

# A NEW APPROACH TO CAUSAL INFERENCE IN MORTALITY STUDIES WITH A SUSTAINED EXPOSURE PERIOD—APPLICATION TO CONTROL OF THE HEALTHY WORKER SURVIVOR EFFECT

JAMES ROBINS

Harvard School of Public Health  
665 Huntington Avenue  
Boston, MA 02115

(Received 19 June 1985; revised 12 January 1986)

**Abstract**—In observational cohort mortality studies with prolonged periods of exposure to the agent under study, it is not uncommon for risk factors for death to be determinants of subsequent exposure. For instance, in occupational mortality studies date of termination of employment is both a determinant of future exposure (since terminated individuals receive no further exposure) and an independent risk factor for death (since disabled individuals tend to leave employment). When current risk factor status determines subsequent exposure and is determined by previous exposure, standard analyses that estimate age-specific mortality rates as a function of cumulative exposure may underestimate the true effect of exposure on mortality whether or not one adjusts for the risk factor in the analysis. This observation raises the question, which if any population parameters can be given a causal interpretation in observational mortality studies?

In answer, we offer a graphical approach to the identification and computation of causal parameters in mortality studies with sustained exposure periods. This approach is shown to be equivalent to an approach in which the observational study is identified with a hypothetical double-blind randomized trial in which data on each subject's assigned treatment protocol has been erased from the data file. Causal inferences can then be made by comparing mortality as a function of treatment protocol, since, in a double-blind randomized trial missing data on treatment protocol, the association of mortality with treatment protocol can still be estimated.

We reanalyze the mortality experience of a cohort of arsenic-exposed copper smelter workers with our method and compare our results with those obtained using standard methods. We find an adverse effect of arsenic exposure on all-cause and lung cancer mortality which standard methods fail to detect.

## NOMENCLATURE

	<i>Section(s) where defined</i>
$A$ -complete Stage 0 reduction .....	7A
$A_{-i_s}$ .....	4A
$A(-i_s)$ .....	4G
alternative designed randomized trial .....	2D, 4F
anti-meld .....	Appendix F
Assumption $\mathcal{G}_1$ .....	11B
Assumption $\mathcal{G}_2$ .....	11A
Assumption $\mathcal{G}_3$ .....	11B
$A(x)$ .....	11B
$B$ -complete PL-sufficient Stage 0 reduction .....	7A
$B$ -complete Stage 0 reduction .....	7A
$c$ .....	9A
$\bar{c}$ .....	9A

C8.3 (a PISTG) .....	8D.2
$C(t_s)$ .....	8D.1
causal melded Stage 0 reduction .....	Appendix F
causal risk factor for death .....	3C, 8A.2
causal risk factor for exposure .....	8A.2
CF8.3 (a PISTG) .....	8D.3
CISTG .....	3B, 4C
coarser .....	3A, 4A
"detailed as the data" .....	4B
$D_{\cdot i, j_s}$ .....	4C
$\Delta t = t_s - t_{s-1}$ .....	4B
$D(t_u)$ .....	4C
"empirical healthy worker survivor effect" .....	3G
$e(t_s)$ .....	2C
$E(t_s)$ , actual .....	2C
$E(t_s)$ , measured .....	3
F8.1 (a PISTG) .....	8C
F8.3 (a PISTG) .....	8D.1
finer .....	3A, 4A
finest FR MCISTG .....	4G
finest MCISTG .....	3B, 4G
FR CISTG .....	4E
FR CISTG $RD_2$ .....	12A
full-independence assumption .....	12A
fundamental PISTG .....	4F
G .....	4D
" $G$ " .....	4A
$G^{3,4}$ as a notational convention .....	3E
$G^A$ .....	4D
$G^{A(i_s)}$ .....	4H
" $G^{A(i_s)}$ " .....	4H
$G^{A(\cdot i_s, t_{s+1})}$ .....	4H
$G$ , as a PEH .....	2C
$G^{A(x)}$ .....	11B
$G$ -causal parameter .....	4D
$G$ -computation algorithm .....	3D
$\gamma_D(t + \Delta t   "G_1")$ .....	4E
$\gamma_D(t + \Delta t   \cdot i_s, j_s)$ .....	4B
generalized minimum latent period .....	11B
generalized treatment .....	3B, 4D
generalized treatment algorithm .....	3B
$\gamma(i_s)$ .....	4B
$\gamma(i_s   G)$ .....	4D
$\gamma(\cdot i_s, j_s)$ .....	4B
$G$ -independence assumption .....	12A, Appendix G
$G$ -null hypothesis .....	6A
" $G$ "-null hypothesis .....	6A
$G$ -null test .....	6A
$G$ -null test for $D_1$ .....	12B
$G(t_s)$ .....	11B
"healthy worker survivor effect" .....	1, 3G
$H(\cdot i_s)$ .....	4E
$HT(\cdot i_s)$ .....	4C
$HT_1(\cdot i_s)$ .....	12A
healthy worker survivor effect is operative for less than $x$ years .....	11A
identifiable temporal assumption .....	2C, 7A, 7B
independent population risk factor .....	8A.2
internodal line .....	3A, 4A
intranodal line .....	3A, 4A
$\cdot i_s$ .....	4A, 4B
$[\cdot i_s]$ .....	4B

$i_s \in G$	4E
$i_s \in "G"$	4E
$i_s j_s$	4A, 4B
$[i_s j_s]$	4B, 4G
' $i_s j_s$ ' as used other than in Sec. 4	6A
' $i_s j_s$ ' as used in Sec. 4	4C
$i_s \in m^A$	7A
$I(t_n)$	4C
$j_s(i_s)$	4A
$J(t_u)$	4C
$K^A$	4D
lagged exposure test algorithm	11C
large sample limiting model 1	5A
large sample limiting model 2	5B
$\lambda_{D_1}(t   "G")$	12A
$\lambda_{D_1}^*(t   "G_1^A")$	12A
$\lambda_D(t   "G")$	4E
$le$	10B
$ls$	11A
$\bar{ls}$	11A
$l(t_s)$	5A
$L(t_s)$	3, 5A
$m$ , as sample size	4B
$M$	4B
$m^A$ as a function defined on a PISTG	4E
$m^A(i_s j_s)$	4E
MCISTG	3B, 4C
melded reduction	Appendix F
minimum latent period	11B
MPISTG	3A, 4B
$N_{i_s}$	4A
$N_{i_s j_s}$	4A
$N[i_s^B]$	5A
nonidentifiable temporal assumption	2C, 7A, 7B
nonparametric test	6B
NPMLE	3D
null hypothesis of no exposure effect controlling for cigarette smoking	9A
OEH	2A
OPISTG	4B
$p(D > t   "G_1")$	4E
PEH	2A
$p(i_s, D > t   "G")$	12A
$p^*(i_s, D > t   "G")$	12A
$p(i_s   G)$	4D, 4E
$p(i_s   "G_1")$	4E
$p(i_s j_s(t_{s+1})   "G_1")$	4E
$p(i_{s+1}   "m^A")$	4E
PISTG	4B
PISTG $A(x)$	11B
PL-sufficient	Appendix A
PL-sufficient Stage 0 reduction	7A
population causal parameter	4D
population predictor of exposure	8A.2
population predictor of future exposure	8A.2
R CISTG	4E
$RD_2$	12A
right circumference points	4A
R SCISTG	4E
SCISTG	4F
$S_{D_1}(t   "G_1^A")$	12A
$S_{D_1}^*(t   "G_1")$	12A

semi-melded Stage 0 reduction .....	Appendix F
semi-monotone deleterious (beneficial) .....	3E
sharp null hypothesis .....	6A
$S(i_s j_s)$ .....	4B, 4D
S08.3 (a PISTG) .....	8D.1
sparse data limiting model .....	5A
Stage 0 counterpart .....	7A
Stage 0 reduction .....	7A
Stage 1 counterpart .....	4H
Stage 1 reduction .....	4H
STG .....	4A
$S(t   G)$ .....	3B, 4D
$S(t   "G_1")$ .....	4E
$S(t   G, i)$ .....	3B
$S(t, K_1^A, K_2^A)$ .....	4D
$S(t_{k+1}   "m^A")$ .....	4E
$S[t_s + \Delta t   E(t_s)]$ .....	6B
$S[t_s + \Delta t   E(t_s - x)]$ .....	11B
$S[t_s + \Delta t   E(t_s), L(t_s)]$ .....	8B
$S(t_s, G_1, G_2)$ .....	3B, 4D
$S(t_s, "G_1", "G_2")$ .....	4E
$S(t_s, G_1, G_2) \equiv 0$ .....	6A
treatment .....	3B, 4C
valid test .....	6B
$w(\cdot i_k, t_s)$ .....	6A
$w_i(m^A, t_k)$ .....	4E
x-doomed .....	11A

## 1. INTRODUCTION

The analysis of occupational cohort mortality studies has traditionally been plagued by the bias resulting from improper comparisons of working populations with the general population. For example, the age-specific mortality rate for arteriosclerotic cardiovascular disease (ASCVD) in an unexposed working population is usually only 60–90% of the rate in the general population. Thus, the general population cannot serve as an appropriate control group when interest is in detecting relative risks in the range 1.5–2.0. Since present day exposures are, in general, lower than exposures experienced in the past, and since most substances associated with the large increases in relative risk may have already been discovered, occupational epidemiology has become increasingly concerned with detecting relative risks less than 2.0.

Recognizing that the general U.S. population is not an adequate control group, occupational epidemiologists have increasingly relied upon comparisons, within a single cohort, among workers who differ in levels of exposure. Unfortunately, if workers at increased risk terminate employment early, standard intracohort methods of analysis that estimate mortality as a function of cumulative exposure can underestimate the true effect of exposure on mortality, whether or not one adjusts for time of termination of employment[1–3]. Thus, even in intracohort analyses, increases in the relative risk in the range 1.5–2.0 due to occupational exposures can be masked by the early termination of workers with poor prognosis (*which we refer to as the healthy worker survivor effect*). In this monograph we present a set of statistical methods specifically designed to control bias due to the healthy worker survivor effect.

Although Gilbert recognized that, for chronic disabling illnesses such as ASCVD, bias due to the healthy worker survivor effect could not be controlled by standard methods, she conjectured that, for diseases for which the interval between clinical manifestation and death is brief, such as lung cancer, any bias due to the healthy worker survivor effect

could be abolished by estimating the association of mortality with cumulative exposure lagged some ten years (that is, for an individual at risk at age  $t$ , any exposure received after age  $t - 10$  is ignored for the purposes of analysis).

In this monograph we test Gilbert's conjecture by reanalyzing data from a large mortality study of copper smelter workers exposed to arsenic, and show it may be incorrect. To stimulate the interest of the reader, we briefly summarize the results of our reanalysis. These results are described in detail in Secs. 11E and 12B. Among the subcohort of smelter workers hired prior to 1935, the lung cancer mortality rate was 3.54 times that of the general U.S. population. In addition, in this subcohort, mortality from lung cancer increased both with increasing cumulative exposure and with increasing cumulative exposure lagged 10–15 years. In the 1920s industrial hygiene controls were introduced and exposures to arsenic reduced. As a consequence, for cohort members hired after 1935, the lung cancer rate was only 1.4 times that of the general U.S. population. Presumably, this residual elevation in risk was either (a) at least partly related to the ongoing, although diminished, arsenic exposures in the smelter or (b) entirely due to the greater degree of cigarette smoking among cohort members than among the general U.S. population.

An intracohort analysis reported in Sec. 11E, restricted to cohort members hired after 1935, showed no association between lung cancer mortality and cumulative exposure to arsenic, even when cumulative exposure was lagged fifteen years ( $\chi^2_1 = 2.2$ ).

In contrast, when we reanalyzed the lung cancer mortality data of cohort members hired after 1935 using our statistical method, we found a marked association between lung cancer mortality and arsenic exposure ( $\chi^2_1 = 18$ ). We conjecture that there exists an adverse effect of arsenic on lung cancer mortality and that the absence of an association between cumulative exposure lagged fifteen years and lung cancer mortality is at least partly a consequence of the fact that smokers leave employment at a greater rate than nonsmokers. If so, smokers will tend to have low cumulative exposures even when lagged 15 years. Since information on cigarette smoking was not obtained from cohort members, any difference in cigarette smoking rates between individuals with high- and low-lagged cumulative exposures could not be adjusted for (nor can our conjecture be directly tested from this data, although, in Sec. 11, we present circumstantial evidence for it). In contrast to Gilbert's proposed method, the analytic method introduced in this paper remains unbiased when cigarette smoking is a determinant of employment status.

Since the post-1935 exposure levels found in the copper smelter are typical of the lower levels of exposure to adverse agents that are the rule in American industry today, it may prove to be generally important to analyze occupational cohort mortality data with the proposed method. In fact, our method may be necessary to control bias in any epidemiologic study in which risk factors are determinants of future exposure. We now proceed to a general outline of the monograph.

In cohort mortality studies in which individuals are exposed to the agent under study for sustained periods of time, independent risk factors for death commonly determine later exposure history. For example, in occupational cohorts, we observe that unexposed individuals who terminate employment at any age (say, 40) prior to age 65 have higher subsequent age-specific mortality rates than unexposed individuals who continue to work past that age (at least, in part, because of the healthy worker survivor effect). It follows that termination status is both a determinant of future exposure (since terminated individuals receive no more exposure) and an independent risk factor for death. As pointed out by Gilbert[1] and Robins[2, 3], if risk factors for death are determinants of subsequent exposure, the association of observed exposure history with mortality may fail to reflect a causal association. If, in addition, past exposure history is a determinant of subsequent risk factor status, the association of observed exposure history with mortality may be noncausal whether or not one adjusts for the risk factor.

The above results raise the question as to which, if any, population parameters can be given causal interpretations in observational studies with sustained exposure periods. This paper will answer that question.

In Sec. 2 we claim that an answer is offered by identifying an observational cohort study with a hypothetical double-blind randomized trial in which data on each subject's assigned treatment protocol has been erased from the data file. Understanding of causal inference in randomized clinical trials is well developed. Thus, we can apply our understanding of causal inference in randomized studies to observational studies.

In Sec. 3 we develop a formal theory of causal inference for observational studies with sustained exposure periods that does not rely on identifying the observational study with a hypothetical randomized trial. Our development in Sec. 3 relies heavily on graphical methods. Graphs, which we shall call *causally interpreted structured tree graphs*, are used to represent those population parameters with causal interpretations.

In Sec. 4 we provide the mathematical formalization of our theory.

In Sec. 5 we apply the approach developed in Secs. 3 and 4 to the estimation of the causal effect of arsenic exposure on total mortality in a cohort of copper smelter workers. We report a small adverse effect of arsenic exposure on survival. In contrast, a standard analysis, which does not control for the healthy worker survivor effect, finds no relationship between mortality and cumulative exposure.

Because our estimate of the causal effect of arsenic on mortality was obtained by fitting statistical models, it may be significantly biased if the models are incorrect (i.e. misspecified). To partly compensate for this, we develop in Sec. 6 a completely nonparametric test of the null hypothesis of no arsenic effect. For many occupational exposures, the central first question is whether exposure has any effect on mortality whatsoever. Fortunately we can often test thus null hypothesis without making any modelling assumptions. Our nonparametric test, in contrast to the standard analysis, suggests a statistically significant adverse effect of arsenic exposure on survival.

In Sec. 7, we show that the approaches to causal inference in observational studies of Secs. 2 and 3 are, in a certain well-defined sense, isomorphic.

In Sec. 8, we formally define the conditions under which the association of mortality with observed exposure history would be causal. In studies in which exposure and covariate status are measured only at start of follow-up, a covariate is defined to be a confounder only if the covariate is an independent risk factor for disease and is associated with exposure. We discuss how to generalize this definition of confounding to studies in which both the exposures and covariates are time dependent.

In cohort mortality studies, nested case control designs are frequently used to save computing and/or data acquisition costs. In Sec. 9 we show that such designs have previously unrecognized limitations when risk factors for death determine subsequent exposure. We show that one may be unable to test, using case-control data, the null hypothesis of no direct exposure effect controlling for cigarette smoking history when the healthy worker survivor effect is present.

If the healthy worker survivor effect is very weak or nonexistent, standard methods that estimate mortality as a function of observed exposure history will be nearly unbiased (and thus, the more complex methods outlined in this paper would not have to be employed). Thus, our first concern is to determine whether the healthy worker survivor effect is indeed operating. In Sec. 10, we discuss two common errors that can lead an investigator to believe that the healthy worker survivor effect is present when it is not.

Gilbert[1] claimed that, even when the healthy worker survivor effect was operative, if (1) the exposure of interest had a biological latent period of ten years from exposure to death and (2) the healthy worker survivor effect was operative for less than 10 years after an individual terminated work, then the association of mortality with observed ex-

posure history lagged 10 years would be causal. Since Gilbert's is the method most commonly used to attempt control of the healthy worker survivor effect, it is important to rigorously determine the conditions under which her approach is valid, and to develop empirical tests of whether these conditions hold. These issues are taken up in Sec. 11.

In Sec. 12, we consider the applicability of our methods to competing risks. We also consider their extension to mortality studies other than occupational studies and to outcomes other than mortality.

Finally Appendices A–G follow Sec. 12. These Appendices contain further generalizations and proofs of some theorems and lemmas.

## 2. OBSERVATIONAL STUDIES AS RANDOMIZED TRIALS MISSING DATA ON TREATMENT PROTOCOL

### A. Association of observed exposure history with mortality may be noncausal

Suppose Fig. 2.1 (where for now we ignore the symbols  $G_H$  and  $G_M$ ) represents the outcome of an occupational mortality study in which at time of hire,  $t_1$ , 400 individuals begin a high-exposure job. 200 of them leave work at  $t_2$ . Of those continuing at high exposure 60 die at  $t_3$ . 100 of those who left work also die. The bottom part of the graph is interpreted similarly. A possible state of nature consistent with this data is that neither exposure level nor employment status per se has an effect on any individual's outcome. Rather, high exposure is simply an irritant causing unhealthy individuals, many of whom are to die anyway, to leave work (i.e. the healthy worker survivor effect is operative). Suppose the difference between high- and moderate-exposure concentrations was less than the difference between moderate- and zero-exposure concentrations, so that the cumulative exposure at end of follow-up will be greater in individuals who remained at moderate exposure throughout the study than in individuals who received high exposure until  $t_2$  and received no exposure while off work thereafter. The study is analyzed in Table 1.

In Table 1, the observed exposure histories are in order of increasing levels of cumulative exposure. Suppose an investigator hypothesized that the effect of exposure, if any, would depend only on lifetime cumulative exposure. Such an investigator would

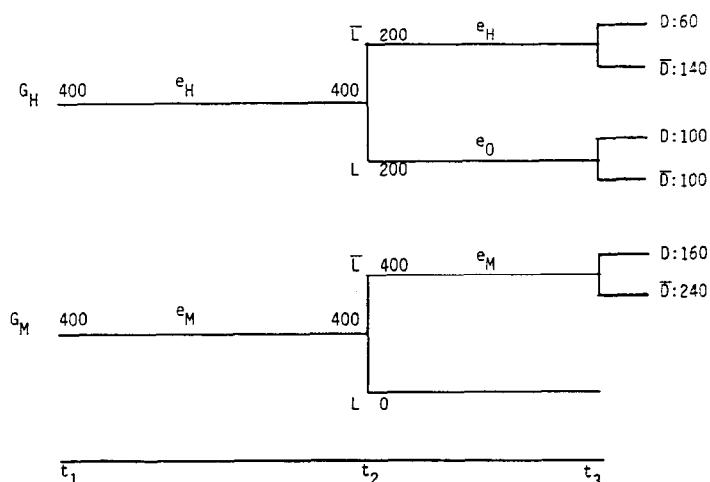


Fig. 2.1. A cohort study. ( $G_H$ : high PEH,  $G_M$ : moderate PEH,  $L$ : terminated at  $t_2$ ,  $D$ : died,  $e_H$ : high observed exposure concentration,  $e_M$ : moderate observed exposure concentration,  $e_0$ : zero observed exposure concentration).

Table 1. Analysis of the cohort study of Fig. 2.1

	HOEP	MOEP	LOEP	$G_H$	$G_M$
$D$	60	160	100	160	160
Total	200	400	200	400	400
	.3	.4	.5		

HOEP: the OEP with highest cumulative exposure (that of individuals remaining on  $G_H$ ).

MOEP: the OEP with moderate cumulative exposure (that of individuals remaining on  $G_M$ ).

LOEP: the OEP with the lowest cumulative exposure (that of individuals leaving  $G_H$  at  $t_2$ ).

falsely conclude that higher levels of cumulative exposure were protective. Although we might hope the association of mortality with observed exposure history controlling for termination status is causal, we see from Table 1 that there remains an apparent protective effect of high cumulative exposure compared to moderate cumulative exposure in the stratum of individuals who never left work.

In a randomized trial, each individual is randomly assigned at start of follow-up to a treatment protocol which gives their planned (i.e. projected) exposure history (PEH) from start to end of follow-up. Because individuals may deviate from their assigned treatment, their observed exposure history may differ from their PEH. If date of termination of protocol is a predictor of mortality, the association of mortality with treatment protocol (i.e. PEH), and not with observed exposure history (i.e. OEH), is a valid causal comparison. This is the basis of the well-known "intention to treat principle" for the analysis of randomized studies in which survival-treatment protocol associations are reported as the causal parameters. To see that the OEH-mortality association may be noncausal in a randomized trial, suppose we had described Fig. 2.1 as a randomized trial in which individuals had been assigned either to a PEH of high exposure throughout the trial ( $G_H$ ) or to a PEH of medium exposure through the trial ( $G_M$ ). Date of leaving employment then becomes date of leaving treatment protocol. In Table 1 we see that the association of mortality with treatment protocol correctly shows no treatment effect. Unfortunately, in an observational study, no subjects have a PEH, since no investigator assigned them to a treatment protocol. In the next subsection we develop a method of circumventing this difficulty.

