**KCOR: A Hazard‑Based Method for Cumulative Outcomes Ratio Analysis**

*Methodological framework and diagnostics for observational mortality comparisons*

Steve Kirsch

Independent Researcher, Los Altos Hills, CA, USA

Correspondence: [stk@alum.mit.edu](mailto:stk@alum.mit.edu)

To do:

1. incorporate the visual description in the documentation section
2. incorporate the limitations in the README section

# Abstract

KCOR (Kirsch Cumulative Outcomes Ratio) is a hazard‑based method for comparing cumulative mortality between fixed cohorts in observational data. It enables us to use retrospective observational data to objectively answer important questions such as “Did this vaccine create a net harm or net benefit as of time t?” The procedure involves (i) choosing an enrollment date for defining fixed cohorts, (ii) slope‑normalizing each cohort to stable reference dates (ii) transforming the adjusted mortality rates to discrete hazards, and (iii) comparing cumulative hazards relative to a baseline‑normalized ratio over calendar time. This paper formalizes the algorithm, assumptions, diagnostics, and uncertainty propagation. The key assumption—that baseline mortality of real-world fixed cohorts is well‑approximated by an exponential trend—can be empirically validated on any dataset of interest, avoiding claims of universal applicability. KCOR is an important new tool because traditional epidemiology often struggles with net benefit assessments in retrospective data without randomization or the ability to perform accurate 1:1 matching. KCOR matches cohorts by their mortality and their slope which arguably provides more accurate matching than can be achieved through traditional methods, especially when vaccinated and unvaccinated cohorts are being compared. [are we allowed to reference the Qatar elife paper here?]

Keywords: KCOR; cumulative hazard; discrete hazard transform; observational data; mortality; cohort methods

# 1. Introduction

KCOR evaluates relative mortality outcomes between two fixed cohorts (e.g., dose groups) on calendar time. It addresses key challenges of observational comparison: time‑varying background mortality and baseline differences between cohorts. After slope‑normalization during quiet periods, KCOR constructs discrete cumulative hazards and tracks their baseline‑normalized ratio over time. Values above 1 indicate net excess cumulative hazard in the numerator cohort since baseline; values below 1 indicate net benefit.

The approach relies only on basic dates (birth, death, vaccination) and is intentionally modular with explicit diagnostics. The exponential baseline assumption is testable by verifying linearity of log‑mortality rates in quiet intervals.

# 2. Notation

|  |  |
| --- | --- |
| Symbol | Meaning |
| t | Week index (calendar time), with t = 0 at enrollment |
| MR(t) | Observed mortality rate in week t for a fixed cohort |
| r | Exponential slope of baseline mortality for the cohort |
| MR\_adj(t) | Slope‑adjusted mortality rate |
| hazard(t) | Discrete hazard in week t |
| CH(t) | Cumulative hazard up to week t |
| t0 | Baseline week used for normalization (default t0 = 4) |
| KCOR(t) | Cumulative‑hazard ratio, normalized at t0 |

# 3. Methods

## 3.1 Mathematical formulation

Slope normalization (per cohort):

Discrete hazard transform and cumulative hazards:

Baseline‑normalized cumulative‑hazard ratio (cohorts v over u):

Log‑scale uncertainty (Poisson approximation Var[CH] ≈ CH):

## 3.2 Expected‑deaths weighting for pooled summaries

For age‑stratified analyses, define weights using expected deaths in a quiet baseline period. Let N\_s(t) be persons at risk in stratum s and hazard\_s(t) the discrete hazard after slope‑adjustment. Then the baseline expected deaths are E\_s = ∑\_{t∈B} N\_s(t)·hazard\_s(t) and weights w\_s = E\_s / ∑\_j E\_j. A pooled cumulative hazard is CH\_pool(t) = ∑\_s w\_s·CH\_s(t).

## 3.3 Algorithm (pseudocode)

Input: Weekly counts or rates for fixed cohorts v and u; enrollment week t\_e=0; anchor windows (start date/end date for slope computation); baseline week t0 (typically week 5).

1) Aggregate to weekly MR\_v(t), MR\_u(t) for fixed cohorts defined at enrollment.

2) Compute slope r per cohort/stratum from quiet‑period anchors (geometric means in windows separated by Δt weeks where Δt is ideally around 52 weeks).

3) Slope‑adjust the mortality rates MR\_adj(t) = MR(t)·exp(−r·(t − t\_e)) for t ≥ 1.

4) Transform to discrete hazards: hazard(t) = −ln(1 − MR\_adj(t)) with clipping MR\_adj(t) ≤ 0.999.

