Journal of Clinical Epidemiology 191 (2026) 112122

COMMENTARY

Defining survival epidemiology: postdiagnosis population science for people living with disease

Raphael E. Cuomo*

School of Medicine, University of California, San Diego, CA, USA

Accepted 23 December 2025; Published online 27 December 2025

Abstract

Objectives: Epidemiology is largely organized to explain who becomes ill, yet many clinical and public health decisions occur after diagnosis. I introduce and formally define survival epidemiology as a new branch of science focused on assessing how people live longer and better with established disease, and I provide justification that prevention estimates should not be assumed to apply postdiagnosis.

Study Design and Setting: Conceptual and methodological commentary synthesizing evidence across cardiovascular, renal, oncologic, pulmonary, and hepatic conditions and integrating causal-inference and time-to-event principles for postdiagnosis questions.

Results: Across diseases, associations measured for incidence often fail to reproduce, and sometimes reverse, among patients with established disease. Diagnosis acts as a causal threshold that changes time scales and bias structures, including conditioning on disease (collider stratification), time-dependent confounding, immortal time bias, and reverse causation. Credible postdiagnosis inference requires designs that emulate randomized trials; explicit alignment of time zero with clinical decision points; strategies defined as used in practice; and handling of competing risks, multistate transitions, and longitudinal biomarkers (including joint models when appropriate). Essential postdiagnosis data include stage, molecular subtype, prior therapy lines, dose intensity and modifications, adverse events, performance status, and patient-reported outcomes. Recommended practice is parallel estimation of prevention and postdiagnosis survival effects for the same exposure−disease pairs and routine reporting of heterogeneity by stage, subtype, treatment pathway, and time since diagnosis.

Conclusion: Prevention and postdiagnosis survival are distinct inferential targets. Journals should require clarity on whether claims pertain to prevention or survival and report target-trial elements; guideline bodies should distinguish prevention from survival recommendations when evidence allows; and funders, training programs, and public communication should support survival-focused methods, data standards, and context-specific messaging for people living with disease.

Keywords: Survival epidemiology; Prevention epidemiology; Clinical Epidemiology; Epidemiology; Population science; Scientific discipline

1. Introduction and definition

The conceptual architecture of epidemiology has historically emphasized explaining who becomes diseased, even though survival and prognosis have long been recognized within the framework of primary, secondary, and tertiary prevention. For more than a century, that prevention first mindset has organized questions, datasets, training, and guidance to the public. Clinical decisions increasingly arise after diagnosis, when pathophysiology, treatment exposures, and competing risks reshape the meaning of familiar risk factors. A growing body of evidence shows that relations observed in incidence studies often fail to reproduce in cohorts of patients with established disease, and in some settings the sign reverses. The so-called obesity paradox in heart failure and chronic kidney disease, the association of higher serum cholesterol with lower short-term mortality in heart failure, and reports that modest alcohol consumption predicts improved survival among patients with cardiovascular disease or certain cancers illustrate that effects estimated for prevention cannot be assumed to hold after diagnosis [1−5]. In the prevention taxonomy, much of this work sits within tertiary prevention, prognosis research, and survivorship care; survival epidemiology is intended to formalize and extend this tradition.

Survival epidemiology describes a complementary, explicitly named branch of population science devoted to outcomes after diagnosis and to the systematic study of how risk factor relations change across the prevention-to-

* Corresponding author. School of Medicine, University of California, 9500 Gilman Dr. La Jolla, San Diego, CA 92093, USA.

E-mail address: racuomo@ucsd.edu.

https://doi.org/10.1016/j.jclinepi.2025.112122

What is new?

Key findings

- Prevention associations often fail to reproduce, and sometimes reverse, after diagnosis across multiple diseases.
- Diagnosis introduces distinct postdiagnosis biases and time scales (e.g., collider stratification, time-dependent confounding, immortal time bias, reverse causation) that require survival-specific designs and analyses.