#### B. Identifying an observational study as a randomized study with data on treatment protocol missing

In Sec. 2C, we will show that in a double-blind randomized trial with data on treatment protocol missing, if data on date of death, observed exposure history and date of leaving treatment protocol are available, mortality as a function of (the missing) treatment protocol can still be estimated. If we could identify an observational study with such a randomized trial, we could draw valid causal inferences by comparing mortality as a function of treatment protocol. In an observational study, data on observed exposure history and date of death are available. Thus it only remains necessary to identify some measured variable with time of termination of protocol. Now, in a double-blind randomized trial, termination of protocol is the first event which may be both a risk factor for death and a determinant of subsequent exposure (since, prior to termination of protocol, subsequent exposures had been determined by a flip of the coin at time of randomization, and not by any risk factors for death). In an observational study, an investigator must subjectively decide which variable will represent the "hypothetical" date of leaving protocol. Suppose we

are willing to assume that, conditional on past exposure, exposure (i.e. job) assignment while at work is unrelated to unmeasured risk factors for death. It then follows from the definition of the healthy worker survivor effect that date of termination of employment should be identified as date of termination of protocol, as it is the first event that both determines future exposure, and is a risk factor for death. We would thus have identified our observational mortality study with a randomized trial with data on treatment protocol missing. (In the mining industry, workers in ill health are selectively transferred to surface jobs. In that case, it follows from the above argument that we would identify date of termination of protocol as the earliest of date of termination of employment and date of transfer to a surface job.)

*Technical note.* The above argument only holds if no individual ever leaves and later returns to work (as would happen if there are individuals on layoff or on disability leave). If individuals leave and later return to work it is unclear how to define date of termination of employment. As we discuss in Sec. 10, it is inappropriate to define date of termination of employment as date of last employment, since death itself can determine date of last employment for individuals on layoff. Rather, if one believes that exposure is received essentially at random at work, conditional on past employment and exposure history, we could view the observational cohort study as a randomized trial in which individuals leave and return to protocol repeatedly. When individuals are at work we consider them to be on protocol, and when out of work to be off protocol. To avoid this added complexity in our early discussions, we suppose that no individual returns to work after he leaves. In Sec. 7 we return to the general case.

### C. Estimating mortality-PEH associations in a double-blind randomized trial with data on PEH missing

This section is somewhat technical. In a randomized trial missing data on treatment assignment we let  $g(t)$  be a subject's unknown planned exposure concentration at time  $t$  and  $e(t)$  a subject's observed exposure concentration. Define  $G(t) = \{g(u); u \leq t\}$  to be the PEH up to  $t$  and  $E(t) = \{e(u); u < t\}$  is the OEH up to  $t$ . Let  $L$  and  $D$  be variables that record the time the individual left treatment protocol and died, respectively. Let  $G$  be an entire PEH defined until end of follow-up. Let  $G \equiv 0$  be the  $G$  with planned exposures being identically zero throughout.

We suppose throughout the remainder of this subsection that:

2C.1. Treatments differ only through exposure.

2C.2. An individual who leaves protocol does so forever (so that  $L$  is well defined).

2C.3. Treatment protocols do not depend on time-dependent covariates measured after time of randomization,  $t_2$  (that is, treatment protocols such as the following are not allowed: continue on a high exposure until  $t_1$ ; if the subject's white count measured at  $t_2$  is greater than 1000, continue on high exposure; if less than 1000, cease further exposure). Note that 2C.1 and 2C.3 are necessary if a treatment protocol is to be uniquely characterized by a single planned exposure history. We drop these restrictions in Sec. 4.

Suppose (as is true in a double-blind randomized trial) that no individual's time of termination of protocol or death is influenced by any planned future exposure that has not or will not be experienced. If so, the following nonidentifiable temporal assumptions would hold for each individual  $i$ .

$$\gamma_L(t | E(t), G, i) = \gamma_L(t | E(t), i) \quad \text{for all } G \text{ with } G(t) = E(t) \quad (2.1)$$

$$\gamma_D(t | E(t_2), L = t_2, G, i) = \gamma_D(t | E(t_2), L = t_2, i) \quad \text{for all } G \text{ with } G(t_2) = E(t_2) \quad (2.2)$$

$$\gamma_D(t | E(t), L > t, G, i) = \gamma_D(t | E(t), L > t, i) \quad \text{for all } G \text{ with } G(t) = E(t) \quad (2.3)$$

where, e.g.  $\gamma_L(t | E(t), i)$  is the incidence of termination of protocol for individual  $i$  at time  $t$  conditional on being alive and on protocol at  $t$  and on having observed exposure history  $E(t)$ , and  $\gamma_D(t | E(t_2), L = t_2, i)$  is the incidence of death at time  $t$  conditional on OEH,  $E(t_2)$ , up to time of termination  $t_2$ .

Note  $E(t)$  must by definition equal  $G(t)$  for all  $t \leq L$  since the individual is still on protocol. The nonidentifiable temporal assumptions are not empirically testable from data on  $G$ ,  $E(t)$ ,  $L$ , and  $D$  (thus, the name "nonidentifiable"). Nonetheless, if in Eqs. (2.1)–(2.3), we do not condition on  $i$ , the resulting population relationships are empirically testable. We call Eqs. (2.1)–(2.3) without conditioning on  $i$  the *identifiable temporal assumptions*. The truth of the nonidentifiable temporal assumptions in a randomized trial implies the truth of the identifiable temporal assumptions (see Lemma B1 in Appendix B). It is easy to find examples to show that the converse is not true.

In a randomized trial with a sustained treatment period it is often supposed, for individuals on protocol, that the exposure received at time  $t$  conditional on past exposure up to  $t$  is not influenced by unmeasured risk factors. In fact, this is in general true only if treatment protocol was assigned at random and the nonidentifiable temporal assumptions hold. More precisely,

**LEMMA 2.1.**  $p(G | i) = p(G)$  and Eqs. (2.1) and (2.3) imply

$$\begin{aligned} \lim_{\Delta t \rightarrow 0} p(E(t + \Delta t) | E(t), L > t + \Delta t, D > t + \Delta t, i) \\ = p(E(t + \Delta t) | E(t), L > (t + \Delta t), D > t + \Delta t) \end{aligned} \quad (2.4)$$

for any random person  $i$  in the population on protocol at  $t + \Delta t$  (irrespective of his risk factors). (We assume changes in exposure are discontinuous.)

*Proof.* Obviously, Eq. (2.4) is true if it is true when  $G$  replaces  $E(t + \Delta t)$  wherever  $E(t + \Delta t)$  occurs in Eq. (2.4). But Eq. (2.4) with  $G$  substituted is true by the argument given in Lemma B1 in Appendix B. It is easy to show by example that the randomization plus the identifiable temporal assumptions do not imply Eq. (2.4).

We prove the following theorem:

**THEOREM.** If the identifiable temporal assumptions hold, the probability of surviving a given number of years as a function of PEH can be consistently estimated in a randomized trial without data on death, OEH and date of termination of protocol are available (provided, of course, that the joint distribution of the observed data can be consistently estimated).

*Proof.* An individual assigned to a particular PEH  $G$  can survive to a time  $t_f$  in any of the following mutually exclusive ways: survive on protocol till  $t_f$  or for any  $t_2 < t_f$  survive on protocol till  $t_2$ , terminate protocol at  $t_2$ , and then survive to  $t_f$ . In symbols, with  $t_1$  as start of follow-up,

$$\begin{aligned} S(t_f | G) \equiv p(D > t_f | G) &= p(L > t_f | G) + \int_{t_1}^{t_f} p(D \\ &> t_f | L = t_2, G) \gamma_L(t_2 | G) p(L > t_2 | G) dt_2. \end{aligned} \quad (2.5)$$

But

$$\begin{aligned} p(D > t_f \mid L = t_2, G) &= \exp \left[ - \int_{t_2}^{t_f} \gamma_D(t \mid L = t_2, E(t_2), G) dt \right] \\ &= \exp \left[ - \int_{t_2}^{t_f} \gamma_D(t \mid L = t_2, E(t_2)) dt \right], \end{aligned} \quad (2.6)$$

$$\begin{aligned} p(L > t_2 \mid G) &= \exp \left[ - \int_{t_1}^{t_2} [\gamma_D(t \mid L > t, E(t), G) + \gamma_L(t \mid E(t), G)] dt \right] \\ &= \exp \left[ - \int_{t_1}^{t_2} [\gamma_D(t \mid L > t, E(t)) + \gamma_L(t \mid E(t))] dt \right], \end{aligned} \quad (2.7)$$

and

$$\gamma_L(t \mid G) = \gamma_L(t \mid E(t)), \quad (2.8)$$

where  $E(t) = G(t)$  if  $t \leq L$  (since the individual is still on protocol) and Eqs. (2.1)–(2.3), without conditioning on  $i$ , justify dropping  $G$  in (2.6)–(2.8). Substituting Eqs. (2.6)–(2.8) into Eq. (2.5) proves the theorem. In fact by Theorem B1 (see Appendix B) even if data on  $G$  was available it would be ignored in the analysis since it is not part of the sufficient statistic.

If no individual leaves protocol, i.e.  $E(t) = G(t)$  for all  $t$ , it follows from Eqs. (2.5)–(2.7) that for any  $G$

$$p(D > t_f \mid G) = \exp \left[ - \int_{t_1}^{t_f} \gamma_D(u \mid E(u)) du \right] \equiv S(t_f \mid E(t_f)). \quad (2.9)$$

Using (2.9), causal parameters of such a randomized trial can be defined in terms of OEH. If all treatment assignments that ever diverge from one another do so at time of randomization,  $p(D > t \mid G)$  can be estimated without data on  $G$  without performing the integrals in Eqs. (2.5)–(2.7). To see this, let  $e(t_1)$  be the initial exposure of some individual following randomization (we assume no one terminates protocol prior to receiving their first exposure). Only one PEH  $G$ , say  $G_{e(t_1)}$ , will have initial exposure  $e(t_1)$  and we can estimate  $p(D > t \mid G_{e(t_1)})$  simply by recording the death times of individuals who received  $e(t_1)$  without regard to data on exposure history past the initial exposure or on date of termination of protocol.

#### D. Alternative designed randomized trials

An observational study is not the only type of study with a sustained exposure period in which individuals do not have treatment protocols assigned at start of follow-up.

Consider the following alternative design for a randomized trial. Each day of the trial, individuals who remain on protocol are given at random a new exposure level for that day. The probability of receiving a particular exposure level on a given day can depend (only) on the exposures received on previous days. If the exposure received on day 1 does not completely determine future exposure, then no individual in the alternative randomized trial has a well-defined treatment assignment  $G$  at the time of initial randomization. Nonetheless, individuals are truly randomized in the sense that no individual's

exposure at day  $t$  while on protocol can, except for sampling variation, be associated with any unmeasured risk factors, conditional on past exposure history [i.e. Eq. (2.4) holds]. More precisely, Eq. (2.4) holds in an alternative designed trial in the limit as the time  $\Delta t$  between exposure assignments goes to zero.

If, in the alternative designed trial, data on observed exposure history, date of termination of protocol, and date of death are available, one can empirically estimate the incidence functions

$$\Gamma = [\gamma_L(t | E(t)), \gamma_D(t | L = t_2, E(t_2)), \gamma_D(t | L > t, E(t))].$$

In an alternative designed randomized trial in which Assumptions 2C.1–2C.2 hold there is a unique double-blind ordinary randomized trial missing data on treatment protocol (i.e. a trial in which a well-defined protocol is assigned to each subject at time of initial randomization) for which  $p[L_i, D_i, E[\min(L_i, D_i)] | i]$  is the same in the alternative designed trial and the double-blind ordinary randomized trial. That ordinary randomized trial is the trial in which

$$\lim_{\Delta t \rightarrow 0} p[G(t + \Delta t) | G(t)] = p[E(t + \Delta t) | E(t), L > t + \Delta t, D > t + \Delta t].$$

Its causal parameters can be estimated by using Eqs. (2.4)–(2.7) from the data obtained in the alternative designed trial. We define the causal parameters of the alternative designed trial to be the causal parameters of that ordinary randomized trial.

#### E. Comparison of survival curves versus incidence differences

Throughout this paper we shall make causal comparisons in terms of comparisons of survival curves rather than incidences (i.e. hazards or conditional survivals). Our choice reflects the fact that, even in a randomized trial where the treatment is given only once at start of follow-up, it is possible that each individual's life may be shortened by treatment and yet the incidence in the untreated group may exceed that in the treated group at certain times  $t$  (since the treated and untreated survivors at time  $t$  may no longer be comparable, due to selective survival).

### 3. GRAPHICAL APPROACH TO CAUSAL INFERENCE IN OBSERVATIONAL SURVIVAL STUDIES

Suppose an investigator has performed an observational cohort study to determine the effect of a particular exposure on total mortality. Information on covariate status (including an exposure of interest) and vital status has been obtained on each study subject at  $S + 1$  times  $t_1, \dots, t_{S+1}$  where  $t_1$  is start of follow-up. We set ourselves the following four tasks.

1. To graphically represent the observed study data. We shall call such graphs measured partially interpreted structured tree graphs (MPISTG).
2. To determine the causal parameters of the study and to represent these parameters by graphs that we shall call measured causally interpreted structured tree graphs (MCISTG). This task requires that we develop a formal theory of causal inference for observational studies with sustained exposure periods. We shall see that investigators may disagree on the causal parameters of the study.
3. To determine which of the causal parameters are of substantive interest, a decision which depends on the purpose and subject matter of the study.

4. To determine the subset of causal parameters that can be consistently estimated from the observed study data; to represent these parameters by graphs that we shall call fully randomized measured causally interpreted structured tree graphs (FR MCISTG); and to develop an algorithm to estimate the causal parameters of an FR MCISTG. We shall see that investigators may disagree as to which causal parameters can be consistently estimated even when they agree on the causal parameters of the study.

In Sec. 4 we shall provide formal mathematical notation for these four tasks. In this section we shall carry them out somewhat informally. As examples, we consider two different observational studies. The first is a standard "point exposure study" in which at start of follow-up each member of the population receives either high ( $H$ ) or zero ( $0$ ) exposure. No further data is obtained until end of follow-up at time  $t_2$ . At  $t_2$ , data on each individual's vital status is obtained.

The second study is an occupational cohort mortality study with a sustained exposure period in which a cohort of workers hired at (calendar) time  $t_1$  (and matched on age at hire) are followed for 40 years. Every six months, exposure concentration, employment status, and vital status are measured. Exposure concentration measured at time  $t_s$  is labelled  $H$  for high and  $0$  for zero. Employment status at  $t_s$  is  $\bar{l}$  if at work and  $l$  if out of work. An individual's measured past employment (equivalent work) and exposure history through  $t_s$  are labelled  $L(t_s)$  and  $E(t_s)$ , respectively. We assume that we do not record (measure) data on exposure concentration and employment status at times between  $t_s$  and  $t_{s+1}$ . Thus, an individual's measured "employment and exposure history" is but part of their total employment and exposure history. Whenever we refer to an individual's covariate history (e.g. exposure and employment history), we shall be referring to their measured covariate history.

#### A. Task 1: A graphical representation of the observed study data

Figures 3.1 and 3.2 are different graphs, each of which represents the data obtained in the point exposure study. Figures 3.3 and 3.4 are graphs representing the data in the first

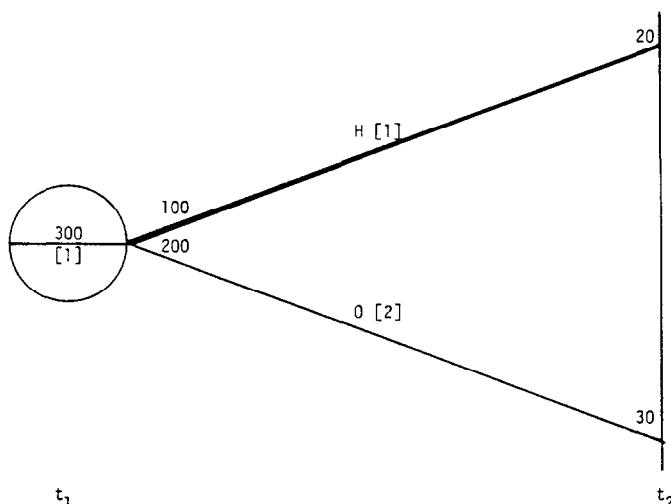


Fig. 3.1. An MPISTG of a point exposure study. Key for Figs. 3.1–3.4:  $H$  = high exposure concentration,  $0$  = unexposed,  $\bar{l}$  = at work,  $l$  = off work, whole numbers = numbers of subjects surviving at the given time with a given covariate history; numbers in [ ] are standard labels (see Sec. 4A of text), fractions in ( ) are conditional probabilities  $\gamma(\cdot i_s)$ ,  $\gamma(\cdot i_s j_s)$ ,  $S(\cdot i_s j_s)$  defined in Sec. 4B of text. Covariates on a given internodal line linking nodes at times  $t_s$  and  $t_{s+1}$  refer to measurements made at  $t_s$ . Highlighted subgraphs represent generalized treatments as defined in the text.

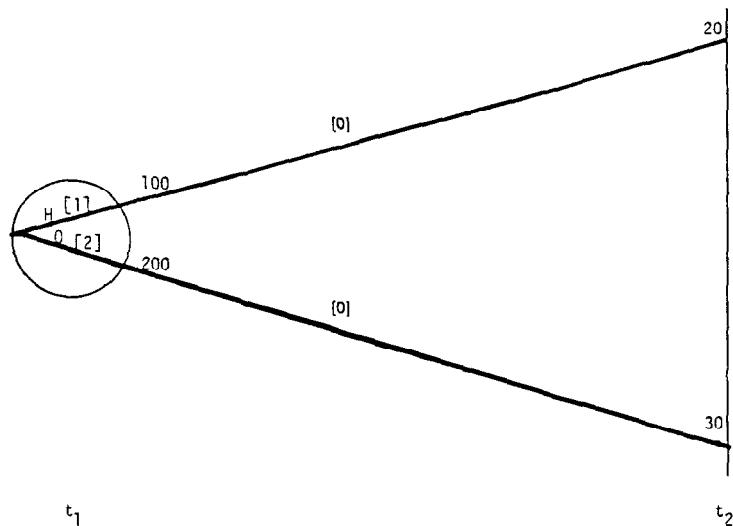


Fig. 3.2. An MPISTG of a point exposure study. (See Key for Fig. 3.1.)

18 months in the workplace study. All four graphs are examples of structured tree graphs (STG).

**Definition 3.1.** A structured tree graph has the following structure. One or more nodes (open circles) represent each time of observation  $t_i$ ,  $i \in (1, \dots, t_{s+1})$ . Successive observations are referred to as generations. Exactly one node is present at  $t_1$ . Each node (except that at  $t_1$ ) receives on its left circumference exactly one internodal line from a node in the previous generation. This line splits into one or more intranodal lines that traverse the interior of the circle and terminate at distinct points on the right circumference. From each of the above points on the right circumference, one or more internodal lines originate and extend to a node in the next generation. As in Fig. 3.1, the final node at  $t_{s+1}$  may be omitted.

