show harm scenario ( $r=1.2$ ), right panels show benefit scenario ( $r=0.8$ ). Top row displays cohort hazard curves with effect window shaded. Bottom row shows  $KCOR(t)$  deviating from 1.0 in the expected direction during the effect window. Uncertainty bands (95% bootstrap intervals; aggregated cohort-time resampling) are shown. X-axis units are weeks since enrollment.

Table S4: Summary of control-test and simulation parameters referenced in Sections S4.2–S4.6. Numeric values are fixed unless otherwise noted; ranges indicate sensitivity grids.

Section	Item	Parameter	Value	Notes
S4.2.1	Synthetic negative control	Data source	example/Frail_cohor	Pathological frailty mixture
S4.2.1	Synthetic negative control	Generation script	code/generate_pathological_neg_control_figs.py	
S4.2.1	Synthetic negative control	Cohort A weights	[0.20, 0.20, 0.20, 0.20, 0.20]	5 frailty groups
S4.2.1	Synthetic negative control	Cohort B weights	[0.30, 0.20, 0.20, 0.20, 0.10]	Shifted mixture
S4.2.1	Synthetic negative control	Frailty values	[1, 2, 4, 6, 10]	Relative multipliers
S4.2.1	Synthetic negative control	Base weekly probability	0.01	
S4.2.1	Synthetic negative control	Weekly log-slope	0.0	Constant baseline during quiet periods
S4.2.1	Synthetic negative control	Skip weeks	2	
S4.2.1	Synthetic negative control	Normalization weeks	4	
S4.2.1	Synthetic negative control	Time horizon	250 weeks	
S4.2.2	Empirical negative control	Data source	Czech admin registry data (KCOR_CM)	Aggregated cohorts
S4.2.2	Empirical negative control	Generation script	test/negative_control/code/generate_negative_con	
S4.2.2	Empirical negative control	Construction	Age strata remapped to pseudo-doses	True null preserved
S4.2.2	Empirical negative control	Age mapping	Dose 0→YoB {1930,1935}; Dose 1→{1940,1945}; Dose 2→{1950,1955}	
S4.2.2	Empirical negative control	Output YoB	1950 (unvax) or 1940 (vax)	
S4.2.2	Empirical negative control	Sheets processed	2021_24, 2022_06	
S4.3	Positive control	Generation script	test/positive_control/code/generate_positive_con	
S4.3	Positive control	Initial cohort size	100,000 per cohort	
S4.3	Positive control	Baseline hazard	0.002 per week	Constant
S4.3	Positive control	Frailty variance	$\theta_0=0.5$ (control), $\theta_1=1.0$ (treatment)	
S4.3	Positive control	Effect window	weeks 20–80	
S4.3	Positive control	Hazard multipliers	$r=1.2$ (harm); $r=0.8$ (benefit)	
S4.3	Positive control	Random seed	42	
S4.3	Positive control	Enrollment date	2021-06-14 (ISO week 2021_24)	
S4.4	Sensitivity analysis	Baseline weeks	[2,3,4,5,6,7,8]	Varied

Section	Item	Parameter	Value	Notes
S4.4	Sensitivity analysis	Quiet-start offsets	[-12,-8,-4,0,+4,+8,+12]	Weeks from 2023-01
S4.4	Sensitivity analysis	Quiet-window end	2023-52	Fixed
S4.4	Sensitivity analysis	Dose pairs	1 vs 0; 2 vs 0; 2 vs 1	
S4.4	Sensitivity analysis	Cohorts	2021_24	
S4.5	Tail-sampling (adversarial)	Generation script	test/sim_grid/code/generate_tail_sampling_sim.py	
S4.5	Tail-sampling (adversarial)	Base frailty distribution	Log-normal, mean 1, variance 0.5	
S4.5	Tail-sampling (adversarial)	Mid-quantile cohort	25th–75th percentile	Renormalized to mean 1
S4.5	Tail-sampling (adversarial)	Tail-mixture cohort	[0–15th] + [85th–100th], equal weights	Weights yield mean 1
S4.5	Tail-sampling (adversarial)	Baseline hazard	0.002 per week	Constant
S4.5	Tail-sampling (adversarial)	Positive-control multiplier	r=1.2 (harm) or r=0.8 (benefit)	
S4.5	Tail-sampling (adversarial)	Effect window	weeks 20–80	
S4.5	Tail-sampling (adversarial)	Random seed	42	
S4.6	Joint frailty + effect	Generation script	test/sim_grid/code/generate_s7_sim.py	
S4.6	Joint frailty + effect	Time horizon	260 weeks	
S4.6	Joint frailty + effect	Cohort size	2,000,000 per cohort	
S4.6	Joint frailty + effect	Frailty distribution	Gamma, mean 1	$\theta_0=0.3$ , $\theta_1=0.8$
S4.6	Joint frailty + effect	Baseline hazard	0.002 per week	Constant
S4.6	Joint frailty + effect	Quiet window	weeks 80–140	Prespecified for frailty estimation
S4.6	Joint frailty + effect	Effect windows	weeks 10–25 (early), 150–190 (late)	Overlap variant: 70–95
S4.6	Joint frailty + effect	Effect shapes	step, ramp, bump	
S4.6	Joint frailty + effect	Effect multiplier	r=1.2 (harm); r=0.8 (benefit)	Applied to treated cohort
S4.6	Joint frailty + effect	Skip weeks	2	
S4.6	Joint frailty + effect	Random seed	42	
S4.6	Joint frailty + effect	Enrollment date	2021-06-14	ISO week 2021_24

#### S4.1 Reference implementation and default operational settings

Table S5: Reference implementation and default operational settings.

Component	Setting	Default value	Notes
Cohort construction	Cohort indexing	Enrollment period × YearOfBirth group × Dose; plus all-ages cohort (YearOfBirth = -2)	Implementation detail
Quiet-period selection	Quiet window	ISO weeks 2023-01 through 2023-52	Calendar year 2023

Component	Setting	Default value	Notes
Early-period stabilization (dynamic HVE)	SKIP_WEEKS	2	Weeks $t < \text{SKIP\_WEEKS}$ are excluded from hazard accumulation (set $\Delta H_d(t) = 0$ for those weeks).
Frailty estimation	Fit method	Nonlinear least squares in cumulative-hazard space	Constraints: $k_d > 0$ , $\theta_d \geq 0$

## S4.2 Negative controls

Negative controls are used to evaluate the behavior of KCOR under settings where the true effect is known to be null, while allowing substantial heterogeneity in baseline risk and selection-induced depletion. Two complementary classes of negative controls are considered: (i) fully synthetic simulations that induce strong depletion curvature through frailty-mixture imbalance, and (ii) empirical registry-based constructions that preserve a true null by repurposing age strata as pseudo-exposures without selective sampling. Together, these controls assess whether KCOR remains stable in the presence of non-proportional hazards arising from selection rather than treatment.