What this adds to what is known?

- Defines ''survival epidemiology'' as a named conceptual and methodological umbrella for postdiagnosis questions, specifying core design and data requirements, including trial emulation, decision-aligned time zero, clinically grounded strategies, competing risks and multistate methods, joint modeling, and capture of stage, subtype, treatment, and toxicity variables.

What is the implication and what should change now?

- Prevention and postdiagnosis survival effects for the same exposure−disease pairs should be estimated separately, and effects derived for one state should not be assumed to apply to the other.
- Journals and guideline bodies should explicitly distinguish prevention versus survival claims and expect target-trial reporting; funders and training programs should prioritize survival-focused methods, reporting standards, and data infrastructure.

survival boundary. It defines a conceptual and methodological umbrella that brings together study designs that condition on diagnosis and follow patients through treatment pathways; the distinctive bias structures that arise when inclusion requires disease; causal inference frameworks tailored to postdiagnosis questions; and the epidemiologic interpretation of survival estimates for clinical and policy decisions. Classical survival analysis provides a family of time-to-event tools that can be used in many settings; survival epidemiology specifies when and how those tools, together with modern causal methods, are applied to people living with disease. It is not survivorship care by another name. Survivorship often denotes services and quality of life after primary therapy and is typically framed as a clinical or psychosocial domain. Survival epidemiology asks how behaviors, environments, and biological traits influence mortality, recurrence, progression, treatment tolerance, functional outcomes, and costs once a person has a defined disease. It treats diagnosis as a causal threshold that introduces new sources of bias, new time scales, and new effect modifiers, including stage, molecular subtype, treatment sequence, toxicity, cachexia, sarcopenia, immune suppression, organ reserve, and the practical constraint that patients must survive long enough to benefit from putatively protective behaviors. The object of inference differs: while preventive epidemiology studies how to avoid disease, survival epidemiology studies how to live longer and better with disease. The aim is not to position prevention and survival epidemiology in opposition, but to make explicit that different questions, time scales, and sources of bias dominate before and after diagnosis. To our knowledge, this commentary offers the first comprehensive definition and formal proposal of ''survival epidemiology'' as a named branch of population science.

This conceptual split is not academic. People move between states across the life course. An exposure that prevents disease in a healthy population can correlate with worse function or tolerance once treatment begins. A nutrient that lowers long-term risk can be unhelpful in the presence of cachexia. The language we use in public health has real clinical effects. Headlines that collapse prevention and survival evidence into one narrative create confusion for patients who are already receiving therapy. A branch named for postdiagnosis outcomes would make the boundary legible and would signal that new standards apply once disease is present. It would also align with the prognosis research agenda that asks questions about outcomes among people who already have disease, bringing a population science lens to what is often treated as purely clinical territory [6].

2. Methods, data, and biases after diagnosis

The tools to support credible postdiagnosis inference now exist and should be considered standard for work in this space. In survival epidemiology, these tools are used not as ends in themselves but as part of a broader design-based strategy for drawing causal inferences among people who already have disease. Classical survival analysis models, such as Cox and flexible parametric models, competing risks models, and multistate models, remain central components of this toolkit but are embedded within target-trial−inspired designs and explicit causal frameworks [7−9]. Target trial emulation, marginal structural models, *g* computation, dynamic treatment regime methods, and joint modeling of longitudinal biomarkers with survival have matured and entered routine use [10−13]. These designs make explicit the eligibility window, align time zero with decision points, define treatment strategies as they would be assigned in practice, and address informative censoring. Cloning, censoring, and weighting can be used to represent alternative strategies in routine care when randomization is