5) Compute cumulative hazards CH(t) = ∑\_{i=0}^{t} hazard(i).

6) Normalize KCOR(t) at t0: divide by CH\_v(t0)/CH\_u(t0).

7) Compute log‑scale CIs with Var[CH] ≈ CH;

8) Report KCOR(t) with 95% bands.

## Implementation Notes

|  |
| --- |
| **Anchor‑selection heuristics**  • Identify two quiet (non‑differential) windows B1, B2 separated by Δt ≈ 8–12 weeks.  • Use ±2‑week windows and geometric means; avoid weeks with known shocks or reporting artifacts.  • Prefer trough‑centered windows when respiratory mortality is seasonal.  • Check both cohorts for parallel log‑MR trends within anchors; if not, adjust windows or stratify further. |
| **Smoothing & defaults**  • Compute MR(t) using a centered 3‑week moving average (fallback: 5‑week if noisy).  • Baseline normalization week t0 = 4 by default.  • Clip MR\_adj(t) at 0.999 before the hazard transform; treat missing MR(t) via local linear interpolation constrained to [0, 1).  • Document choices; include sensitivity to smoothing span and anchor placement. |

# 4. Application workflow

• Define fixed cohorts at enrollment (avoid time‑varying composition).

• Choose anchor windows (B1, B2) used for slope normalization; verify exponential baseline by linearity of log‑MR vs. t within anchors. These times are generally at least 6 months and preferably a year post enrollment to minimize interaction with the intervention. They should be separated by at least 26 weeks and ideally 52 weeks or more.

• Compute slope r per cohort/stratum; apply slope‑normalization, discrete hazard transform, and cumulative hazards.

• Normalize KCOR(t) at t0; produce CI bands and sensitivity overlays (anchors, smoothing).

• For pooled summaries, apply expected‑deaths weights across strata.

# 5. Diagnostics and sensitivity

• Exponential fit check: R² of log‑MR vs. time in quiet ranges; plot residuals.

• Anchor sensitivity: vary windows by ±1–2 weeks and re‑estimate r; overlay KCOR(t).

• Smoothing robustness: 3 vs. 5‑week moving average; optionally LOESS with small span.

• Negative controls: randomized split cohorts (expect KCOR≈1).

• Placebo enrollment dates: shift t\_e by ±k weeks; conclusions should be stable.

• Subgroup stratification: age bands; pooled CH using expected‑deaths weights.

# 6. Benefits and limitations

Benefits: (i) minimal data needs, (ii) mathematically exact discrete‑time cumulative‑hazard comparison, (iii) explicit diagnostics, (iv) interpretable time‑resolved net‑benefit curve, (v) modular and reproducible.

Limitations: (i) requires identifiable quiet periods and approximate exponential baselines, (ii) cohort‑differential shocks within anchors bias r, (iii) as with all observational designs, unmeasured time‑varying confounding may remain.

# 7. Reporting and reproducibility

• Define cohorts and enrollment; provide inclusion/exclusion flow.

• Justify anchor selection against the smoothed raw MR; show quiet‑period fits (log‑MR vs. t) and R².

• Specify smoothing, clipping, interpolation, and t0.

• Present KCOR(t) with 95% CIs, plus anchor/smoothing sensitivity overlays.

• Include negative‑control and placebo‑enrollment analyses.

• Release code/config where permissible for replication.

# 8. Source code and documentation

Full source code and extensive documentation on the method and validation tests is publicly available at <https://github.com/skirsch/KCOR>.

# 9. Figures and tables anchor selection → slope normalization → discrete hazard transform → cumulative hazards → baseline normalization → KCOR(t).

Figure 1. KCOR workflow

# Author Contributions

Conceptualization, methodology, software, validation, writing—original draft, writing—review & editing: S.K.

# Acknowledgments

Thanks to colleagues who provided feedback on discrete‑hazard implementations, anchor selection, and diagnostics.

# References

[1] KCOR README and method notes ([repository documentation](https://github.com/skirsch/KCOR/tree/main)).

[2] Levi R., Ladapo J.A., et al., [Twelve‑Month All‑Cause Mortality after Initial COVID‑19 Vaccination](https://www.medrxiv.org/content/10.1101/2025.04.25.25326460v1), medRxiv (preprint), 2025.

[3] Chemaitelly, H. et al., [Assessing healthy vaccinee effect in COVID-19 vaccine effectiveness studies: a national cohort study in Qatar](https://elifesciences.org/articles/103690)