### S4.2.1 Synthetic negative control: gamma-frailty null

The synthetic negative control (Figure S3) is a fully specified simulation designed to induce **strong selection-induced depletion curvature under a true null effect** by altering only the cohort frailty-mixture weights. KCOR is expected to remain near 1 after depletion normalization despite large differences in cohort-level hazard curvature.

Parameter values and scripts are summarized in Table S4.

Both cohorts share identical per-frailty-group death probabilities; only the mixture weights differ. This induces different cohort-level curvature under the null.

Figure S2 provides a compact summary of KCOR bias as a function of frailty variance  $\theta$  under the same synthetic-null grid used in Table ??.

### S4.2.2 Empirical negative control: age-shift construction

The empirical negative control (Figures ?? and S8) repurposes registry cohorts to create a **true null comparison** while inducing large baseline hazard differences via 10–20 year age shifts. Because these are full-population strata rather than selectively sampled subcohorts, selection-induced depletion is minimal and no gamma-frailty normalization is applied.

Parameter values and scripts are summarized in Table S4.

This construction ensures that dose comparisons are within the same underlying vaccination category, preserving a true null while inducing 10–20 year age differences.

This contrasts with the synthetic negative control (Section S4.2.1), where strong, deliberately induced frailty heterogeneity requires gamma-frailty normalization to recover the null.

## S4.3 Positive control: injected effect

Positive controls are used to verify that KCOR responds appropriately when a true effect is present. Starting from a negative-control simulation with no treatment effect, a known multiplicative hazard shift is injected into one cohort over a prespecified time window. This construction allows direct assessment of whether KCOR detects both the direction and timing of the injected effect while remaining stable outside the effect window.

Parameter values and scripts are summarized in Table S4.

The injection multiplies the treatment cohort’s baseline hazard by factor  $r$  during the effect window, while leaving the control cohort unchanged.

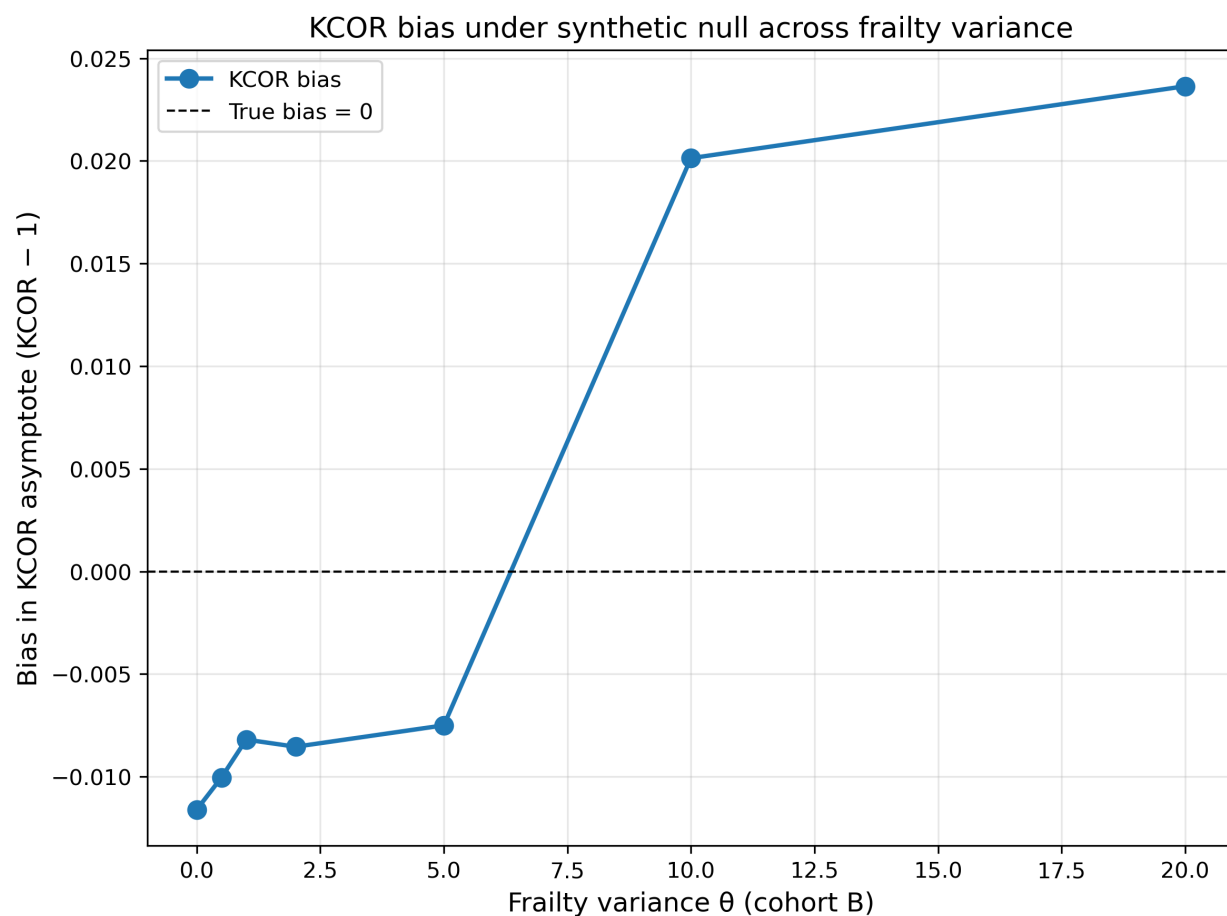


Figure S2: Simulated-null summary: KCOR bias as a function of frailty variance  $\theta$ .