not feasible. Joint models link a biomarker trajectory with a time-to-event outcome so that feedback between the biomarker and treatment choices is handled rather than ignored [13]. Multistate and competing risks frameworks allow investigators to separate the probability of progression from the probability of death without progression, which is essential when toxicities and early complications are part of the pathway [7,8]. In practice, this often involves choosing between cause-specific hazard models, which are closely aligned with etiologic and dynamic treatment questions, and subdistribution hazard models, which directly model cumulative incidence for prognostic purposes. For multistate processes, explicit Markov or semi-Markov assumptions are required; in many postdiagnosis settings, time since entry into the current state (eg, time since progression or initiation of a therapy line) provides a more realistic clock than time since diagnosis alone. Specifying the target trial up front also prevents a series of avoidable errors in observational analyses, including misaligned time zero and self-inflicted immortal time.

In addition to time-dependent predictors, survival epidemiology must routinely accommodate time-dependent effects and nonproportional hazards. Hazards associated with behaviors or biomarkers can change with time since diagnosis, treatment phase, or progression status, so models may require interactions between covariates and flexible functions of time, piecewise hazard specifications, or fully time-varying coefficients. These time-varying effects are often mechanistically meaningful (eg, the effect of body mass index on mortality during intensive chemotherapy vs long-term survivorship) and should be anticipated rather than treated as modeling afterthoughts.

Data infrastructure has also advanced. Clinical data assets can capture stage, residual tumor burden, actionable mutations, prior lines of therapy, performance status, dose reductions, and adverse events, which general population cohorts rarely measure. Linkage across electronic health records, registries, pharmacy dispensing, and patient-reported outcomes, mapped to common data models and connected to biospecimens, allows postdiagnosis questions to be answered with appropriate clinical detail [14]. Performance status measures such as the Eastern Cooperative Oncology Group (ECOG) scale anchor many clinical decisions and should be routinely available in analytic datasets [15]. These resources support the routine estimation of prevention and survival effects for the same exposure−disease pairs and the stratification required to make estimates decision-relevant. They also support time-updated covariates, which lets analysts reflect the way clinicians actually make decisions as patients move through lines of therapy.

Studies that condition on disease status are uniquely vulnerable to biases that can hide or invert real effects. Collider stratification arises when inclusion requires disease. If an exposure increases disease risk, it can appear protective among those selected into the diseased state even when no protective effect exists [16−18]. Time-dependent confounding is common because treatment decisions respond to evolving symptoms or biomarker deterioration. Immortal time bias can inflate benefits when exposures are defined using future information, which is a frequent problem when a therapy or behavior requires survival to a later time point [19−21]. Reverse causation is pervasive because weight loss, low cholesterol, or reduced alcohol intake often reflect subclinical progression or treatment toxicity rather than protective behavior. These biases occur in prevention research, yet they are central and structural after diagnosis. The existence of these structural hazards is a practical reason to formalize a branch whose training and peer review culture keeps them visible.

The field should also adopt a routine test of prevention-to-survival nonequivalence, sometimes termed Cuomo's Paradox in public discussion, within the same exposure and disease pair [1]. This observation conveys the empirical observation that exposure−disease relations estimated for incidence frequently fail to reproduce, and can sometimes reverse, in postdiagnosis survival analyses of the same exposure−disease pair. For a given exposure and disease, studies should delineate the effect of an exposure on incidence with its effect on postdiagnosis survival, preferably within harmonized designs and populations. When effects are not equivalent, guidance should default to survival-specific evidence rather than extrapolation from prevention studies. Effect heterogeneity should be reported in ways that match decisions that patients and clinicians face. Estimates averaged across stages, molecular classes, treatment eras, and time since diagnosis obscure the interactions that govern treatment tolerance and functional outcomes. Stratification by stage, subtype, treatment pathway, and clinically meaningful time windows should be routine rather than exceptional.