An STG that represents the measured covariate data of a particular study we define to be a measured partially interpreted structured tree graph (MPISTG). The MPISTG of Fig. 3.3 is interpreted as follows (where, for the moment, the reader should ignore the highlighted lines, the numbers in square brackets, and the fractions in parentheses). At  $t_1$ , 300 individuals were at work of whom 100 received high exposure and 200 zero exposure. 75 of the 100 receiving high exposure at  $t_1$  survived until  $t_2$ , at which time 15 of the 75 remained at work at high exposure jobs, 30 remained at work at unexposed jobs, and 30 left work and thus were unexposed. Of the 30 who left work at  $t_2$ , 15 survived to  $t_3$ , at which point 6 returned to a high exposure job at work, 6 returned to an unexposed job and 3 remained off work and thus were unexposed. In summary, a particular intranodal line at  $t_s$  (and the corresponding point on the right circumference on which it terminates) represents a unique employment and exposure history through  $t_{s-1}$  plus survival to  $t_s$ . The number written over the intranodal line is the number of subjects with that covariate and survival history. An internodal line connecting a node at  $t_s$  to one at  $t_{s-1}$  represents a unique exposure and work history through  $t_s$ . The number written over the left end of that internodal line is the number of individuals who experienced that covariate history. Figure 3.4 is also an MPISTG of the workplace study. Figure 3.4 differs from Fig. 3.3 only in that, among individuals with identical work and exposure histories through  $t_{s-1}$ , in Fig. 3.4 (but not in Fig. 3.3) the subset of individuals off work at  $t_s$  and the subset of

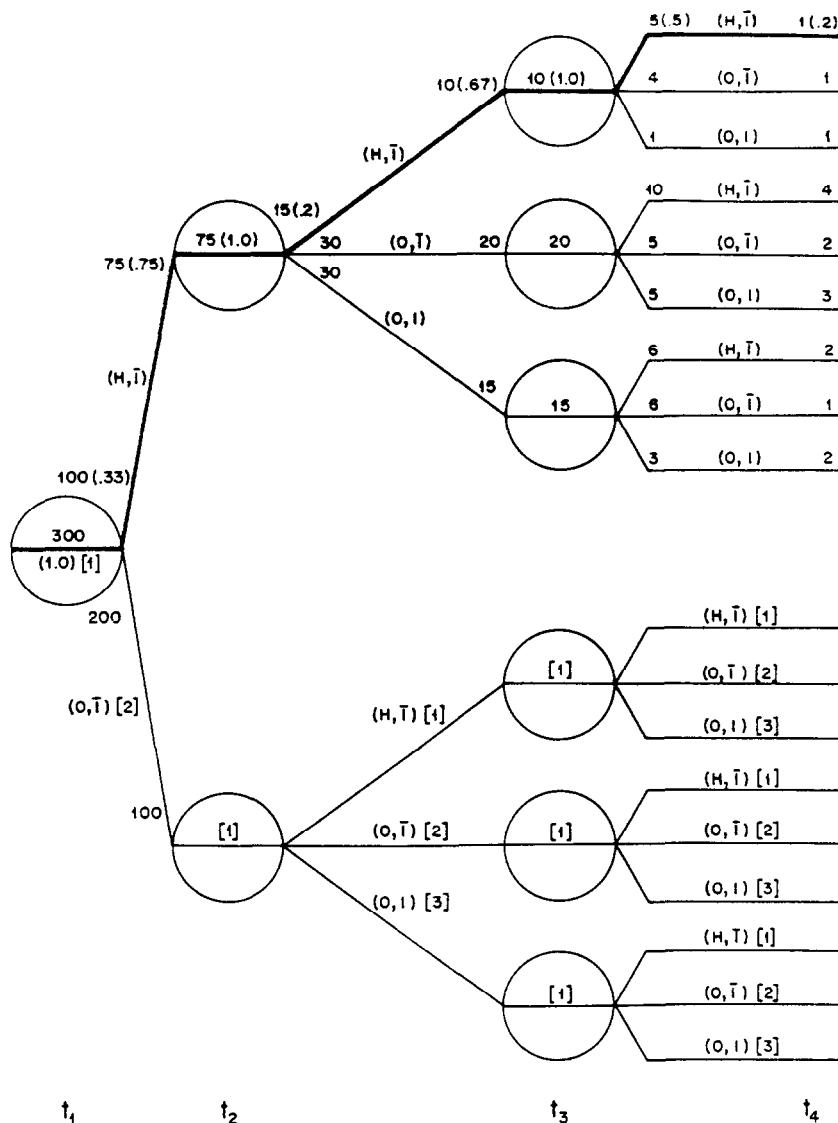


Fig. 3.3. An MPISTG. (See Key for Fig. 3.1.)

individuals at work at  $t_s$  are on distinct intranodal lines (and right circumference points). (Thus in Fig. 3.4 an intranodal line at  $t_s$  represents a unique history of exposure through  $t_{s-1}$  and employment through  $t_s$ .) The two MPISTGs have (1) the same number of nodes in each generation, (2) the same number of internodal lines arising from corresponding nodes, and (3) corresponding internodal lines representing the same covariate histories.

If two nonidentical MPISTGs have these three points in common (as do Figs. 3.3 and 3.4) and if one of the MPISTGs (e.g. Fig. 3.4) always has at least as many intranodal lines per node as the other MPISTG (e.g. Fig. 3.3) we say that the first MPISTG (e.g. Fig. 3.4) is coarser than the latter MPISTG (e.g. Fig. 3.3) and the latter is finer than the former. It will become clear later when we discuss causal parameters why Fig. 3.4 should be considered coarser than Fig. 3.3. At the moment it may seem counterintuitive.

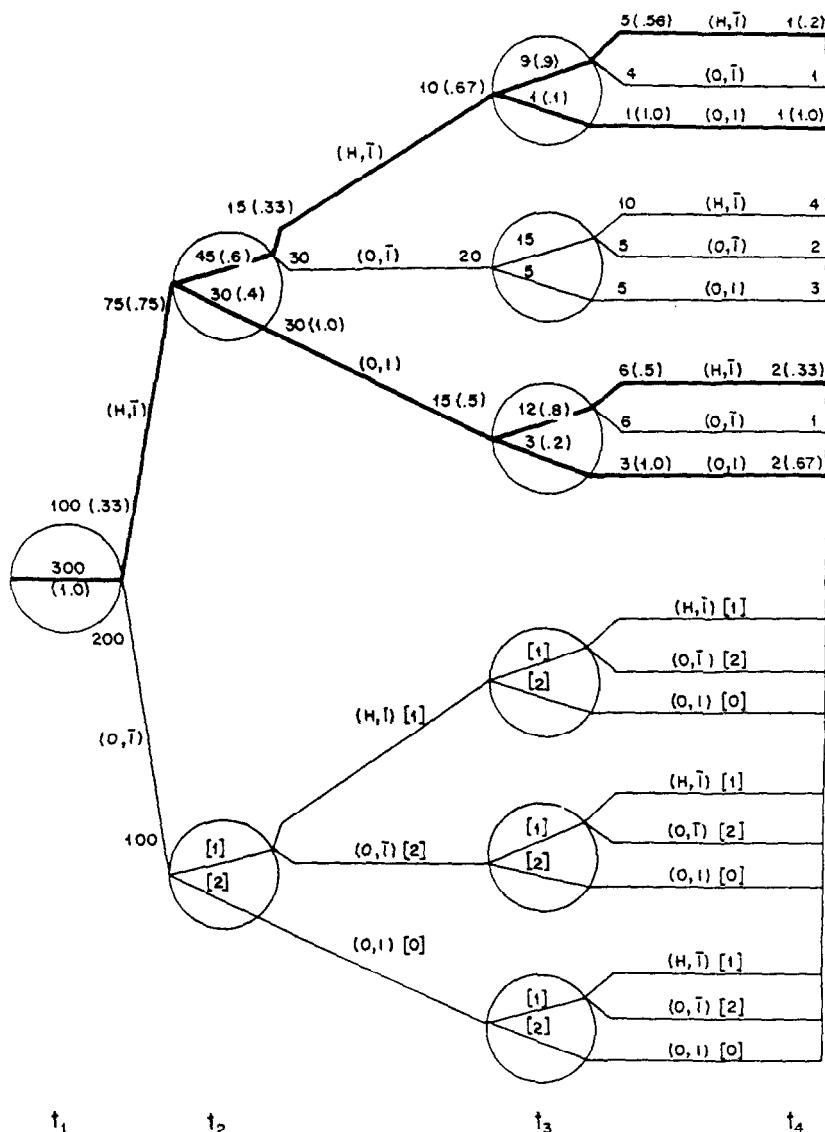


Fig. 3.4. An MPISTG. (See Key for Fig. 3.1.)

### B. Task 2A: Identifying the causal parameters of the study

An observational study differs from a controlled trial in that no investigator influenced the exposures of the study subjects. Rubin[4] has developed a theory of causal inference for point exposure observational studies. We shall extend his theory to include studies with sustained exposure periods. We shall need to define treatments, MCISTGs, and the causal parameters of a study. In our causal theory we shall suppose that nature (our name for the "actor" in an observational study) deterministically decides each individual's covariate and survival status at every time. In contrast Rosenbaum and Rubin[5] assume that in a point exposure study exposure status at  $t_1$  is assigned stochastically by nature.

*Definition of a treatment.* A particular (possibly joint) covariate level at time  $t_s$  (e.g. at work and receiving high exposure at work) is a treatment for a particular individual alive at  $t_s$  if (1) that individual could, at least conceptually, have either received or not

received that covariate level at  $t_s$  and (2) had the subject received that covariate level, the subject's subsequent covariate and survival history would have been well defined. A necessary condition for the existence of a treatment is that, at least conceptually, a controlled trial could have been performed in which at  $t_s$  an investigator superceded nature's deterministic choice of covariate level and gave the individual the covariate level under consideration. (We shall assume that the subsequent covariate and survival history for each individual when receiving a particular treatment at  $t_s$  is uninfluenced by the treatments received by any other individual at any time[4].)

**EXAMPLE.** If in the study represented by Fig. 3.1 high exposure were defined as *running up three flights of stairs*, high exposure would not be a treatment for individuals whose legs were paralyzed.

**Definition.** An MPISTG is an MCISTG if for any right circumference point (a) the covariate levels determining membership on the internodal lines arising from that point are each treatments for any individual with the covariate and survival history represented by the right circumference point, and (b) the subsequent covariate history for any such individual given one of these treatments can be represented by some path of intra- and internodal lines that lie on the MPISTG.

**EXAMPLE.** MPISTG 3.4 would not be an MCISTG if high exposure was defined as "running up three flights of stairs" and, e.g. at  $t_s$  there were individuals at work whose legs were paralyzed. In this paper henceforth we shall assume that exposure refers to exposure to an industrial chemical and that MPISTG 3.4 is an MCISTG. Furthermore, for convenience, we shall usually interpret MPISTG 3.3 to be an MCISTG by making the assumption that any individual off work at  $t_s$  could, conceptually, have been at work at either a high- or zero-exposure job. To satisfy (2) in the definition of a treatment we would need to carefully define what the treatment "being at work" entails for disabled workers, since obviously it cannot entail simply "being at work and performing one's usual task". One possibility is given in Sec. 3E.

### Task 2B. The causal parameters of an MCISTG

Following Rubin, in the point exposure study, we define the population causal effect of exposure on mortality to be the difference between the proportion surviving to  $t_2$  in a hypothetical controlled study in which the entire population received high exposure at  $t_1$  and the proportion surviving in a study in which all had received zero exposure at  $t_1$ . If survival status had been ascertained at multiple times (i.e. at  $t_2, t_3, \dots, t_{s-1}$ ) we would then have compared the proportion of the population surviving at each such time (i.e. we would have compared survival curves). See Sec. 8A for a further discussion.

A natural generalization of the above definition of a causal parameter that applies equally to studies with sustained exposure periods is the following.

**Definition.** A population causal parameter associated with a particular study population is the difference in the population survival curves of two different (usually hypothetical) studies, each with well-defined outcomes. To define the causal parameters of a study, we must determine how, on the basis of the observed study data, to characterize the set of hypothetical studies whose outcomes are believed to be well defined. To do so, we first show that each MCISTG has an associated set of generalized treatments each of which defines a unique (usually hypothetical) study with well-defined outcomes. A particular generalized treatment  $G^A$  of an MCISTG  $A$  is represented by a highlighted subgraph of the MCISTG constructed by use of the following algorithm:

*The Generalized Treatment Algorithm.* Beginning at the left circumference of the  $t_1$  node, highlight all intranodal lines in that node. At each point on the right circumference of the  $t_1$  node, highlight one internodal line. At  $t_2$  highlight all intranodal lines in the nodes on which the highlighted internodal lines originating from the  $t_1$  node terminated. From each point on the right circumference of a  $t_2$  node at which a highlighted intranodal line terminates, highlight one internodal line leading to a node in generation  $t_3$ . Continue in this manner through generation  $t_{S+1}$ .

The highlighted portions of MCISTGs 3.1–3.4 represent generalized treatments. To verbally characterize a particular generalized treatment, one must characterize the internodal line that is highlighted at each right circumference point at which two or more internodal lines were eligible for highlighting. For example, the generalized treatment shown in Fig. 3.4 can be characterized by “if at work at  $t_s$ , receive high exposure” and that of MCISTG 3.3 can be characterized by “if alive at  $t_s$ , remain at work and receive high exposure.”

The controlled study associated with the highlighted subgraph of MCISTG 3.4 would be as follows. At  $t_1$ , an investigator gives all individuals high exposure. Nature determines survival and employment status through  $t_2$ . For individuals at work at  $t_2$  in the hypothetical study, an investigator again gives each high exposure; nature then determines their survival and employment status through  $t_3$ . For individuals off work at  $t_2$ , nature gives each zero exposure at  $t_2$ , and then determines their survival and employment status through  $t_3$ . For those at work at  $t_3$ , an investigator gives each high exposure, etc.

Note that the essential difference between a point exposure and sustained exposure study is that, in a sustained exposure study, covariates measured at times after start of follow-up may also be treatments.

*Definition.* The population (individual)  $G$ -causal parameter comparing two generalized treatments  $G_1$  and  $G_2$  of an MCISTG [written for a particular individual  $i$ , as  $S(t, G_1, G_2, i)$  while for the population parameter, the small  $i$  is dropped] is the difference between the population (individual) survival curve of the hypothetical study defined by generalized treatment  $G_1$  [written as  $S(t | G_1)$  and  $(S(t | G_1, i))$ ] and the survival curve of a hypothetical study defined by generalized treatment  $G_2$ . We shall assume there exists a finest MCISTG (e.g. MCISTG 3.3 in our occupational mortality study) and that the causal parameters of the study are the set of  $G$ -causal parameters of the finest MCISTG and of all coarser MCISTGs. (Note it follows from the definition of an MCISTG that any MPSTG coarser than an MCISTG is itself an MCISTG.) See Sec. 4D for a more precise definition, and Sec. 4G for a caveat.

### C. Task 3: Determining causal parameters of substantive interest

The causal parameters of substantive interest to an investigator depend on the purpose and subject matter of the study.

**EXAMPLE.** Suppose Fig. 3.3 and thus Fig. 3.4 were both MCISTGs and exposure had no causal effect on any individual's mortality after controlling for employment history. This implies that the population survival curve of any two hypothetical controlled trials in which each individual is constrained to remain at work throughout the trial will be the same (irrespective of the exposures selected by the hypothetical investigators). The  $G$ -causal parameter of MCISTG 3.3 comparing these two studies, we define as a population causal effect of exposure controlling for employment history. We also suppose the  $G$ -causal parameter of MCISTG 3.3 representing the causal effect of employment controlling for exposure history demonstrates a strong adverse effect of unemployment on mortality (i.e. by definition, the proportion of study subjects surviving to  $t_s$  in a hypothetical study in which each individual receives zero exposure throughout and is off work from  $t_2$  on-

wards is less than in a study similar in all respects except that each individual continued at work). Then we say that employment history is a causal risk factor controlling for exposure. The adverse effect of unemployment would presumably have been mediated through (unmeasured) factors such as the loss of health insurance and increased financial and psychological stress. Suppose finally that the  $G$ -causal parameter of MCISTG 3.4 comparing the generalized treatment of "high exposure if at work" to "zero exposure if at work" demonstrates an adverse overall effect of high exposure because exposure functioned as an irritant that caused more highly exposed workers to leave work at a higher rate than less highly exposed workers.

An investigator whose purpose was to investigate the biological effect of exposure on mortality would wish to report the direct effect of exposure controlling for the intermediate variable, employment history. (Of course, finding that exposure had no effect on mortality controlling for employment history would not guarantee it had no direct biological effect on mortality since any biological effect may have been counterbalanced by exposure's effects on other unmeasured causal risk factors such as by causing exposed individuals to give up smoking.)

Suppose now that the study's subject matter (i.e. the data) was different. For example, suppose  $l$  in Figs. 3.3 and 3.4 represents cholesterol level instead of leaving work ( $l$  representing elevated cholesterol level, and  $\bar{l}$  representing low or normal cholesterol level). Suppose further that the investigator conceptualized cholesterol level as a treatment because he could imagine two drugs—one that could raise and the other that could lower cholesterol levels without influencing any other biochemical pathway. Such an investigator would believe that Fig. 3.3 was an MCISTG. Even so, the investigator might not wish to control for cholesterol level since the biological effect of exposure operates through its effect on the cholesterol level. In such a case we would report the  $G$ -causal parameters of MCISTG 3.4 in lieu of those of MCISTG 3.3 because the former represents the overall biological effect of exposure on mortality. (Since MCISTG 3.4 implies that individuals with elevated cholesterol levels never receive high exposure, it would be more accurate to say that the  $G$ -causal parameters of the MCISTG shown in Fig. 8.1 in Sec. (8) would be of substantive interest to the investigator.)

The interest of public health officials may lie in reducing exposure-related excess mortality rather than in determining whether exposure has a biological effect on mortality. These officials would be interested in the survival curve of a hypothetical controlled trial in which the investigator determined just those treatments that could actually be controlled by public health regulation in the real world. For example, if a public health official believed that exposure concentration at work could be controlled by regulation, but that neither the rate of leaving work nor the social benefits available for unemployed individuals could be altered, interest might center on the reduction of mortality associated with the generalized treatment "zero exposure if at work" of MCISTG 3.4.

Technically, of course, without making further assumptions, the  $G$ -causal parameters of an MCISTG associated with a particular observed study do not generalize to other study populations or to other study settings in which nature may determine outcomes differently. For example, if the observed study of Fig. 3.4 had been conducted in an identical setting except with different unemployment benefits, the survival curve might have been quite different. The magnitude of this difference could not be empirically predicted from knowledge of the  $G$ -causal parameters of the observed study without further *a priori* assumptions about the causal mechanism.

#### D. Task 4: Determining the causal parameters that can be consistently estimated

Given an MCISTG, we now define a condition that allows us to compute from the observed study data, the survival curve of any hypothetical controlled trial defined by a

generalized treatment of that MCISTG or of any MCISTG coarser than the given MCISTG.

*Definition 3.2.* An MCISTG  $A$  is a fully randomized MCISTG (FR MCISTG) if and only if the subsets of the population represented by the internodal lines arising from any given right circumference point have exactly the same distribution of subsequent covariate and survival histories as one another in any hypothetical study defined by a generalized treatment of MCISTG  $A$  (or another MCISTG coarser than  $A$ ) whose highlighted subgraph passes through that right circumference point. (Informally, the subsets are comparable in the sense that just before receiving their observed treatment at  $t_s$ , they were perfectly balanced with one another on all risk factors predicting future covariate and survival history.)

If an MCISTG  $A$  is an FR MCISTG, we can compute the probability of survival to any time  $t_s$  in the hypothetical study defined by a particular  $G_1$  by the following algorithm.

#### *G-Computation Algorithm*

1. On each intranodal line on the highlighted subgraph  $G_1$  write in parentheses (as we have done on the highlighted graphs in Figs. 3.3 and 3.4) the conditional probability of being in the subset in the observed study defined by that intranodal line, given that one is a member of the subset defined by the node in which the line lies. On the right end of each internodal line on the highlighted subgraph, write in parentheses the conditional probability of surviving to the next node (time) given that (in the observed study) one was in the subset represented by that internodal line.
2. For each highlighted path of intra- and internodal lines that connects a node at  $t_s$  to the left circumference of the  $t_1$  node form the product of the probabilities entered in step 1. (This product is the probability of surviving to  $t_s$  in the hypothetical study defined by  $G_1$  with the covariate history represented by that path of intra- and internodal lines, since, for an FR MCISTG, the conditional probabilities of the observed study defined in Step 1 are also the conditional probabilities of the hypothetical study). The sum of these products is the desired survival probability.

When MCISTG 3.4 is extended to include the entire 40 year follow-up period, there would be  $2^{80}$  terms in the final sum in Step 2 of the *G*-computation algorithm for  $t_{81}$ . Thus, the sum could not be evaluated even with the aid of a high-speed computer. Nonetheless, we can accurately evaluate the sum by using a Monte Carlo algorithm (see Sec. 4I).

We can show by example that we cannot compute the survival curve of the hypothetical study defined by any generalized treatment of an MCISTG that is not an FR MCISTG. Suppose that MCISTG 3.3 was not an FR MCISTG because, among individuals with high exposure at work through  $t_{s-1}$ , those off work at  $t_s$  were less healthy on the average than those individuals still at work at  $t_s$  (i.e. the healthy worker survivor effect is operating). Then, for the group of individuals who were off work at  $t_s$  (in the observed study) we could not compute what their mortality would be had they remained at high exposure jobs (since there is no comparable group of individuals in the observed study who remained at work at high exposure). Thus, we could not estimate the survival curve of the entire population in the hypothetical study defined by the generalized treatment of MCISTG 3.3, "if alive, receive high exposure at work".

In an observational study, no amount of empirical evidence can determine whether an MPISTG is an FR MCISTG (or for that matter whether it is an MCISTG). The assumption that an MPISTG is an FR MCISTG is always subjective. Note that if an MCISTG is an FR MCISTG we can compute exactly the value of the *G*-causal parameters of the MCISTG

(as well as those of any coarser MCISTG). We do not have to estimate these parameters. We know them exactly. Unfortunately, it is unlikely that nature would ever perfectly balance risk factors across the subsets of the population represented by the internodal lines arising from a single right circumference point. As such, no investigator would believe that any MCISTG was exactly an FR MCISTG. Nevertheless, an investigator might be willing to make subjective statements such as, "Although there may be some small association of unmeasured risk factors with treatment at each right circumference point, I do not believe such associations are systematic." We shall give the following formal structure to such subjective statements.

In an observational study, we shall assume that the observed study population has been randomly sampled from a near infinite hypothetical superpopulation. The causal parameters of interest will be those of the superpopulation. We now define an MCISTG  $B$  to be an FR MCISTG if Def. 3.2 holds in the superpopulation. When we subjectively believe that nature did not systematically give treatments to any subset of the population defined by a right circumference point from which two or more internodal lines arise on the basis of unmeasured risk factors for future covariate and survival history, we shall assume that the MCISTG representing the superpopulation is an FR MCISTG. Thus, even for an FR MCISTG, a chance association of treatments with unmeasured risk factors may exist in the observed study population due to sampling variability.

*Remark.* Since we are sampling from a near-infinite superpopulation, we could redefine an MCISTG so that each individual's outcomes (i.e., covariate and vital status history) were stochastic without changing any of our results except for the following technical philosophical difficulty. The controlled trial associated with a generalized treatment (e.g., if at work, receive high exposure) is of the general form—if an individual would be in state  $a$  (e.g., at work at low exposure) at  $t_s$ , put him in state  $b$  at  $t_s$  (e.g., at work at high exposure). Such a trial is, in principle, impossible to precisely implement without precognition, since if we must wait until  $t_s$  to discover  $a$  has occurred, then  $b$  cannot occur at  $t_s$ . Only a deterministic world allows (in principle) for precognition. Obviously, a stochastic model would be adequate if one does not mind being infinitesimally late in placing our subject in state  $b$ .