Bias is defined as  $\text{KCOR}_{\text{asymptote}} - 1$  at the end of follow-up in the synthetic-null grid (no treatment effect), reflecting cumulative deviation under the null rather than instantaneous hazard bias. Points show single-run estimates from the grid; no error bars are shown.

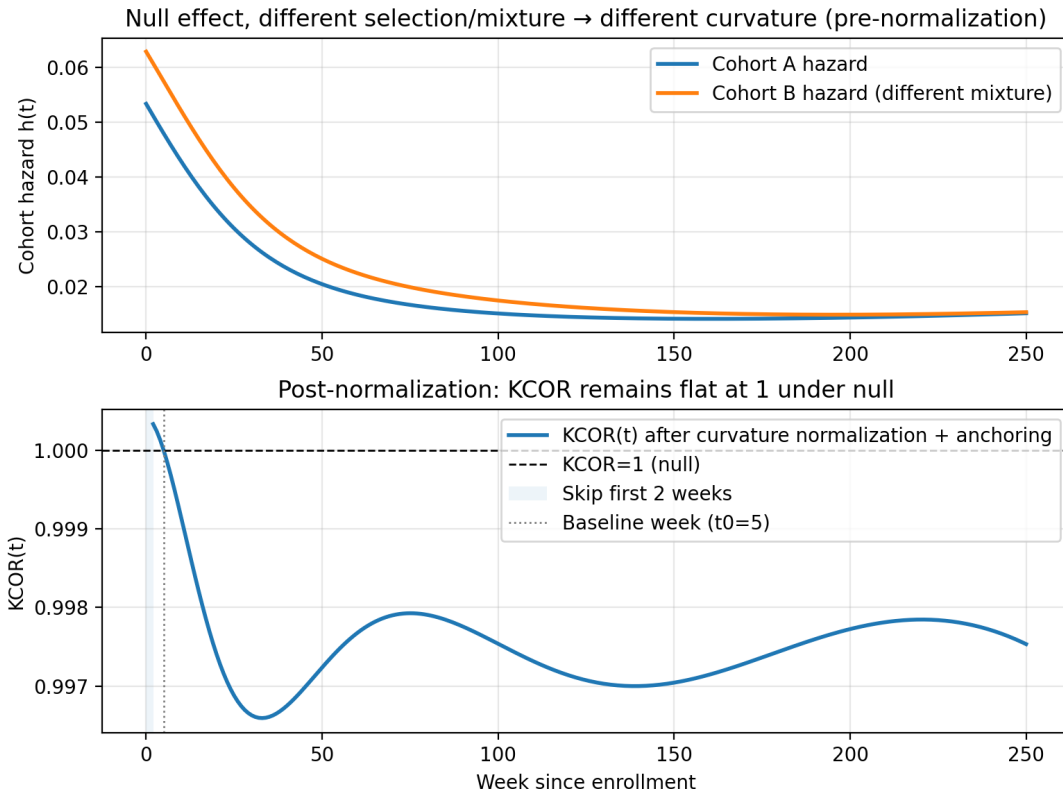


Figure S3: Synthetic negative control under strong selection (different curvature) but no effect:  $KCOR(t)$  remains flat at 1. Top panel shows cohort hazards with different frailty-mixture weights inducing different curvature. Bottom panel shows  $KCOR(t)$  remaining near 1.0 after normalization, demonstrating successful depletion-neutralization under the null. Uncertainty bands (95% bootstrap intervals; aggregated cohort-time resampling) are shown.



#### S4.4 Sensitivity analysis parameters

Sensitivity analyses evaluate the robustness of KCOR estimates to reasonable variation in analysis choices that do not alter the underlying data-generating process. Baseline-window length and quiet-window placement are perturbed over a prespecified range while holding all other parameters fixed. These analyses assess whether KCOR behavior is stable to tuning choices that primarily affect normalization rather than cohort composition.

Parameter values and scripts are summarized in Table S4.

Output grids show KCOR( $t$ ) values for each parameter combination.

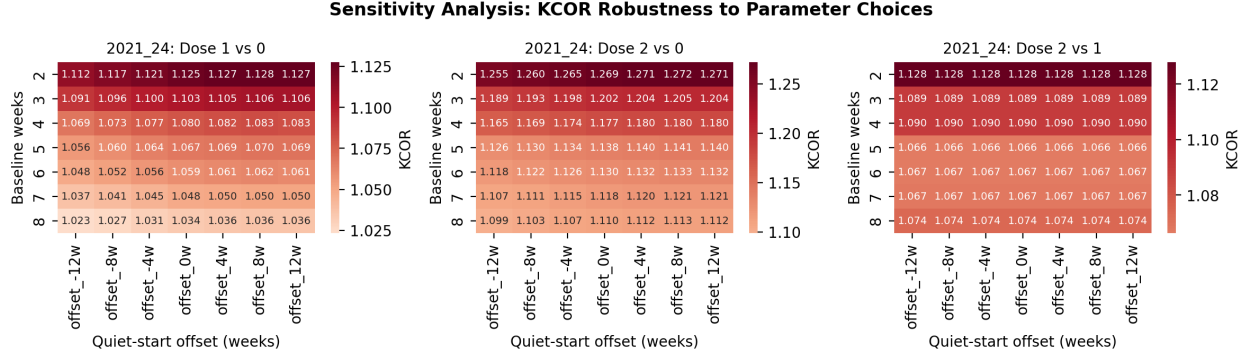


Figure S4: Sensitivity analysis summary showing KCOR( $t$ ) values across parameter grid. Heatmaps display KCOR( $t$ ) estimates for different combinations of baseline weeks (rows) and quiet-window start offsets (columns). Across all comparisons, KCOR( $t$ ) varies smoothly and modestly across a wide range of quiet-start offsets and baseline window lengths, with no qualitative changes in sign or magnitude, indicating robustness to reasonable parameter choices. All panels use a unified color scale centered at 1.0 to enable direct visual comparison across dose comparisons.

#### S4.5 Tail-sampling / bimodal selection (adversarial selection geometry)

This adversarial simulation evaluates KCOR under extreme but controlled violations of typical cohort-selection geometry. Two cohorts are constructed to share identical mean frailty while differing sharply in how risk is distributed, using mid-quantile sampling versus a low/high-tail mixture. This setting stress-tests whether depletion normalization remains effective when frailty heterogeneity is concentrated in the tails rather than smoothly distributed.