3. Applications across diseases

Cancer illustrates the need for survival-specific thinking. Traditional messages emphasize weight control, lipid reduction, and alcohol avoidance to prevent incidence. After diagnosis, caloric sufficiency, maintenance of lean mass, and preservation of organ function can dominate survival. Appetite, nausea control, sleep, and social engagement become relevant survival correlates because they influence treatment adherence and the ability to complete cycles at planned doses. Observational studies of alcohol after breast cancer diagnosis have reported neutral or modestly favorable survival associations at low levels of intake, a pattern that does not mirror prevention messaging and that likely reflects the influence of appetite, comorbidity, and treatment tolerance in the postdiagnosis state [3]. Obesity raises incident risk for several cancers, yet higher body mass index has correlated with improved outcomes among patients treated with immune checkpoint inhibitors in several series. This pattern does not imply that adiposity is beneficial in

general populations. It suggests that inflammatory tone, pharmacokinetics, and symptom thresholds differ once modern immunotherapy is in use. Cancer cachexia further illustrates why prevention targets cannot be applied without modification after diagnosis. Loss of skeletal muscle mass and function is common, predicts poor tolerance of therapy, and shortens survival, which reframes nutritional guidance in oncology [22−24]. Patterns such as these are likely driven by a mixture of mechanisms, including heterogeneous obesity phenotypes, differential frailty and comorbidity, and time-dependent effects, rather than by a simple protective role for higher body mass index. Survival epidemiology therefore treats them as signals that conditioning on diagnosis, treatment selection, and changing hazards can alter observed associations, not as straightforward endorsements of adiposity or alcohol use in the general population.

Cardiovascular disease and chronic kidney disease provide related examples. Obesity increases cardiovascular incidence. In heart failure cohorts, however, higher body mass index has been associated with better survival, and higher serum cholesterol has tracked with lower short-term mortality [2,5]. These associations should not be read as new prevention targets. They are signals that energy reserves, catabolic drive, and treatment tolerance can dominate prognosis in established disease. Admission systolic blood pressure, a prevention hazard in healthy populations, shows the opposite pattern in acute decompensated heart failure, where higher values at presentation are associated with lower mortality [5]. In chronic kidney disease and dialysis populations, so-called reverse epidemiology has been described, where higher blood pressure, higher body mass index, and higher cholesterol are associated with improved survival [4]. This highlights that cachexia, inflammation, and malnutrition can be stronger proximate risks once organ failure is present. Nonproportional hazards and phenotype-specific differences can further contribute to these apparent reversals and are natural targets for survival epidemiology to disentangle.

Pulmonary and liver diseases show similar patterns. In chronic obstructive pulmonary disease, low body mass index predicts poor outcomes, while weight stability and resistance training can preserve function despite traditional prevention messages that stress weight reduction. In cirrhosis, sarcopenia is a strong predictor of mortality and transplant outcomes, which means that blanket sodium or caloric restriction can carry survival tradeoffs if they worsen lean mass. In neurologic disease, weight loss and frailty predict faster functional decline in amyotrophic lateral sclerosis and Parkinson disease. In each case, the prevention narrative cannot simply be repurposed for people living with disease. The effect modifiers are different, the time scales are different, and the relevant outcomes include treatment tolerance and the ability to remain independent at home.

A survival epidemiology lens also clarifies why some well-intended public messages backfire in clinical settings. A healthy adult may use alcohol avoidance to reduce long-term cancer risk. A patient with advanced cancer who is losing weight and has poor appetite may interpret the same message as a reason to avoid oral calorie sources during chemotherapy, which can worsen fatigue and compromise dose intensity. A prevention message that focuses only on average effects in healthy populations can be unsafe when transplanted to people receiving treatment. Survival epidemiology would separate prevention guidance from survival guidance so that clinicians can counsel patients in ways that reflect the state they occupy.