Unless stated otherwise, we shall hereafter assume that MCISTG 3.4 is an FR MCISTG. We shall use the following equivalent terminologies to characterize the fact that MCISTG 3.4 is an FR MCISTG. Conditional on past employment and exposure history, (1) subgroups of the observed study population receiving high exposure at work and zero exposure at work at  $t_s$  are randomized with respect to one another at that time, (2) subgroups receiving high exposure at work and zero exposure at work at  $t_s$  are comparable at  $t_s$ , (3) exposure at work was received at random at  $t_s$ .

The nonparametric maximum likelihood estimator (NPMLE) of the superpopulation survival curve associated with a given generalized treatment of FR MCISTG 3.4 is obtained by first estimating the unknown superpopulation proportions (conditional probabilities) necessary to apply the  $G$ -computation algorithm by the corresponding sample proportions and then applying the algorithm with the sample proportions in place of the superpopulation proportions. Unfortunately, in occupational mortality studies the typical cohort size is approximately 10,000 members. Thus, we cannot hope to consistently estimate the  $2^{80} + 2^{79} - 2$  conditional probabilities associated with the intra- and internodal lines of the highlighted subgraph of FR MCISTG 3.4 (when extended to  $t_{81}$ ) without specifying parsimonious parametric or semiparametric (e.g. Cox) models. A worked example is given in Sec. 5.

Even if we believe that exposure at work at  $t_s$  is received at random conditional on past true exposure and covariate history, we would not necessarily believe that it is received at random conditional on measured past exposure and covariate history when the time between measurements is long. This is because individuals who receive high exposure at  $t_s$  are likely to have received high exposure for some period between  $t_{s-1}$  and  $t_s$  and those receiving zero exposure at  $t_s$  are likely to have received zero exposure for some period between  $t_{s-1}$  and  $t_s$ . As such, the two groups of individuals will not be comparable at  $t_s$  if exposure has an adverse effect on mortality, since they will have different cumulative exposures at  $t_s$ . Even when exposure has no effect on mortality, if high exposure preferentially makes individuals with poor prognosis terminate employment, the same lack of comparability will arise. The only solution is to take measurements at short time intervals. We believe that in occupational mortality studies six months is a short enough period.

*Remark.* In the mining industry, workers in ill health are selectively transferred to unexposed surface jobs. Thus, if our study was of the mining industry, the MCISTG 3.4 would not be a FR MCISTG.

#### E. Estimable causal parameters and some caveats

Suppose one correctly believed that MPISTG 3.3 was an FR MCISTG. That is, among individuals with identical past work and exposure histories at  $t_s$ , the groups "individuals off work," "individuals at high exposure job," and "individuals at work at zero exposure job" do not systematically differ on unmeasured risk factors. Then it follows from the  $G$ -computation algorithm that we can test the null hypothesis of no exposure effect controlling for employment history by determining whether the mortality rate at each time  $t$  depends on observed exposure history when controlling for work history. Suppose, in

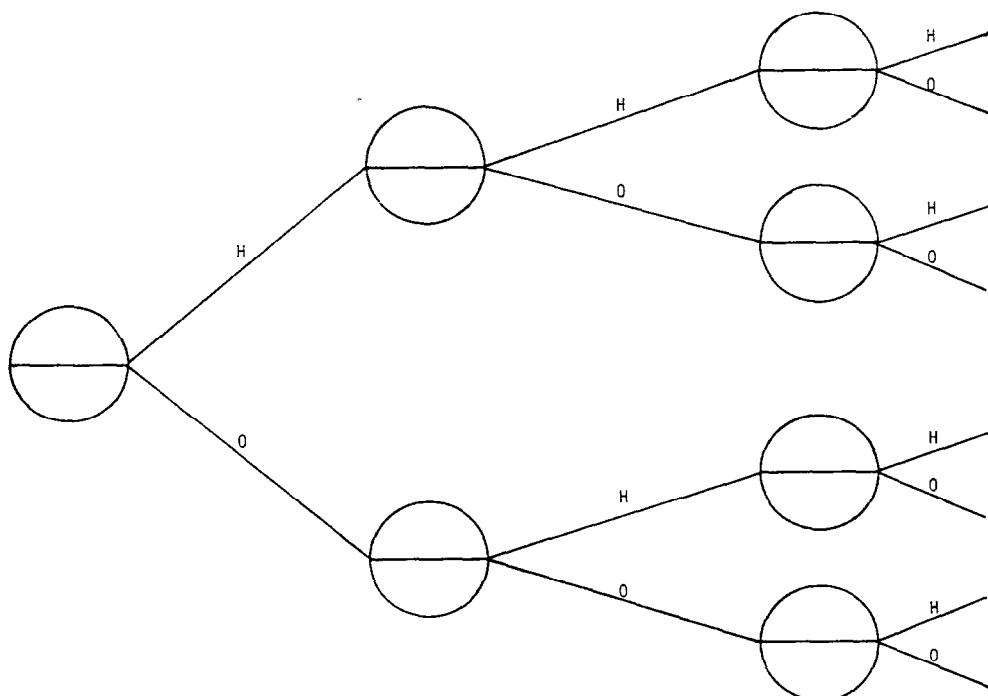


Fig. 3.5. An MPISTG.

addition, that employment history has no effect on an individual's mortality when controlling for exposure history (i.e. loss of health insurance and increased poverty associated with unemployment do not result in increased mortality). Then MPISTG 3.5 would be an FR MCISTG as well (see Theorem 8.1). It follows that a test of the null hypothesis of no causal effect of exposure on overall mortality (or mortality controlling for employment history) is obtained by determining whether the mortality rate at  $t$  depends on observed exposure history (even when we do not control for employment history). Thus, the standard practice in occupational epidemiology is to implicitly assume that MCISTG 3.5 is an FR MCISTG.

Suppose now that due to the fact that disabled individuals leave employment, the investigator correctly believes that MCISTG 3.4, but not MCISTG 3.3, is an FR MCISTG. Then the  $G$ -causal parameters of MCISTG 3.3 are not consistently estimable. In addition, the bias associated with estimating MCISTG 3.3's  $G$ -causal parameters by the  $G$ -computation algorithm applied to the sample data cannot be empirically estimated. The investigator can only estimate the overall effect of exposure when not controlling for employment history from FR MCISTG 3.4. We now examine what one can say, even qualitatively, about the causal parameters of MCISTG 3.3 representing the direct effect of exposure after controlling for employment history from knowledge of the  $G$ -causal parameters of FR MCISTG 3.4. To do so, we shall require the following definitions.

We say an exposure is semimonotone deleterious (beneficial) if there is no individual whose survival would be enhanced (diminished) by having his exposure increased at any time, but there is some individual whose survival is diminished (enhanced) by increasing exposure. That is, "more" is no better (worse) for anyone but is worse (better) for someone. Formally,

*Definition.* Exposure is semimonotone deleterious if for all  $t_k$ ,  $S(t_k, G_1^{3,3}, G_2^{3,3}, i) \leq 0$  (with strict inequality for some  $t_k, i, G_1^{3,3}, G_2^{3,3}$ ) whenever  $G_1^{3,3}, G_2^{3,3}$  are characterized by "stay at work, receiving exposures  $e_1(t_s)$  and  $e_2(t_s)$ ," respectively, such that  $e_2(t_s) \leq e_1(t_s)$  for all  $t_s$  where we have used the following notational convention.

*Notational convention.* A numerical superscript associated with any  $G$  refers to the figure number of the MCISTG for which  $G$  is a generalized treatment.

*Definition.* An exposure is apparently semimonotone deleterious if the above definition holds except with 3.4 substituted for 3.3 and "stay at work" replaced by "if at work" and  $i$  dropped (i.e. this is a population definition). Corresponding definitions of semimonotone beneficial are constructed simply by changing the direction of the inequality sign between  $e_2(t_s)$  and  $e_1(t_s)$ .

When Fig. 3.4 is an FR MCISTG we can empirically determine whether exposure is apparently semimonotone deleterious but not whether it is truly semimonotone deleterious.

It is easy to show that when exposure is apparently semimonotone deleterious and employment history is a causal risk factor controlling for exposure history, it is possible that exposure is semimonotone beneficial, semimonotone deleterious, or has no effect on any individual's mortality controlling for employment history. Thus, one is in a bind if interested in the biological effect of exposure. The causal parameters of FR MCISTG 3.4 can be estimated, but they have little substantive interest. One cannot consistently estimate the causal parameters of MCISTG 3.3 that do have substantive interest.

Suppose now that the investigator correctly assumes *a priori* (based on no empirical evidence if MCISTG 3.3 is not an FR MCISTG) that there is no effect of employment history on any individual's mortality when controlling for exposure history. What can the

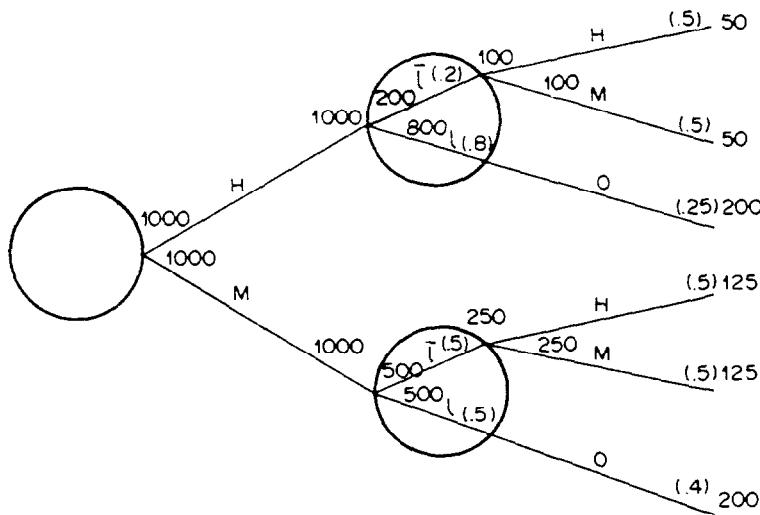


Fig. 3.6. An FR MCISTG: H = high exposure, M = medium exposure, O = zero exposure.

investigator now say concerning the probable biological effects of exposure on mortality from knowledge of the causal parameters of FR MCISTG 3.4?

We will show that it is possible (but unfortunately not identifiable) that even when employment history is not a causal risk factor, exposure can be apparently semimonotone deleterious and yet be semimonotone beneficial.

**EXAMPLE:** Consider a study represented by FR MCISTG 3.6. Exposure is apparently semimonotone deleterious, since

$$\begin{aligned} S(t_3 \mid G^{3.6} = ([H], [H])) &= S(t_3 \mid G^{3.6} = ([H], [M])) = .25 + .825 \\ &= .3 < .55 + .54 = .45 = S[t_3 \mid G^{3.6} = ([M], [H])] = S[t_3 \mid G^{3.6} = ([M], [M])] \end{aligned}$$

where  $([M], [H])$  means  $M$  at  $t_1$ ,  $H$  at  $t_2$ . In fact, exposure appears to be deleterious with a biological latent period of  $\Delta t = |t_2 - t_1| = |t_3 - t_2|$  years. We now show that there

Table 2. A possible state of nature generating the data in FR MCISTG 3.6: Survival experience of 6 homogeneous subgroups of study subjects when treated with various generalized treatments of a modified MCISTG 3.3

Generalized‡ treatment of modified MCISTG 3.3	Group 1†		Group 2		Group 3	
	Subgroup $A_1$	Subgroup $B_1$	Subgroup $A_2$	Subgroup $B_2$	Subgroup $A_3$	Subgroup $B_3$
$t_1$	$t_2$					
H,	0	0	0§	1§	0	0
M,	0	0	0	1	0	0
M,	H	1	0	1	1	0
M,	M	1	0	1	1	0
H,	H	1	0	1	1	1
H,	M	1	0	1	1	1
Number of each type	100	100	200	300	150	150

† Group definitions as given in Sec. 3E.

‡ Represents the generalized treatments (of the *modified* MCISTG 3.3 defined in Sec. 3E) defined by "stay at work and receive the exposures at  $t_1$  and  $t_2$  shown in the columns labelled  $t_1$  and  $t_2$ , respectively."

§ 1 indicates survival past  $t_3$ , 0 indicates death at  $t_3$ .

is a possible state of nature, consistent with the data in FR MCISTG 3.6, such that exposure is semimonotone beneficial.

Suppose the underlying state of nature generating the observed apparent semideleterious exposure effect is such that each exposure group at  $t_1$  consists of the following three distinct groups of individuals: 200 individuals who in the observed study are  $\bar{l}$  at  $t_2$  irrespective of initial exposure (Group 1); 500 individuals who are  $l$  irrespective of initial exposure (Group 2); 300 individuals who are  $l$  only if they received high exposure at  $t_1$  (Group 3). Furthermore, suppose each of these groups were itself a mixture of two subgroups (called  $A$  and  $B$ ) of truly homogeneous individuals with the survival histories shown in Table 2 when treated with generalized treatments of a modified version of MCISTG 3.3 that has 4 internodal lines labelled  $(0, \bar{l})$ ,  $(M, \bar{l})$ ,  $(H, \bar{l})$ ,  $(0, l)$  arising from each node. In Table 2 we can see that exposure is semi-monotone beneficial, since the definition is satisfied for each of the six homogeneous types of individuals. The reader can check that the state of nature depicted in Table 2 would generate the data in MPISTG 3.6 provided  $L$  is not a causal risk factor.

If there are zero exposure jobs in the workplace and  $L$  is not a causal risk factor apparent semimonotone deleterious exposures cannot be semi-monotone beneficial. To see this, in our example, suppose that all the  $M$ s were changed to 0s in Fig. 3.6. Then it is easy to see that exposure is still apparently semi-monotone deleterious but cannot be semi-monotone beneficial, since we have (without any further assumptions) that  $S[t_3 | G^{3,3}] = ([H, \bar{l}], [0, \bar{l}]) = .3 < S[t_3 | G^{3,3}] = ([0, \bar{l}], [0, \bar{l}]) = .45$ .

Fortunately, if neither exposure nor employment history are independent causal risk factors in MCISTG 3.3, then, as one would hope, all  $G$ -causal parameters of MCISTG 3.4 will be 0 at all times.

The above caveats notwithstanding, in this paper (unless stated otherwise) we shall implicitly assume that (1) the effect of employment history on mortality controlling for exposure history is negligible, (2) MCISTG 3.4 is an FR MCISTG, and (3) if the  $G$ -causal parameters of FR MCISTG 3.4 show deleterious (beneficial) overall effect of exposure on the population, then the true individual effects of exposure are monotone deleterious (beneficial). Finally for the sake of convenience of exposition and without loss of generality we shall usually assume that the MPISTG 3.3 is always an MCISTG by, for example, defining for a disabled individual the "treatment of high exposure at work" as follows. (1) The individual is brought to work (in a hospital bed if necessary). (2) The disabled individual is exposed to high levels of the air contaminant (we are assuming the exposure of interest is an air contaminant). (3) His salary and medical benefits continue as if he were actively employed. (4) He receives the standard level of medical care commensurate with his salary and benefits. We shall suppose that (1)–(4) are sufficient to give each disabled individual a well-defined outcome under the treatment of high exposure at work. There is no loss of generality in making the above definition of the treatment "high exposure at work" since because MCISTG 3.3 will not in general be an FR MCISTG we shall not be able to estimate the  $G$ -causal parameters of MCISTG 3.3. As such the inferences that we actually draw from the data will be unaffected by how we define "high exposure at work" for disabled individuals. See Sec. 8A.3 for a related discussion.

#### F. Contrast of our approach with a Bayesian approach

In this paper, we shall formally require the data analyst to make empirically untestable assumptions concerning the randomness of nature's treatment assignments. Such assumptions define the finest FR MCISTG (which, in our case, shall usually be FR MCISTG 3.4). (One could do a sensitivity analysis by analyzing the data under various assumptions as to the finest FR MCISTG.) The data analyst is to formally report "estimates" of the

$G$ -causal parameters of that FR MCISTG even though the analyst's real interest may lie in the  $G$ -causal parameters of a finer MCISTG (e.g. MCISTG 3.3) that is not an FR MCISTG. This approach is formal, and is not meant to preclude informal criticism. For example, even if the analyst treats MCISTG 3.4 as an FR MCISTG in his formal analysis, he may still subjectively believe that MCISTG 3.4 is not quite fully randomized. In such a case, it is perfectly sensible to informally discuss one's belief about the direction and magnitude of the bias of the estimator of the  $G$ -causal parameters of MCISTG 3.4 based on the  $G$ -computation algorithm. Furthermore, it makes sense to informally discuss one's beliefs about the magnitude of the nonidentifiable  $G$ -causal parameters of interest (i.e. those of MCISTG 3.3). Pratt and Schlaifler[6] make a similar point.

In contrast, a Bayesian interested in the causal parameters of MCISTG 3.3 would not be unduly concerned that nature had not given treatments at random. Rather, the Bayesian would simply use the observational data to update his beliefs about the parameters of interest, recognizing that these parameters could not be consistently estimated. Philosophically, we find the strict Bayesian view quite appealing. But it is impractical to implement due to problems with prior specification and computational intractability. Thus, we have proposed our less philosophically appealing but more reasonable approach to causal inference in observational studies.

#### *G. Empirical versus actual healthy worker survivor effect*

*Definition.* The "empirical healthy worker survivor effect" is operative at  $t_s$  if and only if among any set of individuals alive at  $t_s$  in the observed study with identical exposure and employment histories through  $t_{s-1}$ , the probability of survival to any time  $t(t > t_s)$  is greater in the group "at work at  $t_s$ " than in the group "off work at  $t_s$ " when all are treated, starting at  $t_s$ , with the generalized treatment "if at work receive zero exposure".

*Definition.* Employment history is a population risk factor for death controlling for exposure history if and only if the incidence of death at some time  $t$  differs among two groups with the same exposure histories but different employment histories.

*Remark.* Under the assumption that MPISTG 3.4 is an FR MCISTG one can empirically test whether the "empirical healthy worker survivor effect" is operative using the  $G$  computation algorithm beginning at  $t_s$  rather than at  $t_1$ . If employment history is not an independent population risk factor then the empirical healthy worker survivor effect cannot be operating.

*Definition.* We define a group of workers  $A$  to be less healthy than a group of workers  $B$ , both alive at  $t_s$ , for a particular exposure history subsequent to  $t_s$ , say  $\{e_1(t_k); t_k > t_s\}$ , if and only if the probability of survival through any  $t(t > t_s)$  for group  $A$  is less than that for group  $B$  in a hypothetical study defined by the generalized treatment of MCISTG 3.3 "if alive at  $t_k$  receive  $e_1(t_k)$  at work."

*Remark.* When MPISTG 3.3 is not an FR MCISTG and an "empirical healthy worker survivor effect" exists, we cannot empirically determine whether the observed effect is due to the group off work being less healthy than the group on work for the exposure history of "no subsequent exposure" (i.e., whether the healthy worker survivor effect is operative), whether it is due to an effect of the unemployed state per se (mediated, for example, by loss of health insurance), or whether it is due to both.

**CLAIM.** Given that (1) MPISTG 3.4 (but not MPISTG 3.3) is an FR MCISTG and (2) there is no effect of employment history on an individual's mortality controlling for ex-

posure, we can empirically test whether, among workers with identical work and exposure histories through  $t_{s-1}$ , those off work at  $t_s$  are less healthy than those on work at  $t_s$  only for the exposure history of "no subsequent exposure."

*Proof.* We can only estimate the subsequent survival of the groups "on work at  $t_s$ " and "off work at  $t_s$ " (defined above) for generalized treatments of FR MCISTG 3.4. Consider, for example, the generalized treatment "if at work receive high exposure beginning at  $t_s$ ." Then the group "on work at  $t_s$ " will receive higher exposure at  $t_s$  than the group "off work at  $t_s$ ." Thus we cannot compare the survival of the two groups when given the same exposure history unless we treat them beginning at  $t_s$  with "if at work, receive zero exposure."

#### 4. A FORMAL THEORY OF CAUSAL INFERENCE

##### A. Structured tree graphs

An STG (see Def. 3.1) is said to be standardly labelled if (1) within each node, intranodal lines are numbered consecutively from 1; and (2) when two or more internodal lines originate from the same point on a right circumference, they are numbered consecutively beginning with 1, and otherwise, internodal lines are labelled with a 0. The standard numerical label of a particular intranodal or internodal line is read off the graph by recording the numbers in order on the unique path from the left circumference of  $t_1$  to the intranodal or internodal line of interest. Thus, a given intranodal line at  $t_s$  (and the point on the right circumference on which it terminates) can be identified as  $i_1 j_1 i_2 j_2 \dots i_{s-1} j_{s-1} i_s$  where we reserve the  $i$  symbol to represent labels of intranodal lines (and the associated point on the right circumference) and  $j$  to represent those of internodal lines (and the associated point on the left circumference of a node in the next generation). The internodal lines arising from  $\cdot i_s$  can be identified as  $\cdot i_s j_s$  (where we abbreviate the sequence  $i_1 j_1 \dots i_s j_s$  by  $\cdot i_s j_s$  when the meaning is clear). The node on which  $\cdot i_s j_s$  terminates is  $\cdot i_s j_s(t_{s+1})$ . In the above  $i_1 \in \{1, \dots, N\}; j_1(i_1) \in \{1, \dots, N_{i_1}\}$  if  $N_{i_1} > 1$  and  $j_1(i_1) = 0$  if  $N_{i_1} = 1$ ;  $i_s(\cdot i_{s-1} j_{s-1}) \in \{1, \dots, N_{\cdot i_{s-1} j_{s-1}}\}, j_s(\cdot i_s) \in \{1, \dots, N_{\cdot i_s}\}$  if  $N_{\cdot i_s} > 1$  and  $j_s(\cdot i_s) = 0$  if  $N_{\cdot i_s} = 1$ ; where, for example,  $N_{\cdot i_{s-1} j_{s-1}}$  is the number of intranodal lines arising from the node in generation  $t_s$  on which the internodal line  $\cdot i_{s-1} j_{s-1}$  terminates; and  $j_s(\cdot i_s)$  is the number associated with the particular internodal line  $\cdot i_s j_s$ .