- **Mid-sampled cohort:** frailty restricted to central quantiles (e.g., 25th–75th percentile) and renormalized to mean 1.
- **Tail-sampled cohort:** mixture of low and high tails (e.g., 0–15th and 85th–100th percentiles) with mixture weights chosen to yield mean 1.

Parameter values and scripts are summarized in Table S4.

Both cohorts share the same baseline hazard  $h_0(t)$  and have no treatment effect (negative-control version). Positive-control versions are also generated by applying a known hazard multiplier in a prespecified window. The evaluation includes (i) KCOR drift, (ii) quiet-window fit RMSE, (iii) post-normalization linearity, and (iv) parameter stability under window perturbation.

#### S4.6 Joint frailty and treatment-effect simulation

This simulation evaluates KCOR under conditions in which **both selection-induced depletion (frailty heterogeneity)** and a **true treatment effect (harm or benefit)** are present simultaneously. The purpose is to assess whether KCOR can (i) correctly identify and neutralize frailty-driven curvature using a quiet period and (ii) detect a true treatment effect outside that period without confounding the two mechanisms.

This joint simulation combines the selection-induced depletion mechanisms examined in Sections S4.2 and S4.5 with the injected-effect framework of Section S4.3.

Parameter values and scripts for this joint simulation are summarized in Table S4.

## Design

Two fixed cohorts are generated with identical baseline hazards but differing frailty variance. Individual hazards are multiplicatively scaled by a latent frailty term drawn from a gamma distribution with unit mean and cohort-specific variance. A treatment effect is then injected over a prespecified time window that does not overlap the quiet period used for frailty estimation.

Formally, individual hazards are generated as

$$h_i(t) = z_i \cdot h_0(t) \cdot r(t). \quad (2)$$

where  $z_i$  is individual frailty,  $h_0(t)$  is a shared baseline hazard, and  $r(t)$  is a time-localized multiplicative treatment effect applied to one cohort only.

## S5. Additional figures and diagnostics

This section provides diagnostic outputs and evaluation criteria for the simulations and control-test specifications defined in Section S4.

### S5.1 Fit diagnostics

For each cohort  $d$ , the gamma-frailty fit produces diagnostic outputs including:

- **RMSE in  $H$ -space:** Root mean squared error between observed and model-predicted cumulative hazards over the quiet window. Values  $< 0.01$  indicate excellent fit; values  $> 0.05$  may warrant investigation.
- **Fitted parameters:** baseline hazard level and frailty variance. Very small frailty variance ( $< 0.01$ ) indicates minimal detected depletion; very large values ( $> 5$ ) may indicate model stress.
- **Number of fit points:**  $n_{\text{obs}}$  observations in quiet window. Larger  $n_{\text{obs}}$  provides more stable estimates.

When uncertainty bands or bootstrap intervals are reported in this supplement, they are computed using an aggregated-data bootstrap at the cohort–time level (resampling event counts and risk-set sizes within time bins/strata), not by resampling individuals.

Example diagnostic output from the reference implementation:

```
KCOR_FIT,EnrollmentDate=2021_24,YoB=1950,Dose=0,
  k_hat=4.29e-03,theta_hat=8.02e-01,
  RMSE_Hobs=3.37e-03,n_obs=97,success=1
```

### S5.2 Residual analysis

Fit residuals should be examined for. Define residuals:

$$r_d(t) = H_{\text{obs},d}(t) - H_d^{\text{model}}(t; \hat{k}_d, \hat{\theta}_d).$$

- **Systematic patterns:** Residuals should be approximately random around zero. Systematic curvature in residuals suggests model inadequacy.
- **Outliers:** Individual weeks with large residuals may indicate data quality issues or external shocks.
- **Autocorrelation:** Strong autocorrelation in residuals suggests the model is missing time-varying structure.

### S5.3 Parameter stability checks

Robustness of fitted parameters is assessed by:

- **Quiet-window perturbation:** Shift the quiet-window start/end by  $\pm 4$  weeks and re-fit. Stable parameters should vary by  $< 10\%$ .
- **Skip-weeks sensitivity:** Vary SKIP\_WEEKS from 0 to 8 and verify KCOR(t) trajectories remain qualitatively similar.
- **Baseline-shape alternatives:** Compare the default constant baseline over the fit window to mild linear trends and verify normalization is not sensitive to this choice.

### S5.4 Quiet-window overlay plots

Overlaying the prespecified quiet window on hazard and cumulative-hazard time series plots provides a visual diagnostic of window placement relative to mortality dynamics. The fit window should:

- Avoid major epidemic waves or external mortality shocks
- Contain sufficient event counts for stable estimation
- Span a time range where baseline mortality is approximately stationary

Visual inspection of quiet-window placement relative to mortality dynamics is an essential diagnostic step.

### S5.5 Robustness to age stratification

This subsection illustrates robustness of  $\text{KCOR}(t)$  to narrow age stratification by repeating the same fixed-cohort comparison in three single birth-year cohorts spanning advanced ages (1930, 1940, 1950). Across these strata, the trajectories remain qualitatively stable after depletion normalization, supporting the claim that the observed behavior is not an artifact of age aggregation.

#### **Additional empirical negative-control variant (20-year age shift).**

For completeness, the more extreme 20-year age-shift negative control referenced in the main text is included:

## S6. Extended Czech empirical application

### S6.1 Empirical application with diagnostic validation: Czech Republic national registry mortality data

The Czech results do not validate KCOR; they represent an application that satisfies all pre-specified diagnostic criteria. Substantive implications follow only if the identification assumptions hold. Throughout this subsection, observed divergences are interpreted strictly as properties of the estimator under real-world selection, not as intervention effects.

Unless otherwise noted, KCOR curves in the Czech analyses are shown anchored at  $t_0 = 4$  weeks for interpretability.

#### **S6.1.1 Illustrative empirical context: COVID-19 mortality data**

The COVID-19 vaccination period provides a natural empirical regime characterized by strong selection heterogeneity and non-proportional hazards, making it a useful illustration for the KCOR framework. During this period, vaccine uptake was voluntary, rapidly time-varying, and correlated with baseline health status, creating clear examples of selection-induced non-proportional hazards. The Czech Republic national mortality registry data exemplify this regime: voluntary uptake led to asymmetric selection at enrollment, with vaccinated cohorts exhibiting minimal frailty heterogeneity while unvaccinated cohorts retained substantial heterogeneity. This asymmetric pattern reflects the healthy vaccinee effect operating through selective uptake rather than treatment. KCOR normalization removes this selection-induced curvature, enabling interpretable cumulative comparisons. While these examples illustrate KCOR's application, the method is general and applies to any retrospective cohort comparison where selection induces differential depletion dynamics.

