Application to the UAW-GM Cohort

In the simulation study, we showed that under several scenarios compatible with our hypothetical causal structure, the AWKM survival estimator had small bias compared to the WKM estimator. The bias was smaller when the cumulative incidence of the outcome was low and at later follow-up time points. Now, we estimate cancer-free survival in a real-world context. Using data from the UAW-GM Cohort study, we follow 38 553 individuals starting from hire to 40 years after hire for incidence of colon cancer. This duration of follow-up spans the vast majority of individuals' working lifetimes. As in the simulation, the UAW-GM data are longitudinal data with baseline covariates, time-varying covariates, and a survival outcome. We assume that the observed data are iid draws from an underlying causal data generating system.

Since we are working with observational data, the evaluation of the no-interference, causal consistency, and ignorability assumptions are critical for causal inference. The no-interference assumption may be problematized by the fact that there are a finite number of job types in the factory setting. If one worker operates a metalworking machine, then the the other workers are not able to operate that machine at that time. Instead, they may be assigned to assembly tasks, which have lower MWF exposure opportunities. That said, since these factories are quite large, there may be approximate independence. The consistency assumption is also problematic. The MWFs of interest are complex chemical mixtures whose composition changes by design and by nature. Over the last several decades, the formulation of MWFs has changed significantly in reaction to performance needs and toxicity concerns (Mirer 2003; Byers 2006). The composition of MWFs also undergoes unintentional change

Colon cancer incidence was obtained via linkage to the Michigan Cancer Registry, which recorded cancer incidence cases starting on January 1, 1985. The cohort is comprised of individuals hired between 1938 and 1985. Cancer-free survival to the start of the registry is a left-filtering process possibly in the presence of the HWSE as investigated in the simulations. Over the 40 year follow-up period, vital status was obtained through the Social Security Administration, the National Death Index, as well as records provided by the UAW. Baseline confounders included race, sex, plant, and year of hire. Time-varying confounders included age, cumulative time off, and exposure to the metalworking fluids not under study. These terms were included in the estimation of both the treatment and censoring mechanisms, which were estimated with stratification on every year of follow-up. The exposure rules of interest were identical to those in the simulation study: a_0 never exposed and a_1 always exposed while at work. We assume that the effect of exposure on health status, mortality, cancer incidence, etc occurs after a lag of 15 years. Similarly, we assume that the effects of cumulative time off occur after a 15 year lag. Counterfactual survival under rules a_0 and a_1 were estimated using the WKM and AWKM estimators. Summary statistics for the full study population and the colon cancer cases are presented in Table 1.

Byers, Jerry P. 2006. *Metalworking Fluids*. CRC Press.

Mirer, Franklin. 2003. "Updated Epidemiology of Workers Exposed to Metalworking Fluids Provides Sufficient Evidence for Carcinogenicity." *Applied Occupational and Environmental Hygiene* 18 (11): 902–12.

Table 1: Study population characteristics.

	Full cohort		Colon cancer cases	
n (person-years)	38 553	(1 402 372)	267	(9 635)
Race (%)				
Black	7 133	(18.5)	83	(31.1)
White	31 420	(81.5)	184	(68.9)
Sex (%)				
Female	4 757	(12.3)	49	(18.4)
Male	33 796	(87.7)	218	(81.6)
Plant (%)				
Plant 1	9 092	(23.6)	66	(24.7)
Plant 2	17 090	(44.3)	107	(40.1)
Plant 3	12 371	(32.1)	94	(35.2)
Ever exposed to MWF (%)	22193	(57.6)	104	(39.0)
Year of hire (mean (SD))	1963	(11.62)	1968	(8.79)
Age (mean (SD))	63.38	(9.08)	67.78	(9.27)
Cumulative time off (mean (SD))	0.04	(0.14)	0.05	(0.15)