4. Practice and policy implications

If survival epidemiology is formalized, several practical changes follow. At the journal level, editorial guidance can exist that explicitly requests a delineation of claims intended as for the purpose of prevention or for survival after diagnosis and should expect reporting that makes the target trial and time zero explicit. Manuscripts should describe how eligibility, strategies, and censoring were defined; how diagnosis-conditioned selection, time-dependent confounding, and potential immortal time bias were addressed; and which estimand was targeted [10,11,19,20]. Editors and reviewers should expect a clear rationale for any extrapolation from prevention studies to postdiagnosis guidance and, when the same exposure−disease pair has divergent prevention and survival estimates, should give primary weight to survival-specific evidence when counseling people who are already diagnosed.

At the guideline level, bodies should, when survival evidence exists and is decision-relevant, consider clear delineations of prevention and postdiagnosis survival recommendations. When evidence diverges, the difference should be explained in plain language for clinicians and for patients. Professional societies and public health agencies can adopt this split pragmatically by adding survival tables where appropriate. Authoritative survivorship guidance already treats postdiagnosis behavior as a distinct decision space, which provides an immediate template for survival-specific recommendations [25].

At the level of funders and data partners, progress can be accelerated by encouraging analyses that estimate both prevention and postdiagnosis survival effects for the same exposure−disease pairs when scientifically appropriate and logistically feasible. Data partners can prioritize capture of stage, residual disease, mutational status, prior therapies, dose intensity, adverse events, performance status, and patient-reported outcomes, which are often missing from legacy cohorts, so that survival questions can be addressed with adequate clinical detail. Training programs can develop core sequences or stand-alone courses on postdiagnosis causal inference, diagnosis-conditioned selection bias diagnostics, competing risks and multistate modeling, joint modeling of longitudinal biomarkers with survival,

and the integration of clinical detail with population denominators, creating a foundation for future textbooks and tutorials in survival epidemiology. Public communication strategies can then mirror the prevention−survival split by emphasizing that guidance for people living with disease should be based on survival-specific evidence and communicated through clinicians and carefully crafted patient-facing materials rather than through oversimplified headlines.

Survival epidemiology also underscores the distinction between discrimination and calibration in postdiagnosis prediction models. Time-to-event models may retain acceptable discrimination (eg, C-statistics or time-dependent area under the curve) even when they are poorly calibrated, because treatment standards, supportive care, or patient mix have changed, or because competing risks and treatment pathways have been mis-specified. For counseling, shared decision-making, and resource allocation, well-calibrated absolute risks at clinically relevant horizons are as important as rank-ordering, so reporting both discrimination and calibration (ideally stratified by stage, treatment pathway, and time since diagnosis) should become part of routine methodological standards.

The payoff for patients and clinicians is immediate. Advice that is correct for prevention can be counterproductive during treatment. Caloric sufficiency, maintenance of lean mass, and preservation of organ function can dominate survival for many patients. Blood pressure and lipid targets that reduce long term risk in healthy populations may require reframing when cachexia, sarcopenia, and treatment toxicity are present. Survival epidemiology measures what prolongs life and function in the states patients occupy and explains why relations can differ from those seen before diagnosis. Naming the branch, teaching it, funding it, and holding it to its own standards is the next step.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

References

[1] Cuomo RE. The nutritional epidemiology risk−survival paradox. J Nutr 2025;155(10):3148−51.

[2] Curtis JP, Selter JG, Wang Y, Rathore SS, Jovin IS, Jadbabaie F, et al. The obesity paradox: body mass index and outcomes in patients with chronic heart failure. Arch Intern Med 2005;165:55−61.

[3] Kwan ML, Kushi LH, Weltzien E, Tam EK, Castillo A, Sweeney C, et al. Alcohol consumption and breast cancer recurrence and survival among women with early-stage breast cancer: The Life After Cancer Epidemiology Study. J Clin Oncol 2010;28:4410−6.

[4] Kalantar-Zadeh K, Block G, Humphreys MH, Kopple JD. Reverse epidemiology of conventional cardiovascular risk factors in patients with chronic hemodialysis. Kidney Int 2003;63:793−808.