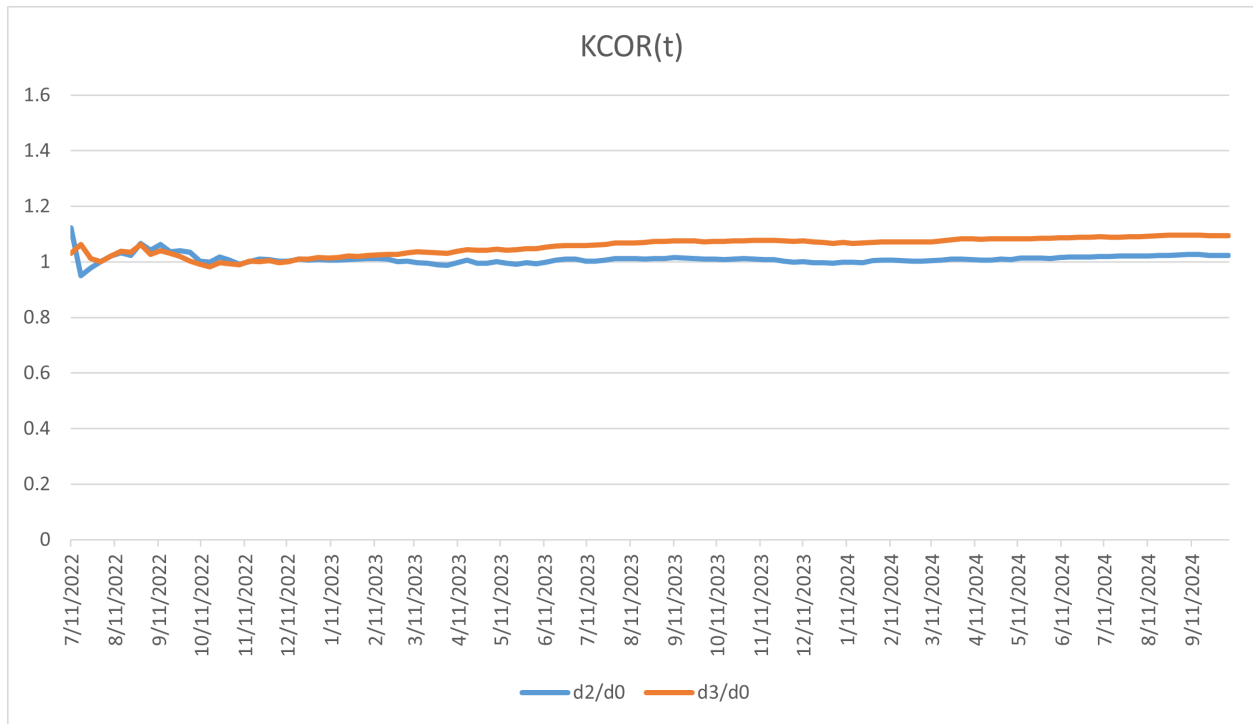


Figure S5: Birth-year cohort 1930: KCOR(t) trajectories comparing dose 2 and dose 3 to dose 0 for cohorts enrolled in ISO week 2022-26 and evaluated over calendar year 2023. KCOR curves are anchored at  $t_0 = 4$  weeks (i.e., plotted as  $KCOR(t; t_0)$ ). This figure is presented as an illustrative application demonstrating estimator behavior on registry data and does not support causal inference. X-axis units are weeks since enrollment.

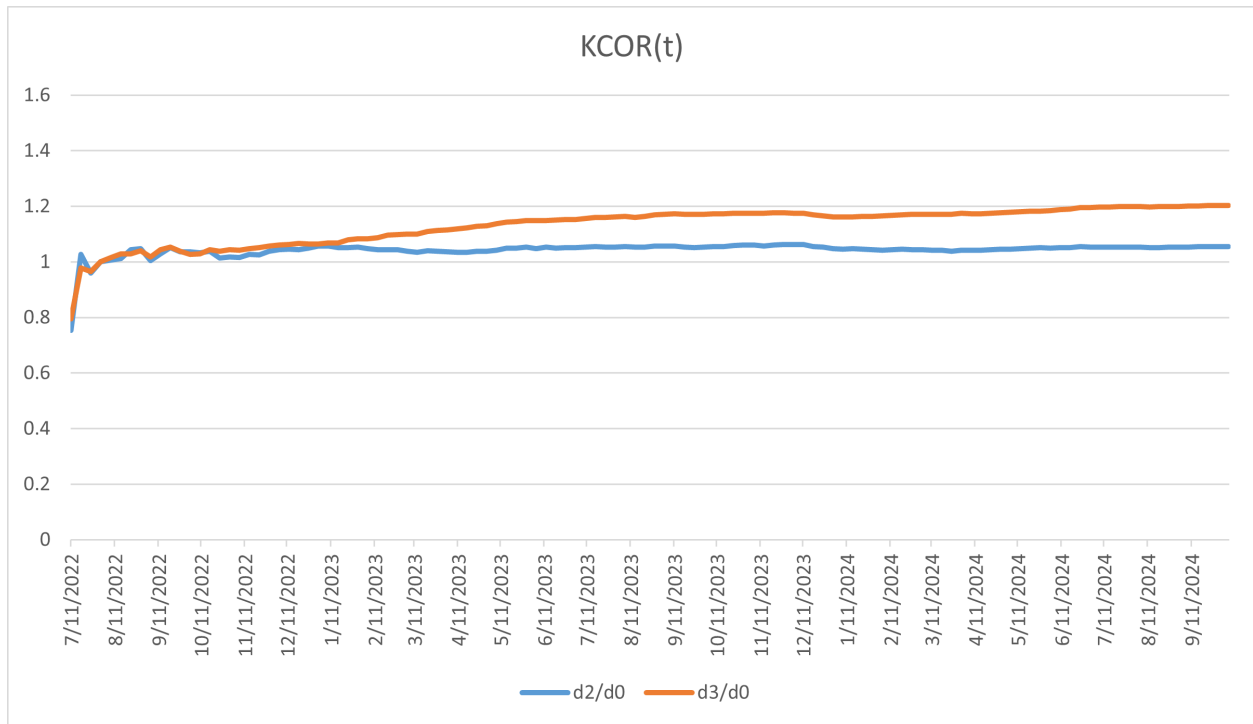


Figure S6: Birth-year cohort 1940: KCOR(t) trajectories comparing dose 2 and dose 3 to dose 0 for cohorts enrolled in ISO week 2022-26 and evaluated over calendar year 2023. KCOR curves are anchored at  $t_0 = 4$  weeks (i.e., plotted as  $\text{KCOR}(t; t_0)$ ). This figure is presented as an illustrative application demonstrating estimator behavior on registry data and does not support causal inference. X-axis units are weeks since enrollment.