**EXAMPLE.** The lower halves of STG 3.1–3.4 have their standard labels within square brackets.

We define a partial ordering among STG.

**Definition 4.1.** Coarseness. We define an STG  $B$  to be at least as coarse as STG  $A$ , and  $A$  at least as fine as  $B$ , if and only if  $B$  can be produced from  $A$  by the following algorithm. For each  $\cdot i_s$  in graph  $A$ , the set of internodal lines are divided into  $K(\cdot i_s)$  mutually exclusive subsets. Each such subset is assigned a separate point on the right circumference of its node (and thus a separate intranodal line). Note that any graph produced by this process from  $A$  is an STG with the same number of nodes and internodal lines per node as  $A$ . If  $K(\cdot i_s)$  is greater than 1 for some  $\cdot i_s$ , we say that the STG  $B$  is coarser than  $A$  ( $A$  is finer than  $B$ ). Formally, we can identify any partition of the internodal lines arising from an  $(\cdot i_s)$  with a particular mutually exclusive and exhaustive partition:

$$A_{\cdot i_s} = (A_{1(\cdot i_s)} \cdots A_{K(\cdot i_s)})$$

of the set  $\{1, \dots, N_{i_s}\}$  such that

$$A_{k(\cdot i_s)} \cap A_{k'(\cdot i_s)} = 0 \text{ if } k \neq k', U_k A_{k(\cdot i_s)} = \{1, \dots, N_{i_s}\}, k \in (1, \dots, K).$$

**Definition 4.2.** A generalized treatment of an STG. A subgraph of an STG  $A$  is a generalized treatment if it can be formed by the generalized treatment algorithm described under Task 2B of Sec. 3. A particular generalized treatment “ $G^A$ ” can be (minimally) specified, using the standard labelling, by the set  $\{j_s(\cdot i_s), j_s \neq 0\}$  of highlighted internodal lines excluding those highlighted lines standardly labelled 0. Let “ $G^A$ ” be the set of all “ $G^A$ . The quotes around “ $G$ ” are used to stress that this use of generalized treatment has no causal interpretation.

### B. Partially interpreted structured tree graphs

Consider an observational cohort survival study in which covariate and vital status data on each individual is collected at  $S + 1$  times  $t_1, \dots, t_{S+1}$ . Time may represent chronological age, time on test, calendar date, etc. The study population of size  $m$  is assumed to have been randomly sampled without replacement from a large superpopulation of size  $M$  and sampling was possibly conditional on time-dependent covariates  $Z(t_{1-}) = [Z_1(t_{1-}), \dots, Z_K(t_{1-})]$  where  $Z(t_{1-}) = \{z(\mu); \mu \leq t_{1-}\}$ ,  $z(\mu)$  is the actual value of a covariate measured at time  $\mu$ , and  $t_{1-}$  is a time infinitesimally less than  $t_1$ . Suppose that for any possibly time-dependent covariate  $Z(t)$ , the potentially available data are of the form  $Z(t_{S+1}) \equiv [Z(t_{1-}), z(t_1), z(t_2), \dots, z(t_{S+1})]$ . By letting  $\Delta t = (t_s - t_{s-1}) \rightarrow 0$  we can approximate continuous time. For notational convenience we assume  $\Delta t$  does not depend on  $s$  and denote  $Z(t_{1-})$  as  $Z(t_0)$ . We consider only the discrete time case both to keep the mathematics simple and to allow graphical representations. In Sec. 6C, we allow for left censoring. Until Sec. 12, right censoring and competing risks are absent.

**Definition 4.3.** An STG is a partially interpreted structured tree graph (PISTG) of a given study if the covariate and survival history of the superpopulation can be represented as follows. Each internodal line  $\cdot i_s j_s$  represents a distinct covariate and survival history defined by being alive at  $t_s$  with a particular value of the vector  $Z(t_s)$ .  $[\cdot i_s j_s]$  is the subset of the population with history  $\cdot i_s j_s$ . A subject has covariate and survival history  $\cdot i_s$  if and only if he has any of the  $N_{i_s}$  histories  $\cdot i_s j_s$ . Thus,  $[\cdot i_s]$  is the union of the  $[\cdot i_s j_s]$ .  $i_s j_s(t_{S+1})$  represents the history “alive at  $t_{S+1}$  with covariate history represented by  $\cdot i_s j_s$ .” The union of the  $N$  sets  $[\cdot i_1]$  is the entire superpopulation. The union of all sets  $[\cdot i_s]$  is the entire population alive at  $t_s$ .  $p(\cdot i_s j_s)$  is the proportion of the superpopulation with history  $\cdot i_s j_s$ . We will usually assume  $p(\cdot i_s j_s | \cdot i_s) > 0$  if  $p(\cdot i_s) > 0$ .

In general, we will encounter a single study whose outcomes were observed and want to use it to predict outcomes in a number of other hypothetical studies in whose outcomes we are interested. A PISTG that represents the data in the study that was actually observed is called an observed PISTG (OPISTG). An OPISTG is an MPISTG if data on the covariates necessary to determine the covariate histories  $\cdot i_s j_s$  of each sampled subject were recorded for data analysis. An MPISTG is *detailed as the data* if whenever two individuals differ on  $Z(t_s)$  they are associated with different internodal lines.

A partial interpretation of an STG  $A$  induces a natural partial interpretation on any STG  $B$  coarser than  $A$ .

For any PISTG  $A$  define

$$\gamma(\cdot i_s | A) \equiv p(\cdot i_s | \cdot i_{s-1} j_{s-1}(t_s), A) \quad (4.1)$$

$$\gamma(\cdot i_s j_s | A) \equiv p(\cdot i_s j_s | \cdot i_s, A) \quad (4.2)$$

$$S(\cdot i_s j_s | A) = p(D > t_{s+1} | D > t_s, \cdot i_s j_s, A) \quad (4.3)$$

$$\gamma_D(t_{s+1} | \cdot i_s j_s, A) \equiv 1 - S(\cdot i_s j_s | A). \quad (4.4)$$

**EXAMPLE.** In Fig. 3.4, the fractions written over the highlighted intranodal lines are  $\gamma(\cdot i_s)$ , the fractions written over the left end of each highlighted internodal line are  $\gamma(\cdot i_s j_s)$ , and the fractions written over the right end of each highlighted internodal line are  $S(\cdot i_s j_s)$  (ignoring sampling variability).

### C. Causally interpreted structured tree graphs (CISTG)

In Sec. 4 only we use primes to indicate outcomes determined by nature.

**Definition 4.4.** A PISTG is a CISTG if each individual in the superpopulation in any set  $[\cdot i'_s]$  has associated a deterministic set  $HT(\cdot i'_s) = \{(D(t_u), I(t_u), J(t_u)); u \in (s+1, \dots, S)\}$  defined as follows: If  $t_{s+1} \leq S$ ,  $(D(t_{s+1}), I(t_{s+1}), J(t_{s+1})) = (\{D_{\cdot i_s j_s}\}, \{\cdot i'_s j_s i'_{s+1}\}, \{\cdot i'_s j_s i'_{s+1} j'_{s+1}\})$ , where in all three sets  $j_s$  takes on each of the values  $(1, \dots, N_{\cdot i_s})$ .  $D_{\cdot i_s j_s} = 1$  if, when in  $\cdot i'_s j_s$ , the individual would die in the interval  $(t_s, t_{s+1}]$  and  $D_{\cdot i_s j_s} = 0$  if the individual would survive past  $t_{s+1}$ .  $\cdot i'_s j_s i'_{s+1}$  and  $\cdot i'_s j_s i'_{s+1} j'_{s+1}$  are the subsets in which nature would deterministically place the individual at  $t_{s+1}$  if at  $t_s$  the individual had been in  $\cdot i'_s j_s$  in a hypothetical study (rather than in  $\cdot i'_s j'_s$ ). [We suppose that nature determines these sets even if the individual died in  $(t_s, t_{s+1}]$ . This latter assumption is used to extend our results to competing risks in Sec. 12. Until Sec. 12 (with the exception of Sec. 8B), all our results will hold without this assumption.] If alive at  $t_{s+1}$  an individual would then be in subset  $\cdot i'_s j'_s i'_{s+1} j'_{s+1}$  in the actual study.

If  $t_{s+2} \leq S$ ,

$$(D(t_{s+2}), I(t_{s+2}), J(t_{s+2})) = (\{D_{\cdot i_s j_s i'_{s+1} j_{s+1}}\}, \{\cdot i'_s j_s i'_{s+1} j_{s+1} i'_{s+2}\}, \{\cdot i'_s j_s i'_{s+1} j_{s+1} i'_{s+2} j'_{s+2}\})$$

where in all three sets  $j_s$  takes on each of the values of  $(1, \dots, N_{\cdot i_s})$  and  $j_{s+1}$  takes on each of the values  $(1, \dots, N_{\cdot i_s j_s i'_{s+1}})$ ;  $\cdot i'_s j'_s i'_{s+1} j_{s+1} i'_{s+2}$  and  $\cdot i'_s j_s i'_{s+1} j_{s+1} i'_{s+2} j'_{s+2}$  are subsets in which nature would place the individual at  $t_{s+2}$  if the individual had been in  $\cdot i'_s j_s i'_{s+1} j_{s+1}$  at  $t_{s+1}$ .  $D_{\cdot i_s j_s i'_{s+1} j_{s+1}} = 1$  if  $D_{\cdot i_s j_s} = 1$  or the individual would die in the interval  $(t_{s+1}, t_{s+2}]$  when in  $\cdot i'_s j_s i'_{s+1} j_{s+1}$ ; and  $D_{\cdot i_s j_s i'_{s+1} j_{s+1}} = 0$  if  $D_{\cdot i_s j_s} = 0$  and the individual would survive past  $t_{s+2}$  when in  $\cdot i'_s j_s i'_{s+1} j_{s+1}$ . We continue in this fashion incrementing by one until  $t_{s+k} = t_S$ .

If a PISTG is an CISTG, the covariate and death history for any individual are uniquely determined through  $t_{s+1}$  when a hypothetical investigator specifies the treatments,  $j_s(\cdot i_s)$ , for any subset of the times  $t_1$  through  $t_S$  and allows nature to select the treatment at all other times. OPISTGs and MPISTGs that are CISTGs, we call OCISTGs and MCISTGs.

**LEMMA.** If PISTG  $A$  is an CISTG, then PISTG  $B$  coarser than  $A$  is an CISTG.

**Proof.** This follows immediately from the existence of the sets  $J(t_u)$  defined above.

### D. The causal parameters of a CISTG

**Definition.** The generalized treatments of a CISTG. A general treatment " $G_1^A$ " of the STG of a CISTG  $A$  defines a unique hypothetical study  $G_1^A$  in which: (1) Each individual

with a particular history  $i_1$  is treated with unique treatment  $\cdot i_1 j_1$  on the highlighted graph " $G_1^A$ ". (2) If alive with history  $\cdot i_2$  at  $t_2$ , the subject receives the highlighted  $\cdot i_2 j_2$  of " $G_1^A$ ", etc. From the definition of a CISTG, each subject has a unique well-defined covariate and survival history in such a study. We describe this study by saying that each subject was treated with generalized treatment  $G_1^A$  (where we drop the quotes surrounding " $G_1^A$ " when referring to the hypothetical study defined by  $G_1^A$ ). Throughout this paper, when no ambiguity is likely to arise, we shall, in a slight abuse of notation, refer to the subgraph of the PISTG represented by " $G_1^A$ " as  $G_1^A$ .

We need to describe some notation. Given a CISTG  $A$ ,  $S(t | G_1^A)$  is the survival curve of the hypothetical study defined by  $G_1^A$ . For a particular  $\cdot i_s$  and  $\cdot i_s j_s$ ,  $\gamma(\cdot i_s j_s | G_1^A) \equiv \gamma(\cdot i_s j_s^A | G_1^A)$  is the conditional probability of having covariate history  $\cdot i_s j_s^A$  given one has covariate history  $\cdot i_s^A$  in the hypothetical study defined by  $G_1^A$ , where  $\cdot i_s j_s^A$  is the covariate history represented by the internodal line " $\cdot i_s j_s$ " in PISTG  $A$  (when PISTG  $A$  is standardly labelled). In equations that include both the study represented by  $A$  and the study represented by a particular  $G_1^A$  whenever a superscript is missing from any symbol the superscript is assumed to be an  $A$ . Also, in the same spirit,  $S(\cdot i_s j_s) \equiv S(\cdot i_s j_s | A)$ .

*Definition.* The set of  $G$ -causal parameters of a CISTG is the set  $\{S(t, G_1^A, G_2^A) \equiv p(D > t | G_1^A) - p(D > t | G_2^A) \equiv S(t | G_1^A) - S(t | G_2^A); G_1^A, G_2^A \in G^A\}$ . Let  $K^A$  represent a study in which each individual in the population is treated with a particular, possibly different,  $G_k^A \in G^A$ . Let  $K^A$  be the collection of all such  $K^A$ .

*Definition.* The set of population causal parameters of CISTG  $A$  is the set  $\{S(t, K_1^A, K_2^A) \equiv S(t | K_1^A) - S(t | K_2^A), K_1^A, K_2^A \in K^A\}$ .

**LEMMA.** If CISTG  $B$  is coarser than CISTG  $A$  then for all  $K_1^B, K_2^B, S(t, K_1^B, K_2^B) = S(t, K_1^A, K_2^A)$  for some  $K_1^A, K_2^A \in K^A$ .

In general, most of the population causal parameters of an CISTG  $A$  are nonidentifiable. We proceed to define the  $S(t, K_1^A, K_2^A)$  considered identifiable by a particular investigator.

### E. Fully randomized CISTG

Given a CISTG  $A$  for individuals in  $\cdot i_s'$  define  $H^A(\cdot i_s') = \{D(t_u), I(t_u), u \in (s+1, \dots, S)\}$ . For each individual in  $(\cdot i_s')$  of a CISTG  $A$ ,  $H(\cdot i_s')$  is the record of that individual's subsequent deterministic covariate and survival history for each generalized treatment whose highlighted subgraph at  $t_s$  passes through  $(\cdot i_s')$ .

Suppose there exists an CISTG such that on the basis of our subjective beliefs we assume for each  $\cdot i_s$ ,

$$p[\cdot i_s j_s | \cdot i_s, H(\cdot i_s)] = p(\cdot i_s j_s | \cdot i_s). \quad (4.5)$$

Note that Eq. (4.5) is a weaker assumption than

$$p(\cdot i_s j_s | \cdot i_s, HT(\cdot i_s)) = p(\cdot i_s j_s | \cdot i_s). \quad (4.6)$$

*Definition.* Given a PISTG  $A$ , an assumption is nonidentifiable if complete knowledge of the  $\gamma(\cdot i_s)$ ,  $\gamma(\cdot i_s j_s)$ ,  $S(\cdot i_s j_s)$  is not sufficient to establish whether the assumption is true or false.

*Definition.* Any CISTG satisfying Eq. (4.5) is called a random CISTG (R CISTG).

*Definition.* An CISTG satisfying Eq. (4.6) is called a fully random CISTG (FR CISTG).

Equations (4.5) and (4.6) are nonidentifiable.

**THEOREM.** Given an FR CISTG  $B$ , any CISTG coarser than  $B$  is an FR CISTG.

*Proof.* Follows immediately from the definition of an FR CISTG. An CISTG  $A$  coarser than an R CISTG  $B$  need not be an R CISTG.

**THEOREM 4.1.** Let  $G^B$  be a generalized treatment of an R CISTG  $B$ ; then  $S(\cdot i_s j_s \mid G^B) = S(\cdot i_s j_s \mid B)$ ,  $\gamma(\cdot i_s \mid G^B) = \gamma(\cdot i_s \mid B)$ .

*Proof.* See Appendix C.

*Definition.* An OCISTG (MCISTG) that is an FR CISTG is called an FR OCISTG (FR MCISTG).

*Definition.* Given a generalized treatment “ $G_1$ ” of the STG of a PISTG, define

$$\begin{aligned} S(t_{k+1} \mid "G_1") &\equiv p(D > t_{k+1} \mid "G_1") \\ &\equiv \sum_{i_1=1}^N \gamma(i_1) S(i_1 j_1) \left[ \sum_{i_2=1}^{N_{i_1 j_1}} \gamma(i_1 j_1 i_2) S(i_1 j_1 i_2 j_2) \left[ \cdots \left[ \sum_{i_k=1}^{N_{i_{k-1} j_{k-1}}} \gamma(i_k) S(i_k j_k) \right] \cdots \right] \right] \end{aligned} \quad (4.7)$$

where for any  $\cdot i_s$ , the choice of  $\cdot i_s j_s$  is uniquely determined by “ $G_1$ ”. Note the PISTG need not be a CISTG.

*Definition.*  $S(t, "G_1, "G_2") = S(t \mid "G_1") - S(t \mid "G_2")$

**Corollary 4.1.** Let  $G_1$  be a generalized treatment of an R CISTG; then  $p(D > t_k \mid G_1) = p(D > t_k \mid "G_1")$ .

*Proof.* Using Theorem 4.1, it follows that the right side of Eq. (4.7) is simply a sum over all possible ways of surviving to  $t_{k+1}$  when a population is treated with  $G_1$ .  $S(t \mid "G_1")$  is nothing but the  $G$ -computation algorithm of Sec. 3 written in a compact and computationally efficient form. Note that  $p(D > t_k \mid "G_1")$  is not a survival probability unless PISTG  $A$  is an R CISTG. Rather, it is, by definition, the population parameter of Eq. (4.7). This notational device will be used throughout this paper. Its utility will be obvious. We now give other examples of this notation.

*Definition.* Given a PISTG and  $\cdot i_{s+1}$  on “ $G$ ”,  $p(\cdot i_1 \mid "G") \equiv \gamma(\cdot i_1)$  and  $p(\cdot i_s j_s i_{s+1} \mid "G") \equiv \gamma(\cdot i_{s+1}) S(\cdot i_s j_s) p(\cdot i_s \mid "G")$ . From this recursive definition we can see that  $p(\cdot i_s \mid "G")$  is the product of the intranodal probabilities  $\gamma(\cdot i_k)$  and conditional survival probabilities  $S(\cdot i_k j_k)$  of the sequence of intra- and internodal lines connecting the left circumference of  $t_1$  to  $\cdot i_s$ .

*Definition.*  $p[\cdot i_s j_s(t_{s+1}) \mid "G"]$  is the product of the intrandal probabilities  $\gamma(\cdot i_k)$  and conditional survival probabilities  $S(\cdot i_k j_k)$  [ending with  $S(\cdot i_s j_s)$ ] on the sequence of inter- and internodal lines connecting the left circumference at  $t_1$  to the node  $\cdot i_s j_s(t_{s+1})$ .

*Definition.* We say  $\cdot i_s \in "G^A"$  (alternately  $\cdot i_s j_s \in "G^A"$ ) if  $\cdot i_s$  (alternately  $\cdot i_s j_s$ ) lies on the highlighted subgraph “ $G^A$ ”. Using the above conventions we can rewrite  $p(D > t_k \mid "G_1")$  as  $\sum_{i_{k-1} j_{k-1} \in G_1^A} p[\cdot i_{k-1} j_{k-1}(t_k) \mid "G_1"]$ . In fact, this form is exactly the  $G$ -

computation algorithm. We also define

$$\gamma_D(t_s + \Delta t | "G_1") \equiv \frac{\sum_{i_s j_s \in G_1} \gamma_D(t_s + \Delta t | i_s j_s) p(i_s | "G_1")}{p(D > t_s | "G_1")}.$$

If the PISTG under study is an R CISTG and the “ ” are removed from the left side of each of the above propositions, the definitions become true probability statements about the study defined by  $G_1$ .

Let  $m^A$  be a function that assigns to each internodal line of a PISTG  $A$  a number  $m^A(\cdot i_s j_s)$  subject to the constraint  $\sum_{j_s(\cdot i_s)=1}^{N_{i_s}} m^A(\cdot i_s j_s) = 1$  ( $m^A(\cdot i_s j_s) = 1$  if  $j_s(\cdot i_s) = 0$ ) (“ $G^A$ ” can be identified with any  $m^A$  such that  $m^A(\cdot i_s j_s) = 1$  if  $i_s j_s$  is on the highlighted subgraph “ $G^A$ ”). Let  $i \in (1, \dots, I)$  index the “ $G_i^A$ ”  $\in \mathbf{G}^A$ .“

*Definition.* Given a PISTG  $A$ ,

$$S(t_{k+1} | "m^A") \equiv \sum_i w_i(m^A, t_k) S(t_{k+1} | "G_i^A")$$

where

$$w_i^A(m^A, t_k) \equiv \prod_{i_1=1}^{N_{i_1}} m^A(\cdot i_1 j_1) \left[ \prod_{i_2=1}^{N_{i_2 j_1}} m^A(\cdot i_1 j_1 i_2 j_2) \left[ \dots \left[ \prod_{i_k=1}^{N_{i_k j_{k-1}}} m^A(\cdot i_k j_k) \right] \dots \right] \right]$$

where for any  $\cdot i_s j_s$  the numerical value of  $j_s$  is uniquely determined by  $G_i^A$ . Graphically  $w_i(m^A, t_k)$  is obtained by first writing on each internodal line  $\cdot i_s j_s$  of STG  $A$ , the quantity  $m^A(\cdot i_s j_s)$ , and secondly forming the product of all the  $m^A(\cdot i_s j_s)$  on the highlighted subgraph representing  $G_i^A$  up to the time  $t_k$  in question. Note that  $\sum_i w_i^A(m^A, t_k) = 1$ .