[5] Kalantar-Zadeh K, Block G, Horwich TB, Fonarow GC. Reverse epidemiology of conventional cardiovascular risk factors in patients with chronic heart failure. J Am Coll Cardiol 2004;43(8):1439−44.

[6] Hemingway H, Croft P, Perel P, Hayden JA, Abrams K, Timmis A, et al. Prognosis research strategy (PROGRESS) 1: a framework for researching clinical outcomes. BMJ 2013;346:e5595.

[7] Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. J Am Stat Assoc 1999;94(446):496−509.

[8] Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: competing risks and multi state models. Stat Med 2007;26(11):2389−430.

[9] Ng R, Kornas K, Sutradhar R, Wodchis WP, Rosella LC. The current application of the Royston−Parmar model for prognostic modeling in health research: a scoping review. Diagn Progn Res 2018;2(1):4.

[10] Hernán MA, Robins JM. Using big data to emulate a target trial when a randomised trial is not available. Am J Epidemiol 2016;183(8):758−64.

[11] Hernán MA, Robins JM. Causal inference: what if. Boca Raton: Chapman and Hall/CRC; 2020.

[12] Young JG, Cain LE, Robins JM, O'Reilly EJ, Hernán MA. Comparative effectiveness of dynamic treatment regimes. Stat Med 2011;30(11):1230−50.

[13] Rizopoulos D. Joint models for longitudinal and time-to-event data: with applications in R. Boca Raton: Chapman and Hall/CRC; 2012.

[14] Hripcsak G, Duke JD, Shah NH, Reich CG, Huser V, Schuemie MJ, et al. Observational health data sciences and informatics. Stud Health Technol Inform 2015;216:574−8.

[15] Oken MM, Creech RH, Tormey DC, Horton J, Davis TE, McFadden ET, Carbone PP. Toxicity and response criteria of the Eastern Cooperative Oncology Group. Am J Clin Oncol 1982;5(6):649−55.

[16] Cole SR, Platt RW, Schisterman EF, Chu H, Westreich D, Richardson D, Poole C. Illustrating bias due to conditioning on a collider. Int J Epidemiol 2010;39:417−20.

[17] Munafò MR, Tilling K, Taylor AE, Evans DM, Davey Smith G. Collider scope: when selection bias can substantially influence observed associations. Int J Epidemiol 2018;47(1):226−35.

[18] Hernán MA, Monge S. Selection bias due to conditioning on a collider. BMJ 2023;381:p1135.

[19] Suissa S. Immortal time bias in pharmacoepidemiology. Am J Epidemiol 2008;167(4):492−9.

[20] Lévesque LE, Hanley JA, Kezouh A, Suissa S. Problem of immortal time bias in cohort studies: example using statins and mortality. BMJ 2010;340:b5087.

[21] Tyrer F, Bhaskaran K, Okoli GN, et al. Immortal time bias for life-long conditions in retrospective observational studies using electronic health records. BMC Med Res Methodol 2022;22:219.

[22] Fearon K, Strasser F, Anker SD, Bosaeus I, Bruera E, Fainsinger RL, et al. Definition and classification of cancer cachexia: an international consensus. Lancet Oncol 2011;12(5):489−95.

[23] Prado CM, Lieffers JR, McCargar LJ, Reiman T, Sawyer MB, Martin L, Baracos VE. Prevalence and clinical implications of sarcopenic obesity in patients with solid tumours of the respiratory and gastrointestinal tracts. Lancet Oncol 2008;9(7):629−35.

[24] Baracos VE, Martin L, Korc M, Guttridge DC, Fearon KCH. Cancer-associated cachexia. Nat Rev Dis Primers 2018;4:17105.

[25] Rock CL, Thomson C, Sullivan K, Howe CL, Kushi LH, Caan BJ, et al. Nutrition and physical activity guideline for cancer survivors. CA Cancer J Clin 2022;72(3):230−62.