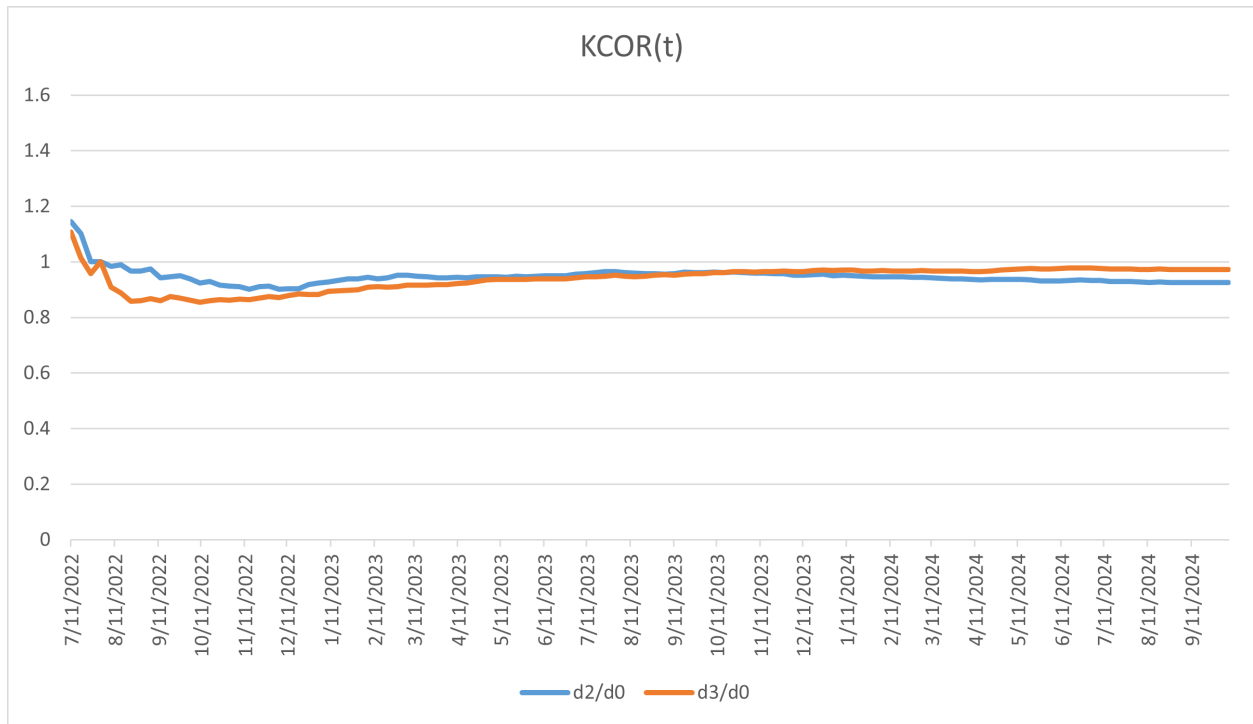


Figure S7: Birth-year cohort 1950: KCOR(t) trajectories comparing dose 2 and dose 3 to dose 0 for cohorts enrolled in ISO week 2022-26 and evaluated over calendar year 2023. KCOR curves are anchored at  $t_0 = 4$  weeks (i.e., plotted as  $KCOR(t; t_0)$ ). This figure is presented as an illustrative application demonstrating estimator behavior on registry data and does not support causal inference. X-axis units are weeks since enrollment.