*Definition.*  $p(\cdot i_{k+1} | "m^A")$  is defined exactly like  $S(t_{k+1} | "m^A")$  except  $p(\cdot i_{k+1} | "G_i^A")$  replaces  $S(t_{k+1} | "G_i^A")$ .

LEMMA 4.2.  $p(D > t | A) = S(t | "m^A")$  when  $m^A(\cdot i_s j_s) = \gamma(\cdot i_s j_s)$ .

*Proof.* Direct calculation.

This lemma implies that the observed survival experience of the entire study population in an R MCISTG  $A$  can be written as a weighted average of the survivals that would be observed in controlled trials defined by the  $G^A$ .

LEMMA 4.3. If PISTG  $B$  is coarser than PISTG  $A$ , then, for any “ $G^B$ ,”  $S(t_{k+1} | "G^B") = S(t_{k+1} | "m^A")$  for  $m^A(\cdot i_s j_s)$  chosen as follows: highlight on STG  $A$  those internodal lines whose isomorphic counterparts are on the highlighted subgraph  $G^B$  of STG  $B$ . Each highlighted internodal line is a member of exactly one of the subsets  $A_{k(\cdot i_s)}$  that were used to form  $B$  from  $A$ . For each highlighted line on  $A$  choose  $m^A(\cdot i_s j_s) = \sum_{j_s(\cdot i_s) \in A_{k(\cdot i_s)}} \gamma(\cdot i_s j_s)$ . (That is, in each subset  $A_{k(\cdot i_s)}$  we, as it were, give all the weight to the highlighted line.)

*Proof.* Direct calculation.

COROLLARY 4.3. If the  $G$ -causal parameters of an FR CISTG  $A$  are identically zero, then so are the  $G$ -causal parameters of all coarser FR CISTG.

*Definition.* A set of causal parameters  $B$  are more basic than a set  $A$  if, whenever all the causal parameters in set  $B$  are zero, the causal parameters in set  $A$  must be zero but not vice versa.

Thus, the  $G$ -causal parameters of a FR CISTG  $A$  are more basic than those of a coarser graph  $B$ .

#### F. Causal inference in alternative designed randomized trials

We generalize the characterization of an alternative designed randomized trial given in Sec. 2 as follows: (1) at each time  $t$ , subjects may be assigned treatments that include interventions in addition to the exposure under study; (2) the probability of being assigned to a particular treatment at  $t$  may depend on the subject's past (measured) covariate history (including history of being off protocol). We discretize the timeline for graphical purposes, and assume treatments are given to individuals on protocol only at discrete times  $t_1, t_2, \dots$  with assignment probabilities at  $t_s$  depending on the value of covariates measured at times  $t_1, t_2, \dots$  up to  $t_s$ . We assume that every individual on protocol at  $t_s$  received their assigned treatment. Subjects may leave protocol and later return.

*Definition.* An MPISTG as “detailed as the data” of an alternative designed randomized trial is the *fundamental* MPISTG of the trial if: (1) any subset of the population that was eligible to be randomly assigned to one of  $K$  treatments at  $t_s$  is represented by a right circumference point at  $t_s$  and (2) the  $K$  treatments are given  $K$  separate internodal lines originating from that point.

*Definition.* A PISTG is a sharp causally interpreted tree graph (SCISTG) if each individual in any set  $\cdot i_s'$  has associated a deterministic  $H(\cdot i_s')$  as defined previously.

We shall consider the fundamental MPISTG of an alternative designed randomized trial to be an SCISTG but not a CISTG, since the treatment nature would have given to a subject, if the investigator had not performed the trial, is unlikely to lie on the MPISTG of the observed trial. (For example, if not entered in the trial, no individual may obtain either treatment  $A$  or  $B$  under study.) By the definition of an alternative designed randomized trial, Eq. (4.5) holds for the fundamental SCISTG. The only important difference is that coarser graphs formed from the fundamental SCISTG are not SCISTGs or CISTGs. Otherwise, all theorems for R CISTGs hold for R SCISTGs.

#### G. Nonexistence of a finest MCISTG and FR MCISTG as detailed as the data

**LEMMA 1.** If there exists an MPISTG as detailed as the data, there need not be a finest MCISTG as detailed as the data.

**LEMMA 2.** If there exists a finest MCISTG as detailed as the data, there need not exist a finest FR MCISTG as detailed as the data.

**LEMMA 3.** If there exists a finest FR MCISTG  $A$  as detailed as the data, the comparison of survival curves of the generalized treatments of the finest FR MCISTG and those of coarser FR MCISTGs need not exhaust the causal parameters considered identifiable by a given investigator.

*Remarks.* Lemma 1 follows because there may be a right circumference point from which three internodal lines arise on an MPISTG such that the set  $HT(\cdot i_s)$  is not considered defined by the investigator. Nonetheless, the investigator might consider  $HT(\cdot i_s)$  to be defined for the coarser graphs based on separating individuals with the covariate history represented by the top most line from the bottom two or the bottom line from the superior two. This lack of the associative property means that a finest MCISTG as detailed as the data will not, in general, exist. Similarly, Eq. (4.6) [as well as (4.5)] do not have the associative property. Finally, Lemma 3 follows because an investigator may believe Eq. (4.6) holds for an MCISTG less detailed than the data but for no MCISTG as detailed as the data.

Nonetheless, in practice, we shall suppose that there exists a finest FR MCISTG as detailed as the data and that its  $G$ -causal parameters and those of the FR MCISTGs coarser than it constitute the set of causal parameters one believes can be identified from the data without further *a priori* assumptions.

#### H. Conditional survivals

One may be interested in the  $G$ -causal parameters of a CISTG conditional on being a member of a particular subset of the population (for example, males or, as another example, individuals who have survived to  $t_s$  with a particular employment and exposure history). We associate with any  $\cdot i_s$  of an STG  $B$  the subgraph  $B(\cdot i_s)$  representing the part of the STG  $B$  to the right of that  $\cdot i_s$  but connected to  $\cdot i_s$ . That is,  $B(\cdot i_s)$  includes  $\cdot i_s$  and all internodal lines arising from  $\cdot i_s$ , all nodes on which these terminate, all intranodal lines within these nodes, etc. If  $B$  is a PISTG, CISTG, or FR CISTG,  $B(\cdot i_s)$  can be interpreted to be so as well. For example, the PISTG  $B(\cdot i_s)$  represents the covariate and survival history subsequent to  $t_s$  for individuals in  $\cdot i_s$  in  $B$ .

LEMMA. If  $B$  is an R CISTG then there exists  $G_1^B$ ,  $G_2^B$  such that

$$S[t, G_1^{B(\cdot i_s)}, G_2^{B(\cdot i_s)}] = \frac{S(t, G_1^B, G_2^B)}{p(\cdot i_s | G_1^B)}.$$

*Proof.* The theorem is true by direct calculation for  $S[t, "G_1^{B(\cdot i_s)}", "G_2^{B(\cdot i_s)}"]$ ,  $S(t, "G_1^B", "G_2^B")$ , and  $p(\cdot i_s | "G_1^B")$  where  $G_1^B$ ,  $G_2^B$  agree for all nodes not in  $B(\cdot i_s)$  and  $G_1^B$  agrees with  $G_1^{B(\cdot i_s)}$  and  $G_2^B$  agrees with  $G_2^{B(\cdot i_s)}$  on subgraph  $B(\cdot i_s)$ .

The lemma is false if CISTG  $B$  is not an R CISTG. Graphs  $B(\cdot i_s j_s)$  and  $B(\cdot i_s j_s(t_{s-1}))$  can be defined in a manner similar to  $B(\cdot i_s)$ .

#### I. Computational considerations

The computational burden in computing  $S(t_k | "G_1")$  from Eq. (4.7) increases with increasing numbers of internodal lines originating from nodes at  $t_{k-1}$  on the highlighted subgraph " $G_1$ ." The number of such internodal lines may be so large that even if we knew  $\gamma(\cdot i_s)$  and  $S(\cdot i_s j_s)$  for all  $\cdot i_s$  and  $\cdot i_s j_s$  on " $G_1$ ", we cannot compute  $S(t_k | "G_1")$  even with the aid of a high-speed computer. In such instances we can resort to Monte Carlo methods.

EXAMPLE. Suppose, in our occupational mortality study, that Fig. 3.4 is the finest FR MCISTG for the first  $1\frac{1}{2}$  years of follow-up. The FR MCISTG for the entire 40 year follow-

up period (which we will still call FR MCISTG 3.4) would have  $S = 81$  generations. There are  $2^{80}$  internodal lines between  $t_{80}$  and  $t_{81}$  on the highlighted subgraph representing the generalized treatment of this FR MCISTG. We can estimate  $S(t_{81} | G)$  for any generalized treatment  $G$  by sampling of the employment history paths as follows. For a given generalized treatment  $G$  of interest we choose at random a path  $L_i(t_{80})$  from the  $2^{80}$  possible employment history paths connecting  $t_1$  to  $t_{81}$  by flipping a coin 80 times with success probabilities  $P_1, \dots, P_{80}$  chosen in such a way that  $P_{80}$  depends on the outcomes of the preceding 79 tosses  $H_1, \dots, H_{79}$  [where  $H_j = 1$  (0) if toss  $j$  was a success (failure)] as follows:  $P_1 = p[l(t_1) = 1]$ ,  $P_2 = p[l(t_2) = 1 | L_i(t_1) = H_1, E(t_1)], \dots$ , and  $P_{80} = p[l(t_{80}) = 1 | L_i(t_{79}) = (H_1, \dots, H_{79}), E(t_{79})]$ , where  $L_i(t_s)$  is the randomly selected employment history through  $t_s$ ,  $l(t_s) = 1$  if out of work at  $t_s$  and  $E(t_s)$  is uniquely determined by  $G$  and  $L_i(t_s)$ . Note  $P_s = \gamma(\cdot i_s)$  for  $\cdot i_s$  defined by  $l(t_s) = 1, L_i(t_{s-1}), E(t_{s-1})$ .  $(H_1, \dots, H_{80})$  determine a unique path  $L_i(t_{80})$ . We compute

$$S[L_i(t_{80})] \equiv \prod_{s=1}^{80} p[D > t_{s+1} | L_i(t_s) = (H_1, \dots, H_s), E(t_s), D > t_s] \equiv \prod_{s=1}^{80} S(\cdot i_s j_s)$$

where  $\cdot i_s j_s$  is defined by  $L_i(t_s)$  and  $E(t_s)$ . Then,  $\sum_{i=1}^N S[L_i(t_{80})]/N$  converges in probability to  $p(D > t_{81} | G)$  as  $N \rightarrow \infty$ , where  $N$  is the number of random paths chosen as above. In practice, of course, we will only have estimates of the  $S(\cdot i_s j_s)$  and  $\gamma(\cdot i_s)$  with which to estimate  $S(t_{81} | G)$ .

We can sometimes lessen our computational burden by applying an algorithm that takes as input a PISTG and generates as output a PISTG that, if different from the input, is guaranteed to have fewer internodal lines between  $t_s$  and  $t_{s+1}$ .

*Stage 1 reduction algorithm.* We represent the algorithm in Fig. 4.1. Figure 4.1a is the original PISTG.

Step 1: Merge into a single line all internodal lines originating from  $t_1$  standardly labelled with  $j = 0$  (i.e. merge all internodal lines for which  $N_{i_1} = 1$ ). Merge the associated points of origin on the right circumference of  $t_1$  and merge into one line all the intranodal lines terminating on these points. Merge into a single node all those nodes at  $t_2$  on which the merged internodal lines previously terminated. Step 2: Retain on the right circumference of this single merged node the set of distinct points (and corresponding internodal lines) associated with each of the separate  $t_2$  nodes that were merged. Repetition: For each remaining node at  $t_2$  (which includes the newly constructed nodes), merge all internodal lines labelled with  $j = 0$  and repeat the above steps with  $t_2$  in the role of  $t_1$  and  $t_3$  in the role of  $t_2$ . Continue through  $t_S$ . Figure 4.1b shows Steps 1 and 2 applied to the nodes at  $t_2$ . Figure 4.1c shows Steps 1 and 2 applied to the nodes at  $t_3$ .

Since the Stage 1 reduction algorithm does not affect any  $\cdot i_s$  for which  $N_{i_s} > 1$ , any “ $G_1^A$ ” has a natural counterpart “ $G_1^B$ ” in the Stage 1 reduced graph  $B$  generated from  $A$ . It is the generalized treatment whose set  $\{\cdot i_s j_s, j_s(\cdot i_s) \neq 0\}$  represents the same covariate and survival histories as those in “ $G_1^A$ ”.

**LEMMA.**  $S(t | "G_1^A") = S(t | "G_1^B")$  where  $B$  is the Stage 1 reduction of  $A$  and “ $G_1^B$ ” is the Stage 1 counterpart of “ $G_1^A$ ”.

*Proof.* Direct computation.

Note in Fig. 4.1 the computational savings gained by computing  $S(t | "G_1^B")$  in lieu of  $S(t | "G_1^A")$ . Unfortunately, with PISTG 3.4, the Stage 1 reduction algorithm does not reduce the graph further.

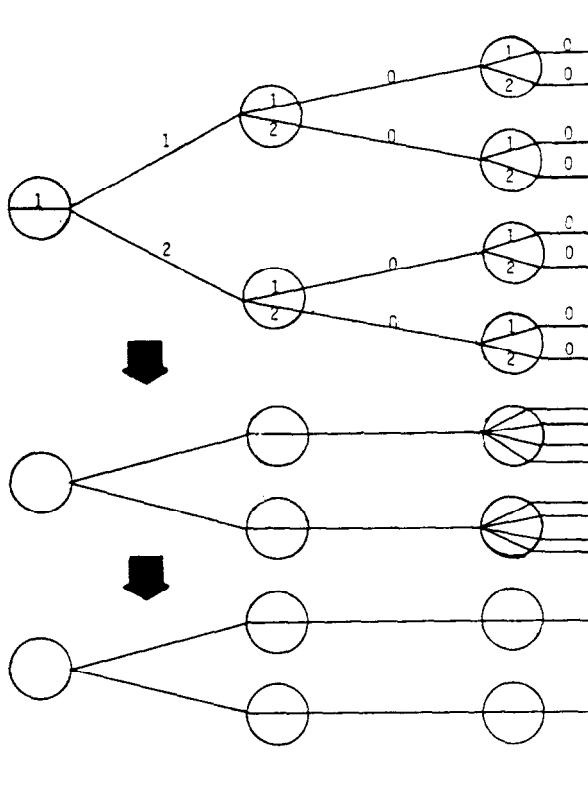


Fig. 4.1. Graphical representation of Stage 1 reduction algorithm.

## 5. ESTIMATION OF THE $G$ -CAUSAL PARAMETERS

### A. Two large sample limiting models

Throughout this paper we will suppose that the sampling fraction  $m/M$  is almost zero so that we may consider the outcomes of the sampled individuals to be statistically independent. Given an MPISTG, the nonparametric maximum likelihood estimators of  $\gamma(\cdot i_s)$  and  $S(\cdot i_s j_s)$  are the corresponding sample proportions.

**LEMMA.** The PL-sufficient statistic (as defined in Appendix A) for  $S(t \mid "G^A")$  is  $N[\cdot i_s j_s^B]$ ,  $N[\cdot i_s j_s^B(t_{s+1})]$ ,  $N[\cdot i_s^B]$  for  $\cdot i_s j_s^B$ ,  $\cdot i_s^B \in "G^B"$ , and " $G^B$ " is the counterpart of " $G^A$ " in the Stage 1 reduced graph  $B$ , and  $N[\cdot i_s^B]$  is the number of sampled individuals observed to be in  $\cdot i_s^B$ .

*Proof.* (Ref. [7]) The import of this lemma is that, in general, given a PISTG  $A$  (and no further identifiable *a priori* assumptions) no information is lost by applying the Stage 1 reduction algorithm. Because we have made no modelling restrictions, we obtain the NPMLE of  $S(t \mid "G^A")$  whether we apply the  $G$ -computation algorithm to PISTG  $A$  or apply it to PISTG  $B$  and estimate  $S(t \mid "G^B")$ .

**Large sample limiting model 1.** Consider a sequence of studies, each represented by an MPISTG  $A_m$ , and indexed by their sample size  $m$ . Let  $A1_m$  be the Stage 1 reduction of  $A_m$ . Under limiting model 1, we suppose that as  $m \rightarrow \infty$ ,  $m/M \rightarrow 0$ , the total number of internodal lines on the  $A1_m$  remains bounded, and, provided  $p_m[\cdot i_s^{A1m}] > 0$ ,  $mp_m(\cdot i_s j_s^{A1m}) \rightarrow \infty$  where  $\cdot i_s j_s^{A1m}$  is the covariate and survival history represented by the internodal line  $\cdot i_s j_s$  (standardly labelled) of  $A1_m$ . This limiting model roughly says that

the number of subjects on each internodal line of the Stage 1 reduction of the MPISTG  $A_m$  becomes large.

Under limiting model 1, the NPMLE of  $S(t_s | \cdot \cdot G_1 \cdot)$  computed from Eq. (4.7) upon replacing  $\gamma(\cdot i_s)$  and  $S(\cdot i_s j_s)$  by their NPMLE will be consistent and asymptotically normal. In realistic cohort mortality studies, the total number of internodal lines between  $t_s$  and  $t_{s+1}$  on the Stage 1 reduced MPISTG (in our example, 2<sup>80</sup>) may be large compared to the number of study subjects. Such sample size limitations require that we model the  $\gamma(\cdot i_s)$  and  $S(\cdot i_s j_s)$ . (As an example, we shall consider the use of statistical models to estimate the  $G$ -causal parameters of the FR MCISTG of Fig. 3.4 in the following subsection.) Furthermore, the asymptotics of limiting model 1 are not relevant. Rather, we will consider the asymptotic properties of estimators under limiting models in which there exists a sequence of  $\cdot i_s j_s^{A^1 m}$  (defined in the previous paragraph) such that  $mp_m(\cdot i_s j_s^{A^1 m})$  is bounded as  $m \rightarrow \infty$ . We shall call any such limiting model a "sparse data limiting model for the sequence of MPISTGs".

### B. Estimation of $S(t | G^{3,4})$ in a cohort of arsenic smelter workers

Lee and Fraumeni[8] assembled data on 8047 arsenic-exposed white males who worked at a Montana copper smelter for at least one year between 1937 and 1956. A worker was entered into follow-up at the latter of January 1, 1938 and his date of first hire (except workers hired prior to January 1, 1938 who were on layoff on January 1, 1938 were not entered into follow-up even if they subsequently returned to the workforce). Our analysis is based on the 5947 workers who were hired subsequent to January 1, 1935 among whom there were 1784 deaths of which 116 were due to lung cancer. Data on each worker consisted of a job and employment history from date of hire to end of follow-up in 1977. Each job was coded as having high, medium, or low arsenic exposure. Subjects off work were considered unexposed. We abstracted information for each worker on current exposure concentration and employment status at six-monthly intervals. Approximately 700 members of our subcohort had unknown vital status at end of follow-up. These workers were considered to be alive at end of follow-up. Such an assumption is common in occupational epidemiology[9] because a search of Social Security Administration records is fairly efficient in verifying deaths but not in verifying survival.

Since we defined  $t_1$  of an MPISTG to be the time at which follow-up was begun and we condition on all relevant history up to  $t_1$ , technically we will have a separate MPISTG for each subset of workers with a particular year of birth, age at hire, and, if hired prior to 1938, a particular exposure and employment history from date of hire to 1938. (None-the-less, our statistical models will "borrow information" across tree graphs.) (We could, but choose not to, represent the entire cohort on a single MPISTG with a separate internodal line at  $t_1$  for each of the above subsets if we used as our time scale the number of years since start of follow-up.) Since  $E(t)$ ,  $L(t)$  are time-dependent covariates a viable modelling strategy is to use the time-dependent Cox proportional hazards model. Since we obtained data only at discrete times, we shall use the discrete failure time regression model of Cox[10] that specifies a linear log odds model for conditional probabilities (discrete hazards) at each failure time. Examples of the type of models one might use for the conditional probabilities of MPISTG 3.4 (modified so that there are three internodal lines arising from the superior right circumference point in each node representing low, medium and high exposure at work) are

$$\text{logit}[\gamma_{D,i}[t + \Delta t | E(t), L(t), Z(t_1)]] = \beta_{0,D,i,t} + \boldsymbol{\beta}_D \cdot \mathbf{X}_D \quad (5.1)$$

$$\begin{aligned} \text{logit}[\gamma_{L,i}[t + \Delta t | E(t), l(t)] \\ = 0, D > t + \Delta t, L(t - \Delta t), Z(t_1)]] = \beta_{0,L,i,t} + \boldsymbol{\beta}_L \cdot \mathbf{X}_L \end{aligned} \quad (5.2)$$

$$\text{logit}[\gamma_{R,i}[t + \Delta t | E(t), D > t + \Delta t, L(t - \Delta t), l(t) \\ = 1, Z(t_1)]] = \beta_{0,R,i,t} + \beta_R \cdot X_R \quad (5.3)$$

where

- (1)  $i$  is a stratum indicator and we have stratified jointly by five-year intervals of age at hire and five year intervals of calendar year of hire.
- (2)  $t$  is age and we allow  $t$  to take on only the values  $\Delta t, 2\Delta t, 3\Delta t, \dots$  with  $\Delta t = 0.5$  years.
- (3)  $\gamma_{D,i}(t | \cdot)$  is the conditional probability of being dead by age  $t$ , given that one was alive at age  $t - \Delta t$  in stratum  $i$  with covariate history represented by  $\cdot$ .  $Z(t_1)$  is all information on the subject up to age at start of follow-up.
- (4)  $\gamma_{L,i}(t | \cdot)$  is the conditional probability of leaving work at age  $t$ ,  $t > t_1$ , given one is alive at  $t$  and was at work at  $t - \Delta t$  [i.e.  $l(t - \Delta t) = 0$ ] in stratum  $i$  given  $\cdot$ .
- (5)  $\gamma_{R,i}(t | \cdot)$  is the conditional probability of returning to work at age  $t$  ( $t > t_1$ ) given one was out of work at  $t - \Delta t$  and alive at  $t$  in stratum  $i$  given  $\cdot$ .
- (6) With  $J$  labelling failure type (i.e.  $J \in \{L, D, R\}$ ),  $X_J$  is a known vector of covariates determined by  $E(t), L(t), Z(t_1)$ .  $\beta_J$  is a vector of unknown regression coefficients and  $\beta_J \cdot X_J$  is their inner product. The  $\beta_{0,J,i,t}$  are unknown coefficients which can vary with age and stratum level as well as with failure type. Note the slightly different meanings for  $L$  in  $\gamma_L$  and in  $L(t)$ .