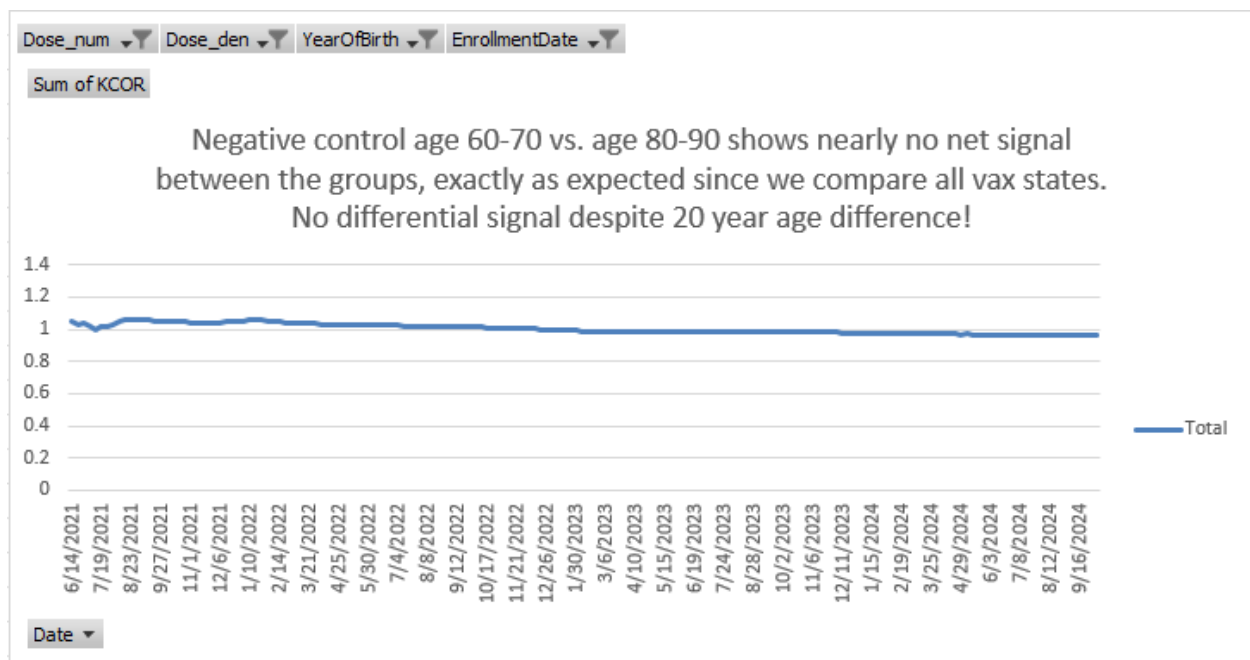


Figure S8: Empirical negative control with approximately 20-year age difference between cohorts. Even under extreme composition differences,  $KCOR(t)$  exhibits no systematic drift, consistent with robustness to selection-induced curvature.  $KCOR$  curves are anchored at  $t_0 = 4$  weeks (i.e., plotted as  $KCOR(t; t_0)$ ). Uncertainty bands (95% bootstrap intervals) are shown. Data source: Czech Republic mortality and vaccination dataset processed into  $KCOR\_CMR$  aggregated format (negative-control construction; see Supplementary Information, SI). X-axis units are weeks since enrollment.

### S6.1.2 Frailty normalization behavior under empirical validation

Across examined age strata in the Czech Republic mortality dataset, fitted frailty parameters exhibit a pronounced asymmetry across cohorts. Some cohorts show negligible estimated frailty variance:

$$\hat{\theta}_d \approx 0. \quad (3)$$

while others exhibit substantial frailty-driven depletion. This pattern reflects differences in selection-induced hazard curvature at cohort entry rather than any prespecified cohort identity.

As a consequence, KCOR normalization leaves some cohorts' cumulative hazards nearly unchanged, while substantially increasing the depletion-neutralized baseline cumulative hazard for others. This behavior is consistent with curvature-driven normalization rather than cohort identity. This pattern is visible directly in depletion-neutralized versus observed cumulative hazard plots and is summarized quantitatively in the fitted-parameter logs (see `KCOR_summary.log`).

After frailty normalization, the depletion-neutralized baseline cumulative hazards are approximately linear in event time. Residual deviations from linearity reflect real time-varying risk—such as seasonality or epidemic waves—rather than selection-induced depletion. This linearization is a diagnostic consistent with successful removal of depletion-driven curvature under the working model; persistent nonlinearity or parameter instability indicates model stress or quiet-window contamination.

Table S6 summarizes these diagnostic checks across age strata.



Table S6: Diagnostic gate for Czech application: KCOR results reported only where diagnostics pass.

Age band (years)	Quiet window valid	Post-normalization linearity	Parameter stability	KCOR reported
40–49	Yes	Yes	Yes	Yes
50–59	Yes	Yes	Yes	Yes
60–69	Yes	Yes	Yes	Yes
70–79	Yes	Yes	Yes	Yes
80–89	Yes	Yes	Yes	Yes
90–99	Yes	Yes	Yes	Yes
All ages	Yes	Yes	Yes	Yes

All age strata in the Czech application satisfied the prespecified diagnostic criteria, permitting KCOR computation and reporting. KCOR results are not reported for any age stratum where diagnostics indicate non-identifiability.

**Interpretation:** In this application, unvaccinated cohorts exhibit frailty heterogeneity, while Dose 2 cohorts show near-zero estimated frailty across all age bands, consistent with selective uptake prior to follow-up:

$$\hat{\theta}_d > 0. \quad (4)$$

for Dose 0 cohorts and

$$\hat{\theta}_d \approx 0. \quad (5)$$

for Dose 2 cohorts. Estimated frailty heterogeneity can appear larger at younger ages because baseline hazards are low, so proportional differences across latent risk strata translate into visibly different short-term hazards before depletion compresses the risk distribution. At older ages, higher baseline hazard and stronger ongoing depletion can reduce the apparent dispersion of remaining risk, yielding smaller fitted  $\theta$  even if latent heterogeneity is not literally smaller. Frailty variance is largest at younger ages, where low baseline mortality amplifies the impact of heterogeneity on cumulative hazard curvature, and declines at older ages where mortality is compressed and survivors are more homogeneous. Because Table S7 demonstrates selection-induced heterogeneity, unadjusted cumulative outcome contrasts are expected to conflate depletion effects with any true treatment differences; see Table S8 for raw cumulative hazards reported as a pre-normalization diagnostic. KCOR normalization removes the depletion component, enabling interpretable comparison of the remaining differences.

These raw contrasts reflect both selection and depletion effects and are not interpreted causally.

Table S7: Estimated gamma-frailty variance (fitted frailty variance) by age band and vaccination status for Czech cohorts enrolled in 2021\_24.

Age band (years)	Fitted frailty variance (Dose 0)	Fitted frailty variance (Dose 2)
40–49	16.79	$2.66 \times 10^{-6}$
50–59	23.02	$1.87 \times 10^{-4}$
60–69	13.13	$7.01 \times 10^{-18}$
70–79	6.98	$3.46 \times 10^{-17}$
80–89	2.97	$2.03 \times 10^{-11}$
90–99	0.80	$8.66 \times 10^{-16}$
All ages (full population)	4.98	$1.02 \times 10^{-11}$

**Notes:** - The fitted frailty variance quantifies unobserved frailty heterogeneity and selection-induced depletion within cohorts. Near-zero values indicate effectively linear cumulative hazards over the quiet window and are typical of strongly pre-selected cohorts. - Each entry reports a single fitted gamma-frailty variance for the specified age band and vaccination status within the 2021\_24 enrollment cohort. - The “All ages (full population)” row corresponds to an independent fit over the full pooled age range, included as a global diagnostic. - Table S8 reports raw outcome contrasts for ages 40+ (YOB  $\leq$  1980) where event counts are stable.

**Diagnostic checks:** - **Dose ordering:** the fitted frailty variance is positive for Dose 0 and collapses toward zero for Dose 2 across all age strata, consistent with selective uptake. - **Magnitude separation:** Dose 2 estimates are effectively zero relative to Dose 0, indicating near-linear cumulative hazards rather than forced curvature. - **Age coherence:** the fitted frailty variance decreases at older ages as baseline mortality rises and survivor populations become more homogeneous; monotonicity is not imposed. - **Stability:** No sign reversals, boundary pathologies, or numerical instabilities are observed. - **Falsifiability:** Failure of any one of these checks would constitute evidence against model adequacy.