We now describe our approach to estimating the parameters of Eqs. (5.1)–(5.3). Note that if we knew the conditional probabilities on the left side of Eqs. (5.1)–(5.3) we would know  $S(\cdot | i, j_s)$  and  $\gamma(\cdot | i_s)$  for each MPISTG 3.4 [indexed by  $(Z(t_1))$ ] of the copper smelter workers study. In order to save computing costs, for failures of type  $J$  at  $t$ , we sampled controls without replacement from the noncases at risk for failure  $J$  at time  $t$  (matched to the case on stratum  $i$ )[11–13]. We sampled five controls per case. Because each individual may leave and return to work on a number of different occasions, in our cohort the total number of “failures” exceeded 10,000 for both “leaving” and “returning.” To further save computing costs we analyzed only random samples of 700 of these two failure types.

Because of the large number of stratum specific baseline discrete hazards  $\gamma_{J,i}(t) \equiv e^{\beta_{0,J,i,t}} / (1 + e^{\beta_{0,J,i,t}})$ , we fit models (5.1)–(5.3) using conditional logistic regression[13]. (Technically, the conditional logistic regression estimates are the maximum partial likelihood estimates of the  $\beta_J$ [12].) The parameters  $\beta_{0,J,i,t}$  do not appear in the conditional likelihood function. We then estimated  $\gamma_{J,i}(t)$  by Breslow’s[14] estimator  $d_{J,i,t} / \sum_k e^{\beta_J \cdot X_{J,k}}$ , where  $d_{J,i,t}$  is the number of failures of type  $J$  in stratum  $i$  at age  $t$ ,  $k$  indexes the individuals in the risk set of the  $d_{J,i,t}$  cases, and  $\beta_J$  are the maximum partial likelihood estimates. We computed Breslow’s estimator from the full cohort data. Because Breslow’s estimator is noniterative, the computational burden is not excessive. Breslow’s estimator is only a good estimator of  $\gamma_{J,i}(t)$  if the  $\gamma_{J,i}(t | E(t - \Delta t), L(t - \Delta t), Z(t_1))$  are less than 10%. This required us to choose  $\Delta t$  to be as short as 6 months. Even so, near retirement (i.e. age 65) the Breslow estimator may be a poor estimator of the baseline hazard of leaving work unless one chooses  $(t_s - t_{s-1})$  to be even less than 6 months for ages near 65.

The conditional probabilities estimated in Eqs. (5.1)–(5.3) can then be used to estimate the probabilities  $S(\cdot | i, j_s)$  and the  $\gamma(\cdot | i_s)$  required by the Monte Carlo algorithm estimator of  $S(t | "G^{3,4}")$ . To compute the overall sampling error of the estimate of  $S(t | "G^{3,4}")$  we would have to use bootstrap resampling methods.

*Large sample limiting model 2 (a sparse data limiting model).* Under limiting model 2, given a sequence of MPISTGs indexed by sample size  $m$ , as  $m \rightarrow \infty$  we suppose  $m/M \rightarrow 0$ ,  $m\Delta t$  is bounded, and the number of individuals at risk for event  $J$  in each stratum  $i$  and at each age  $t$  increases without bound.

Under limiting model 2, under suitable regularity conditions, the Monte Carlo estimator of  $S(t | "G_1^{3.4}")$  will be consistent asymptotically normal as  $m \rightarrow \infty$  and  $N \rightarrow \infty$  (where  $N$  is the number of Monte Carlo trials) provided the models given in Eqs. (5.1)–(5.3) are correctly specified. A formal proof would require Martingale methods and is not given.

If in Eqs. (5.1)–(5.3) we assumed a linear logistic form for the nuisance hazards, that is,

$$\beta_{0,j,i,t} = \beta_{0,i,j} + \beta'_{0,i,j} \cdot t. \quad (5.4)$$

then, an estimator of  $S(t | "G_1^{3.4}")$  based on the fit of Eqs. (5.1)–(5.3) [incorporating Eq. (5.4)] by the method of “unconditional maximum likelihood” would be consistent under a large sample limiting model in which the number of individuals in each stratum  $i$  increased without bound. Provided models (5.1)–(5.3) are correctly specified, for any  $G_1$  of interest and for several  $t_k$  of interest, e.g.  $t_k = 55, 60, 65, 70$  years of age, joint confidence intervals for the  $S(t_k, G_1^{3.4}, G^{3.4} \equiv 0)$  can be constructed by bootstrapping (i.e. by refitting repeated samples of the original data). Similarly, valid tests of the null hypothesis that  $S(t, G^{3.4}, G^{3.4} \equiv 0) = 0$  for all  $G^{3.4}$  and  $t$  might be based on constructing joint confidence intervals as above for several different combinations of  $G^{3.4}$  and  $t_k$ .

In order that readers who are uninterested in the details of the analysis of the smelter worker cohort are not inconvenienced, we defer to Appendix D the specification of models (5.1)–(5.3) and the results of fitting those models.

In practice, estimates of  $S(t | "G_1^{3.4}")$  can be shown to be quite sensitive to the specifications of the models in Eqs. (5.1)–(5.3). Unfortunately, in practice, model misspecification is unavoidable. Often, with many occupational exposures, the central first question is whether the exposure under study has any effect on the mortality experience of any individual. Fortunately, under circumstances described below, nonparametric tests of the null hypothesis of no exposure effect on any individual's mortality can be constructed.

## 6. NONPARAMETRIC TESTS FOR THE CAUSAL EFFECT OF EXPOSURE

### A. *G*-null tests

Often the null hypothesis that exposure has no effect on any individual's mortality can be expressed as the “sharp null hypothesis” for a particular MCISTG.

*Definition 6.1.* The sharp null hypothesis of a CISTG holds if and only if for all individuals  $i$  and for all  $G_1^A, G_2^A \in \mathbf{G}^A$ ,  $S(t | G_1^A, i) = S(t | G_2^A, i)$  for  $t_1 \leq t \leq t_{S+1}$ . That is, no individual's survival history is influenced by the generalized treatment they might receive. It is important to note that an investigator's null hypothesis of substantive interest might not be the sharp null hypothesis of any CISTG  $A$ .

**EXAMPLE.** Suppose that Fig. 3.3 was believed to be an MCISTG and the null hypothesis of interest was “exposure had no effect on any individual's mortality, controlling for exposure history.” This null hypothesis need not be a sharp null hypothesis, since it could hold even if the  $G$ -causal parameter comparing the outcomes of the hypothetical studies defined by the generalized treatments of MCISTG 3.3 “if alive at  $t_s$ , receive zero exposure at work” and “if alive at  $t_s$ , receive zero exposure off work” were nonzero. On the other hand, if the null hypothesis of interest to the investigator was that there was “no overall effect of exposure on any individual's mortality,” this null hypothesis would be the sharp null hypothesis of MCISTG 3.4.

*Definition.* The  $G$ -null hypothesis for a CISTG  $A$  holds if and only if the population  $G$ -causal parameters of CISTG  $A$  are all identically zero which we write  $S(t, G_1^A, G_2^A) \equiv 0$  (This is a departure from our usual use of  $\equiv$  as meaning "is defined equal to"). If the sharp null hypothesis holds for an MCISTG  $A$ , then the  $G$ -null hypothesis holds for MCISTG  $A$  and for all MCISTG coarser than  $A$ . Therefore, any nonparametric test of the  $G$ -null hypotheses of  $A$  and of all MCISTG coarser than  $A$  is a test of the sharp null hypothesis. When MCISTG  $A$  is an FR MCISTG, by Corollary 4.3, we need only to test the  $G$ -null hypothesis of  $A$ . [It is easy to show by counterexample that Corollary 4.3 is false if CISTG  $A$  is not an FR CISTG (provided the sharp null hypothesis does not hold for  $A$ ).]

*Definition.* The " $G$ "-null hypothesis holds for a PISTG [written  $S(t, "G_1^A", "G_2^A") \equiv 0$ ] if  $S(t, "G_1^A", "G_2^A") = 0$  for all  $t, "G_1^A", "G_2^A"$ .

**LEMMA.** If  $A$  is an R CISTG, the  $G$ -null hypothesis holds if and only if the " $G$ "-null hypothesis holds.

We will use the following notational convention, which we will continue to use throughout the paper.

*Notational convention.* Given a PISTG,  $\cdot i_s j_s$  and  $\cdot i'_s j'_s$  are two arbitrary internodal lines arising from the same right circumference point  $\cdot i_s$ .  $\cdot i'_s j'_s$  and  $\cdot i_s j_s$  are two arbitrary internodal lines arising from possibly different right circumference points of a single node at  $t_s$ .  $\cdot i'_s j'_s$ ,  $\cdot i_s j_s$  are two arbitrary internodal lines at  $t_s$  that may or may not be in the same node and may or may not arise from the same right circumference point.

**THEOREM 6.1.** Given an R CISTG  $A$ , the  $G$ -null hypothesis holds if and only if for each  $\cdot i_s$ ,  $S(t | \cdot i_s j_s) \equiv p(D > t | \cdot i_s j_s) = p(D > t | \cdot i'_s j'_s) \equiv S(t | \cdot i'_s j'_s)$  whenever  $j_s(\cdot i_s) \neq j'_s(\cdot i_s)$ .

*Proof.* See Appendix E.

Theorem 6.1 says that the  $G$ -null hypothesis holds for an FR CISTG if and only if, for any right circumference point, the subsequent survivals (in the observed study) for the various treatment groups (defined by the internodal lines arising from that right circumference point) are the same. We show, by example, how Theorem 6.1 can be used to construct nonparametric tests of the null hypothesis that all  $G$ -causal parameters of an FR MCISTG are identically zero.

**EXAMPLE.** Suppose MPISTG 3.4 is an FR MCISTG. It is easy to check that at each time  $t_s$  ( $s > 1$ ) there are  $2 \times 3^{s-2}$  right circumference points  $\cdot i_s$  from which two or more internodal lines originate. There are a total of  $3^{s-1}$  such right circumference points on the graph. At each such  $\cdot i_s$ , we can calculate the numerator of the ordinary log rank test comparing the subsequent survivals of individuals (in that  $\cdot i_s$ ) who received high exposure at  $t_s$  with the survivals of those who received zero exposure at work at  $t_s$ . Under the  $G$ -null hypothesis each log rank numerator has expectation zero. In Theorem E1 of Appendix E it is shown that the log rank numerators corresponding to different  $\cdot i_s$  are uncorrelated. Thus, a one degree of freedom summary test based on a sum of these log rank numerators standardized by the square root of the sums of the usual log rank variances will, under mild regularity conditions, be asymptotic normally distributed with expectation zero and variance one. We need not actually compute all  $3^{s-1}$  log rank numerators, as most will be zero.

Rather, we use the following *G-null test algorithm*. (1) For each time  $t_s$  at which a death occurs [i.e. has occurred in the interval  $(t_{s-1}, t_s]$ ] determine for each individual who died at  $t_s$  the subsets  $\cdot i_k$  with  $N_{\cdot i_k} > 1$ , in which he had been a member (i.e. those from which two or more internodal lines arise). (In the case of FR MCISTG 3.4 this would mean determining the times at which each case failing at  $t_s$  had been at work.) (2) For each such  $\cdot i_k$  determine if there are any members of the subset  $\cdot i_k$  who survived past  $t_s$ . (3) If so, construct a  $2 \times N_{\cdot i_k}$  contingency table denoted  $(\cdot i_k, t_s)$  in which the two rows are cases (individuals in  $\cdot i_k$  who died between  $t_{s-1}$  and  $t_s$ ) and controls (individuals in  $\cdot i_k$  who survived past  $t_s$ ). The columns represent the  $N_{\cdot i_k}$  possible treatments at  $t_k$  for individuals in  $\cdot i_k$ . If  $N_{\cdot i_k} = 2$ , then the log rank numerator, say  $O(\cdot i_k, t_s) - E(\cdot i_k, t_s)$  (i.e. the Mantel-Haenszel numerator) is computed for table  $(\cdot i_k, t_s)$ . If  $N_{\cdot i_k} > 2$  and the treatments have been quantitatively scored, the numerator of the log rank (Mantel-Haenszel) test for trend is computed. The appropriate Mantel-Haenszel (i.e. hypergeometric) variance is computed. (If  $N_{\cdot i_k} > 2$  and the treatments have no natural ordering, one could compute the  $(N_{\cdot i_k} - 1)$ -dimensional numerator of the  $N_{\cdot i_k}$ -sample log rank test.) (4) Sum the numerators  $O(\cdot i_k, t_s) - E(\cdot i_k, t_s)$  from each table, possibly after multiplication by weight functions  $w(\cdot i_k, t_s)$  chosen so as to increase power against alternatives felt to be *a priori* likely. Multiply the hypergeometric variance associated with each table by the square of  $w(\cdot i_k, t_s)$  and sum over tables. (5) Divide the numerator sum by the square root of the variance sum and compare to the standard normal distribution (see Theorem E1). A program, "G-Null Test", written in PASCAL by Donald Blevins and the author is available upon request. It will compute the *G*-null test for any FR MCISTG, any exposure scoring function, and any weight function. If at each time  $t_s$  at which a case occurs, only a random sample of potential controls are selected from among study subjects who survived past  $t_s$ , the validity of the *G*-null test is not affected. The above algorithm for such case-control data can be modified so that at step (2), one need only determine whether any member of the "matched controls" of cases failing at  $t_s$  are in  $\cdot i_k$ .

### B. Remarks on the power of *G*-null tests in sparse data

We would like to compare in sparse data the power of representative members (indexed by differing table weight and exposure scoring functions) of the class of *G*-null tests of the *G*-null hypothesis to the power of a test of the *G*-null hypothesis that is optimal or near optimal against specified alternatives.

**EXAMPLE.** Suppose, as is implicitly assumed in standard analyses of occupational mortality studies, that Fig. 3.5 is an FR MCISTG. Then tests of  $\beta_1 = 0$  in the model

$$\text{logit}\gamma_D[t + \Delta t | E(t)] = \beta_{0,t} + \beta_1 ce(t) \quad (6.1)$$

are tests of the *G*-null hypothesis for FR MCISTG 3.5 [where  $ce(t)$  is cumulative (measured) exposure up to  $t$ ]. The Cox (partial likelihood) score statistic is

$$\sum \left[ \frac{dce(t_s)}{d(t_s)} - \bar{ce}(t_s) \right] \quad (6.2)$$

where the sum is over the distinct death times,  $dce(t_s)$  is the total cumulative exposure among the  $d(t_s)$  individuals dying in  $(t_s - \Delta t, t_s]$ , and  $\bar{ce}(t_s)$  is the average cumulative exposure among individuals at risk at  $t_s - \Delta t$ . Tests based on the partial likelihood score are nearly optimal against alternatives parameterized by Eq. (6.1) for  $\beta_1$  close to zero.

When the data is sparse,  $G$ -null tests (i.e. tests based on the  $G$ -null test algorithm) may have very poor power properties compared to tests based on the partial likelihood score against alternatives defined by Eq. (6.1). As an extreme example, suppose, at each time  $t_s$ , there are 100,000 rather than two possible exposure levels, and that no two sampled study subjects received identical exposures at  $t_1$ . (This might be the case if individual exposure measurements were made with a precise measuring instrument.) In such a case, no right circumference point beyond  $t_1$  of FR MCISTG 3.5 (modified so as to reflect 100,000 exposure levels) will represent more than one study subject. Therefore,  $G$ -null tests will depend only on exposures received at  $t_1$ . If this initial exposure is uncorrelated or inversely correlated with cumulative exposure at representative death times  $t$ , a one-sided  $G$ -null test will have power against the alternatives of Eq. (6.1) equal to or less than the  $\alpha$ -level of the test.

As a second example, suppose now there are but four or five exposure levels; data is collected every few hours [so  $(t_s - t_{s-1})$  is small]; and for each individual there is a small hour-to-hour variation in exposure level around the individual's daily mean. Then, within a few months from start of follow-up, no right circumference point on FR MCISTG 3.5 will contain more than one study subject. Again, our  $G$ -null test will depend largely on exposures received near start of follow-up. Fortunately, for FR MCISTG 3.5, tests based on Eq. (6.2) will retain good power against alternatives in which exposures received at times past start of follow-up influence mortality.

Unfortunately, if the STG of our (Stage 1 reduced) FR MCISTG has a nonnegligible fraction of nodes with more than one intranodal line (e.g. as would be the case when the healthy worker survivor effect is operative and our finest FR MCISTG is Fig. 3.4), and if we make no *a priori* assumptions beyond the nonidentifiable randomization assumptions of Eq. (4.6); then, under sparse data asymptotics, only the  $G$ -null tests will reject the  $G$ -null hypothesis at the nominal rate for all states of nature consistent with the  $G$ -null hypothesis. In this paper, tests with the above property shall be called nonparametric tests of the  $G$ -null hypothesis. This is a slight abuse of common usage.

As an example of a test that will fail to reject at the nominal coverage rate, consider the test described in Sec. 5B, which is based on estimating  $S(t \mid \cdot^G)$  for various  $G$  based on models (5.1)–(5.3) and then bootstrapping to obtain standard errors. Since  $RR(t + \Delta t \mid \cdot^G) = \gamma_D(t + \Delta t \mid \cdot^G)/\gamma_D(t + \Delta t \mid \cdot^G = 0)$  cannot be written as a function of

$$\frac{\gamma_L[t + \Delta t \mid E(t), L(t)]}{\gamma_L[t + \Delta t \mid E(t) = 0, L(t)]}, \frac{\gamma_D[t + \Delta t \mid E(t), L(t)]}{\gamma_D[t + \Delta t \mid E(t) = 0, L(t)]}, \frac{\gamma_R[t + \Delta t \mid E(t), L(t)]}{\gamma_R[t + \Delta t \mid E(t) = 0, L(t)]}$$

(i.e.  $RR(t \mid \cdot^G)$  depends on the nuisance hazards for returning, leaving, and death), the “ $G$ ”-null hypothesis may hold even though all of the above hazard ratios differ from unity [as would be the case if high exposure functioned as an irritant causing individuals with poor prognosis to terminate employment, but neither exposure nor work status per se (causally) influenced mortality]. It follows that when models for Eqs. (5.1)–(5.3) are misspecified, the true  $\alpha$ -level of the test based on fitting Eqs. (5.1)–(5.3) and bootstrapping can differ from the nominal  $\alpha$ -level. Contrast this result to that in the situation in which Fig. 3.5 is assumed to be an FR MCISTG. In that case, although estimates of the  $G$ -causal parameters  $S[t + \Delta t \mid E(t)] - [S(t + \Delta t \mid E(t) = 0)]$ , where

$$S[t_s + \Delta t \mid E(t_s)] = \prod_{k=1}^s [1 - \gamma_D(t_k + \Delta t \mid E(t_k))] \quad (6.3)$$

[where  $E(t_k)$  is the initial part of  $E(t_s)$ ]

will be biased by misspecification of the relative risk model  $\gamma_D[t + \Delta t | E(t)]/\gamma_D[t + \Delta t | E(t) = 0] = RR[t + \Delta t | E(t)]$ , tests of the  $G$ -null hypothesis that  $RR(t + \Delta t | E(t)) = 1$  will be *valid*, i.e., will reject at its nominal level under the *null*.

Thus, for FR MCISTG 3.4, how might we increase power of the  $G$ -null tests in sparse data against alternatives in which exposures received long after start of follow-up influence mortality. We could take one of several approaches. First, we could give sharply increased weight to those relatively few tables associated with nodes far from start of follow-up that contribute information to the test statistic. Of course, giving large weight to a few small tables may itself decrease power. Alternatively, one could choose to group the measured exposure levels into fewer categories and/or to record data at less frequent intervals. [One could also redefine measured exposure at  $t_s$  to be a category of average exposure over the interval  $(t_{s-1}, t_s]$ . This device could be useful in our second example above if we also increase  $\Delta t$ .] That is, by one method or another one could reduce the number of internodal lines arising from each right circumference point and also increase  $\Delta t$ . Unfortunately, if higher exposure tends to make unhealthy individuals preferentially leave work, then, even under the  $G$ -null hypothesis, an MCISTG formed from R MCISTG 3.4 by either grouping exposure levels or by recording data at less frequent intervals will not itself be an R MCISTG. Thus, “ $G$ -null tests” based on such an MCISTG may fail to reject at the nominal rate under the  $G$ -null hypothesis. We may be faced with the usual trade-off between power and bias. We stress the “may” in the last sentence because, depending on the true state of nature and the true correlation between early and late exposure concentrations, the effect of grouping exposure and/or recording data at less frequent intervals could be to decrease power.