Table S8: Ratio of observed cumulative mortality hazards for unvaccinated (Dose 0) versus fully vaccinated (Dose 2) Czech cohorts enrolled in 2021\_24. (Note: the all-ages row reflects aggregation effects and is not directly comparable to age-stratified rows.)

Age band (years)	Dose 0 cumulative hazard	Dose 2 cumulative hazard	Ratio
40–49	0.005260	0.004117	1.2776
50–59	0.014969	0.009582	1.5622
60–69	0.045475	0.023136	1.9655
70–79	0.123097	0.057675	2.1343
80–89	0.307169	0.167345	1.8355
90–99	0.776341	0.517284	1.5008
All ages (full population)	0.023160	0.073323	0.3159

This table reports unadjusted cumulative hazards derived directly from the raw data, prior to any frailty normalization or depletion correction, and is shown to illustrate the magnitude and direction of selection-induced curvature addressed by KCOR.

Values reflect raw cumulative outcome differences prior to KCOR normalization and are not interpreted causally due to cohort non-exchangeability. Cumulative hazards were integrated from cohort enrollment through the end of available follow-up for the 2021\_24 enrollment window (through week 2024-16), identically for Dose 0 and Dose 2 cohorts.

### S6.1.3 Illustrative application to national registry mortality data

A brief illustrative application is included to demonstrate end-to-end KCOR behavior on real registry mortality data in a setting that minimizes timing-driven shocks and window-tuning sensitivity. Cohorts were enrolled in ISO week 2022-26, and evaluation was restricted to calendar year 2023, yielding a 26-week post-enrollment buffer before slope estimation and a prespecified full-year window for assessment. Frailty parameters were estimated using a prespecified epidemiologically quiet window (calendar year 2023) to minimize wave-related hazard variation. This example is intended to illustrate estimator behavior under real-world selection and heterogeneity and does not support causal inference.

Figure S9 shows  $KCOR(t)$  trajectories for dose 2 and dose 3 relative to dose 0 for an all-ages analysis. An all-ages analysis is presented as a high-heterogeneity stress test, since aggregation across age induces substantial baseline hazard and frailty variation.

## S7. Computational environment and runtime notes

**Environment.** Python 3.11; key dependencies include numpy, scipy, pandas, and lifelines (for Cox-model comparisons), with plotting via matplotlib.

**Compute requirements.** The full simulation grid reproduces in approximately 1 hour 26 minutes on a 20-core CPU with 128 GB RAM; smaller subsets reproduce in minutes.

**Reproduction.** Running `make paper` (or the repository’s top-level build command) regenerates all artifacts from a clean checkout.

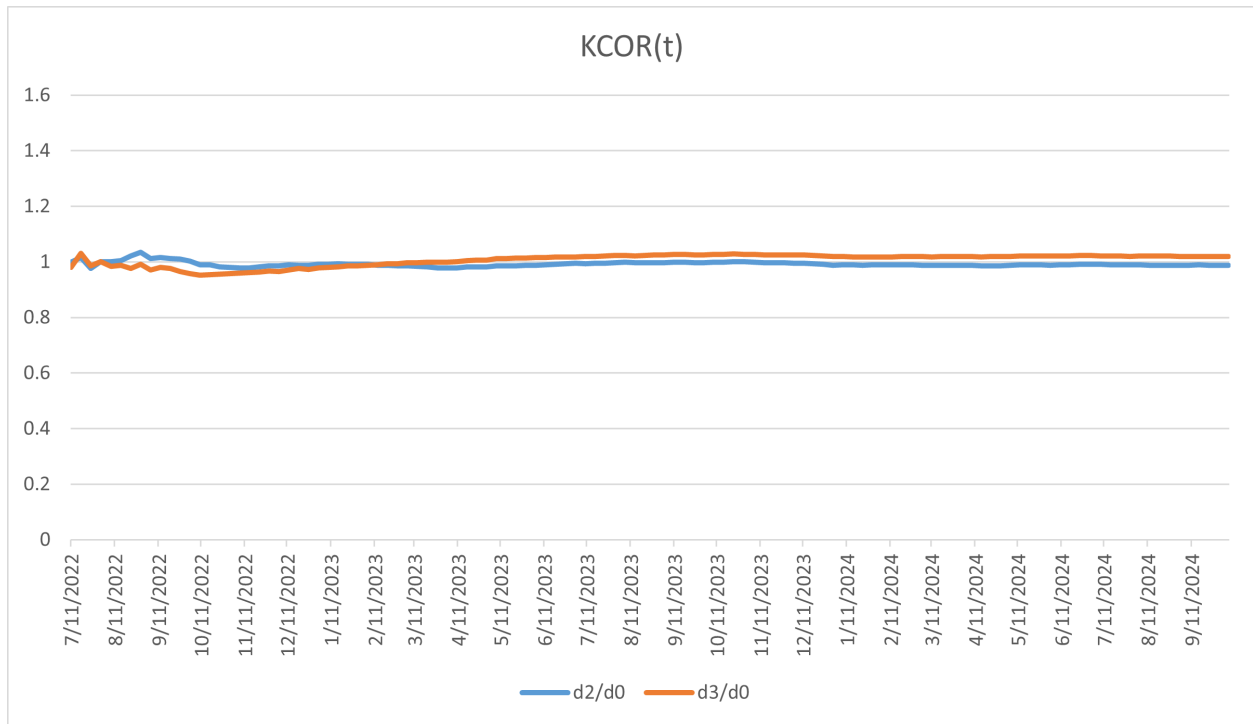


Figure S9: All-ages stress test:  $KCOR(t)$  trajectories comparing dose 2 and dose 3 to dose 0 for cohorts enrolled in ISO week 2022-26 and evaluated over calendar year 2023.  $KCOR$  curves are anchored at  $t_0 = 4$  weeks (i.e., plotted as  $KCOR(t; t_0)$ ). This figure is presented as an illustrative application demonstrating estimator behavior under extreme heterogeneity and does not support causal inference. X-axis units are weeks since enrollment.