### C. Selection bias caused by cohort definition

The definition of a PISTG in Section 4 required that  $t_1$  be start of follow-up, defined as the first time at which an individual, had he died, would have had his death recorded for data analysis. In fact, such a restriction is unnecessary. When selection into follow-up is unrelated to health status, the requirement that  $t_1$  represents start of follow-up can result in severe loss of efficiency as the example below indicates.

**EXAMPLE.** Consider the occupational study represented by FR MCISTG 3.4 extended to  $t_{81}$  with  $\Delta t = \frac{1}{2}$  year. Suppose mortality data for the time period  $t_1$  to  $t_{50}$  were not recorded, but all exposure and employment data from time  $t_1$  onwards were available. If the workers were age 20 at time of hire ( $t_1$ ), few deaths would be lost. Now, groups of workers with different levels of  $[E(t_{50}), L(t_{50})]$  would not, in general, be comparable at  $t_{50}$ , even under the  $G$ -null hypothesis, since exposure may, for example, be an irritant that makes sick individuals tend to leave work. It is possible that no two workers have the same vector  $[E(t_{50}), L(t_{50})]$ . Thus, any FR MCISTG starting at  $t_{50}$  (i.e. an FR MCISTG 3.4( $i_{49}j_{49}(t_{50})$ )) defined by a particular history  $[E(t_{50}), L(t_{50})]$ , would contain no right circumference point with two or more subjects. Thus, no test of the  $G$ -null hypothesis would be possible under the above restriction.

But this is nonsense, of course, because the standard  $G$ -null test for FR MCISTG 3.4 applied to a worker dying at, say,  $t_{60}$  only requires vital status information on individuals who were alive at  $t_{59}$  (although it requires exposure and work history data from  $t_1$ ). Thus, it would be irrelevant whether follow-up started at  $t_1$  or  $t_{50}$ . In particular, if exposure is received at random, conditional on past work and employment history, and, as in our cohort, follow-up begins in 1938, we may allow  $t_1$  on FR MCISTG 3.4 to be year of hire, say 1905 (matching, as usual, on risk factors such as age at hire), provided subjects are not selected into follow-up on the basis of their health status.

Unfortunately, in many occupational studies, our arsenic study being one, the cohort is defined in such a way that follow-up is initiated on workers hired prior to a given year

only if they were at work in that year (in our study 1938). This method of cohort definition can invalidate the contribution to the  $G$ -null tests from tables  $(\cdot i_k, t_s)$  with  $t_k$  (the time the exposure is received) prior to the selection date (1938). To see why, suppose that, although exposure has no effect on mortality, high exposure functions as an irritant that causes less healthy individuals to leave work. Consider two 35 year old individuals hired in 1935, one of whom receives high exposure and the other zero exposure at that time. Suppose both have been selected into the cohort by virtue of their being at work in 1938 and one dies in 1940. For the contribution to the  $G$ -null test from the 1935 exposures to be valid, both must be equally likely to have died in 1940 under the null. But of the two, the one who received high exposure in 1935 is more likely to have lived past 1940, since (by virtue of his being employed in 1938) he is likely to be healthy. We are presently investigating the magnitude of the selection effect associated with this method of cohort definition. In the meanwhile, the only safe option is to match on exposure and work history until time of selection. In our arsenic data, this means matching an exposure and work history up to 1938.

If the healthy worker survivor effect were not operating (i.e. MPISTG 3.3 and/or 3.5 were an FR MCISTG), no bias is introduced by this method of cohort definition.

#### D. An example

The results of applying the  $G$ -null test algorithm to the 1784 deaths that occurred in the 5947 copper smelter workers described in Sec. 5B are reported in Table 3. We used several modified versions of FR MCISTG 3.4. The modified STGs 3.4 are based on either three exposure categories at work (H, M, L) or two exposure categories at work (H or M versus L), where here L stands for low exposure. The time scale for the causal tree graphs was chosen to be time since hire.  $\Delta t$  was chosen to be 6 months, 1 year, 2 years, or 3 years. The weight functions  $w(i_k, t_s)$  were chosen to be identically 1 and  $1/|t_s - t_k|$ . Comparisons were made between choosing 50, 25 and 5 controls per case. Controls were matched to cases on time since hire, to within 6 months on age at hire, and to within 3 years on calendar period of hire. To avoid any bias due to the fact that workers hired prior to 1938 were selected into follow-up only if employed in 1938, no individual was allowed to contribute to the  $G$ -null test statistic until 1938. This effectively matches on

Table 3. Z-scores<sup>†</sup> of  $G$ -null tests for the effect of arsenic exposure on total mortality in a cohort of copper smelter workers

Exposure score	Exposure grouping					
	H M L <sup>‡</sup>			H or M L		
	3	2	1	2	1	
	Weight functions		Weight functions		Weight functions	
	$1/ t_s - t_k $ <sup>§</sup>		Constant = 1		Constant = 1	
	No. of controls/case		No. of controls/case		No. of controls/case	
	50		50	20	5	50
$\Delta t$						
6 months	3.13		2.61 (1.16 <sup>¶</sup> )			
1 year	3.46		3.01	2.85	1.33	2.83 <sup>†</sup>
2 years	4.4		3.9			
3 years	4.7		4.23			

<sup>†</sup> Standard normal deviate under the  $G$ -null hypothesis.

<sup>‡</sup> H, M, L—high, medium, low exposure concentration.

<sup>§</sup> [Time from exposure to death]<sup>-1</sup>.

<sup>¶</sup> Summary Mantel-Haenzel odds ratio based on scoring H or M = 1, L = 0.

exposure and employment history till selection date. An unweighted summary Mantel-Haenszel odds ratio was computed over all tables with high or medium exposure constituting the exposed and low exposure, the unexposed.

Although the value of the Mantel-Haenszel odds ratio would not be of much scientific interest in itself [since it is essentially a weighted average of functions of the  $G$ -causal parameters of FR MCISTGs (one for each  $\cdot i_s$ ) coarser than the FR MCISTGs  $3.4(\cdot i_s)$ ]; nonetheless, we report it in order to facilitate the interpretation of our test of the null hypothesis. In nonexperimental studies all assumptions concerning comparability are certain to be at least slightly wrong. As such, even if the null hypothesis of no exposure effect is true for MCISTG 3.4, with a sufficiently large sample size any test will reject the null hypothesis with near certainty due to residual confounding (that is, due to the fact that MCISTG 3.4 will never exactly be an FR MCISTG). The  $p$ -value of, say, 0.005 is more likely to be due to residual confounding when associated with a Mantel-Haenszel odds ratio of 1.1, say, than when associated with a Mantel-Haenszel odds ratio of 3. This reflects the fact that in the latter case a strong confounder must have been overlooked, which is *a priori* less likely.

The main points of interest in Table 3 are (1) there does seem to be an adverse effect of arsenic on total mortality, at least on the assumption that MCISTG 3.4 is an FR MCISTG and (2) it seems clear that five controls per case are too few. Our results are consistent with those obtained in Appendix D based on fitting models (5.1)–(5.3). The standard analysis based on the Cox score of Eq. (6.2) had a  $z$ -score of 0.001, 1.25, 1.5 based on scoring exposure 1, 2, 3; 0, 1, 2; 0, 1, 4, respectively, for low, medium, and high exposure. Thus, the standard analysis failed to demonstrate an adverse effect of arsenic on total mortality, whereas our approach did so.

## 7. AN ISOMORPHISM BETWEEN THE VIEWS OF CAUSAL INFERENCE EXPRESSED IN SECTIONS 2 AND 3

### A. Incorporation of further *a priori* assumptions

Until this point, we have made no *a priori* assumptions other than the nonidentifiable assumptions necessary to define an MPISTG to be an FR MCISTG (with the exception that in Sec. 5 estimation of  $G$ -causal parameters from sparse data required that we make additional *a priori* assumptions as reflected in the specification of statistical models). In this section, we consider the use of further identifiable *a priori* assumptions in order to allow us both to more efficiently estimate the causal parameters of R MCISTGs, and to estimate the causal parameters of R OCISTGs that are not R MCISTGs (because data on certain covariates have not been obtained). We shall see that, given particular *a priori* information concerning an FR OCISTG  $A$ , there may (under circumstances defined below) exist an MPISTG that represents the PL-sufficient statistic for estimating the  $G$ -causal parameters of FR OCISTG  $A$  (see Appendix A). This MPISTG, which will be called the “ $A$ -complete Stage 0 PL-sufficient reduced graph of R CISTG  $A$ ”, we shall now define. We will require a number of preliminary definitions.

**Definition.** A PISTG  $B$  is a Stage 0 reduction of a PISTG  $A$  if and only if each subset  $[\cdot i_s^B]$  is the union of subsets  $[\cdot i_s^A]$  and each  $[\cdot i_s j_s^B]$  is the union of subsets  $[\cdot i_s j_s^A]$  [that is, the subset of the population represented by any intranodal (internodal) line at  $t_s$  on PISTG  $B$  is the union of subsets of the population represented by intranodal (internodal) lines at  $t_s$  on PISTG  $A$ ].

**EXAMPLE.** PISTG 3.5 is a Stage 0 reduction of PISTG 3.4. For instance, any internodal

line of 3.5 at  $t_s$  is specified by a particular exposure history through  $t_s$ . Any internodal line at  $t_s$  on PISTG 3.4 is specified by a particular exposure and employment history through  $t_s$ . Thus, one need only take the union over internodal lines on 3.4 with the same exposure history (but different employment histories).

*Definition.* If PISTG  $B$  is a Stage 0 reduction of PISTG  $A$  then  $G_1^A$  has a Stage 0 counterpart  $G_1^B$  if whenever  $[i_s^A] \in G_1^A$  and  $[i_s'^A] \in G_1^A$  are in the same  $[i_s^B]$ , then  $[i_s j_s^A] \in G_1^A$  and  $[i_s' j_s'^A] \in G_1^A$  are in the same  $[i_s j_s^B]$  (where we have used the notational convention described in Sec. 6A). The Stage 0 counterpart  $G_1^B$  is defined by the property  $i_s j_s^B \in G_1^B$  if for some  $i_s j_s^A \in G_1^A$ ,  $[i_s j_s^A] \subset [i_s j_s^B]$ . We also call  $G_1^A$  a Stage 0 counterpart of  $G_1^B$ . A  $G_1^B$  may have several Stage 0 counterparts.

We describe a graphical procedure for determining whether a given  $G^A$  has a Stage 0 counterpart  $G^B$ . Each intranodal line  $i_s^A$  and internodal line  $i_s j_s^A$  has a counterpart in the Stage 0 reduction  $B$ . These counterparts are, respectively, the  $i_s^B$  and  $i_s j_s^B$  in which the  $i_s^A$  and  $i_s j_s^A$  are contained. Now, a given  $G^A$  has a counterpart  $G^B$  if the counterpart of the set of intra- and internodal lines on the highlighted subgraph  $G^A$  forms a subgraph on PISTG  $B$  which constitutes a generalized treatment  $G^B$  (that is, the subgraph on  $B$  could have been generated by the generalized treatment algorithm).

*EXAMPLE.* “ $G_1^{3.4}$ ” defined by “if at work receive zero exposure” has a Stage 0 counterpart “ $G_1^{3.5}$ ”, defined by “if alive receive zero exposure”, while “ $G_2^{3.4}$ ” defined by “if at work, receive high exposure” has no Stage 0 counterpart “ $G_2^{3.5}$ ”. Conversely,  $G_2^{3.5}$  defined by “if alive, receive high exposure” has no Stage 0 counterpart  $G^{3.4}$ . On the other hand, PISTG 3.5 is also the Stage 0 reduction of PISTG 8.1 (see Fig. 8.1 in Sec. 8).  $G_2^{3.5}$ , defined above, does have a Stage 0 counterpart  $G^{8.1}$  defined by “if alive receive high exposure.”

*Definition.* If  $B$  is a Stage 0 reduction of PISTG  $A$  and each  $G^B$  ( $G^A$ ) is the Stage 0 counterpart of some  $G^A$  ( $G^B$ ), we say  $B$  is a  $B(A)$ -complete Stage 0 reduction of  $A$ . A Stage 0 reduction is  $AB$ -complete if it is  $A$ -complete and  $B$ -complete.

*Definition.* If  $B$  is a  $B$ -complete Stage 0 reduction of PISTG  $A$  such that each  $G^B$  has a unique counterpart  $G^A$ , we say that  $B$  is the unique  $B$ -complete Stage 0 reduction of  $A$ .

*EXAMPLE.* PISTG 3.5 is a unique  $B$ -complete Stage 0 reduction of PISTG 8.1 although it is not  $A$ -complete, since any  $G^{8.1}$  for which the exposure to be received at  $t_s$  depends on employment status at  $t_s$  has no Stage 0 counterpart  $G^{3.5}$ . As we have seen above, PISTG 3.5 is not a  $B$ -complete Stage 0 reduction of PISTG 3.4. PISTG 3.5 is a  $B$ -complete stage 0 reduction of PISTG 8.3 (see Fig. 8.3), but is not a unique  $B$ -complete Stage 0 reduction. This is because the  $G^{3.5}$  defined by “if alive, receive high exposure” is the counterpart of two different generalized treatments of PISTG 8.3—“if alive, receive high exposure and  $c$  (i.e. smoke)” and of “if alive, receive high exposure and  $\bar{c}$ ”.

PISTG 7.1b is an  $AB$ -complete Stage 0 reduction of PISTG 7.1a. But, suppose in Fig. 7.1a the label on the inferior internodal line arising from the inferior node at  $t_2$  were  $H$  rather than 0, and its Stage 0 reduction, 7.1b, had a second internodal line, labelled by  $H$ , arising from the inferior right circumference point at  $t_2$ . In that case, 7.1b would be a  $A$ -complete (but not  $B$ -complete) Stage 0 reduction of PISTG 7.1a, since  $G^{7.1b}$  characterized by “if  $l$  at  $t_2$ , receive zero exposure; and if  $l$  at  $t_2$ , receive high exposure” would have no counterpart  $G^{7.1a}$ .

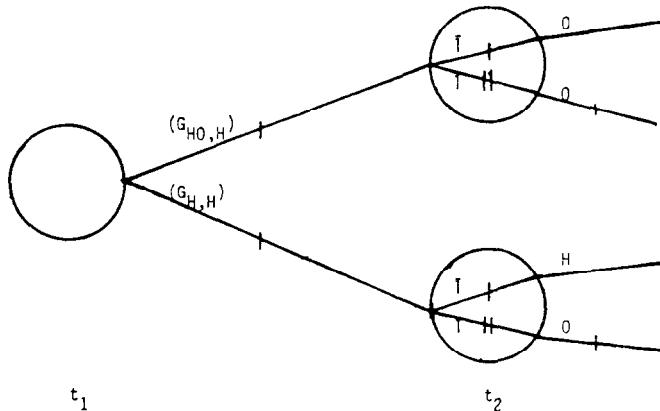


Fig. 7.1a. The fundamental MPISTG of an ordinary designed randomized trial:  $G_{H0}$  = treatment protocol of high exposure until  $t_2$ , zero exposure thereafter,  $G_H$  = treatment protocol of continuous high exposure,  $H$  = high exposure,  $I$  = at work,  $l$  = off work,  $0$  = zero exposure.

*Definition:* *Highlighted subgraph representing a function  $m^A$ .* We can represent any  $m^A$  by a highlighted subgraph of PISTG  $A$  as follows. Highlight all of PISTG  $A$  except, for any  $\cdot i_s j_s$  for which  $m^A(\cdot i_s j_s) = 0$ , do not highlight any part of subgraph  $A(\cdot i_s j_s)$  (i.e. the STG beginning at  $\cdot i_s j_s$ ).

*EXAMPLE.* If  $m^A$  is a function identified with a particular  $G^A$  (see Sec. 4E), then the highlighted subgraph associated with  $m^A$  is identical with that associated with  $G^A$ . In general, different functions  $m^A$  may have the same highlighted subgraph.

*Definition.* We say  $\cdot i_s^A \in m^A$  if  $\cdot i_s^A$  lies on the highlighted subgraph represented by  $m^A$ . We define  $\cdot i_s j_s \in m^A$  similarly.

Suppose now that we have a PISTG  $A$  and a set (which may be empty)  $R$  of true assumptions concerning the identifiable parameters of the sampled superpopulation. Then,

*Definition.* PISTG  $B$  is the Stage 0 PL-sufficient reduction of PISTG  $A$  under Assumptions  $R$  if and only if  $B$  is a Stage 0 reduction of  $A$ , and (1) for each " $G_1^A$ " there exists a  $m^B$  such that  $S(t | "G_1^A") = S(t | "m^B")$  and  $p(\cdot i_s^B | "G_1^A") = p(\cdot i_s^B | "m^B")$  where

$$p(\cdot i_s^B | "G_1^A") \equiv \sum_{[\cdot i_s^B] \subset [\cdot i_s^A]} p(\cdot i_s^A | "G_1^A") \quad \text{for } \cdot i_s^A \in "G_1^A"$$

(2) given that Assumptions  $R$  are known to hold *a priori* and given the data represented



Fig. 7.1b. A Stage 0 reduction of MPISTG 7.1a:  $H$  = high exposure,  $I$  = at work,  $l$  = off work,  $0$  = zero exposure.

by PISTG A (i.e. the data that would be available if PISTG A were an MPISTG A), then for any known function  $m^B$  the PL-sufficient statistic for  $S(t_k \mid "m^B")$  and  $p(\cdot i_k^B \mid "m^B")$  are  $N[\cdot i_s j_s^B]$ ,  $N[\cdot i_s j_s^B(t_{s+1})]$ ,  $N[\cdot i_s^B]$  for the  $\cdot i_s j_s^B$  and  $\cdot i_s^B$  on the highlighted subgraph  $m^B$ , ( $k \geq s$ ), where  $N[\cdot i_s^B]$  is the number of sampled subjects in  $[\cdot i_s^B]$ .

*Definition.* A Stage 0 PL-sufficient reduction  $B$  of PISTG A under Assumption R is A-complete ( $B$ -complete) if for each  $G^A$  ( $G^B$ ), there exists a  $G^B$  ( $G^A$ ) such that  $D(t \mid "G^A") = S(t \mid "G^B")$ , and  $p(\cdot i_s^B \mid "G^A") = p(\cdot i_s^B \mid "G^B")$ .

*Definition.* A Stage 0 PL-sufficient reduction is  $AB$ -complete if it is A-complete and  $B$ -complete.

*Remark.* The importance of Stage 0 PL-sufficiency is that if MPISTG B is the A-complete Stage 0 PL-sufficient reduction of R OCISTG A under some assumptions R, then the  $G$ -causal parameters of A are identifiable and can be estimated without loss of information from data on MPISTG B, providing assumptions R hold.

EXAMPLE 1. If Assumptions R constitute the “empty set” then the Stage 1 reduction of PISTG A is the  $AB$ -complete Stage 0 PL-sufficient reduction of PISTG A.

EXAMPLE 2. Consider the R MSCISTG associated with an ordinary designed randomized trial whose “fundamental MPISTG” is shown in Fig. 7.1a, where being off work,  $l$ , is “being off protocol”. We assume that the treatment protocols  $G_H$  and  $G_{H0}$  are assigned at  $t_1$  and contain information on the treatments (exposures) to be received at  $t_1$  and  $t_2$  (but no other times).

The crosshatches on Fig. 7.1a are to be interpreted as follows. If in generation  $t_s$ ,  $\gamma(\cdot i_s) = \gamma(\cdot i'_s)$  is known *a priori*, intranodal lines  $\cdot i_s$  and  $\cdot i'_s$  are marked with an identical number of crosshatches. Similarly, if  $S(\cdot i_s j_s) = S(\cdot i'_s j'_s)$ , internodal lines  $\cdot i_s j_s$  and  $\cdot i'_s j'_s$  are marked with an equal number of crosshatches. No significance is attached to the fact that lines in different generations are marked with an equal number of crosshatches or that an intranodal line and an internodal line in the same generation have an equal number of crosshatches. The pattern of crosshatches seen in Fig. 7.1a represents the discrete time version of the identifiable temporal assumption (see Sec. 2) applied to the R MSCISTG representing the ordinary randomized trial. The discrete time version of Theorem B1 in Appendix B shows that we can merge the intranodal lines and internodal lines in the same generation that have an equal number of crosshatches to give the  $AB$ -complete Stage 0 PL-sufficient reduction MPISTG 7.1b under the identifiable temporal assumptions, where, for example,  $S(t \mid G_{H(t_1)}^{7.1a}) = S(t \mid "G_{H(t_2)}^{7.1b}")$  where “ $G_{H(t_2)}^{7.1b}$ ” is the generalized treatment of the MPISTG 7.1b characterized by “if  $l$  at  $t_2$ , receive high exposure at  $t_2$ ”. (Theorem F1 of Appendix F generalizes Theorem B1 and can also be used to show that MPISTG 7.1b is a Stage 0 PL-sufficient reduction.)

If data on  $G_{H0}$  and  $G_H$  are not available, Fig. 7.1a is an R OSCISTG. Nonetheless, if the identifiable temporal assumption holds, the  $G$ -causal parameters of the R OSCISTG 7.1a are identifiable based on MPISTG 7.1b.

Given that 7.1a is R SCISTG, is 7.1b an R SCISTG (or even an SCISTG)? In general, the Stage 0 reduction of an SCISTG (or CISTG) need not itself be an SCISTG (CISTG).

EXAMPLE. Suppose MPISTG 3.4 is an MCISTG such that  $H$  is the treatment “walk up three flights of stairs” and 0 is the treatment “remain sedentary”. Suppose some

individuals who are  $l$  at time  $t$  are paralyzed at that time. Then clearly the Stage 0 reduction, MPISTG 3.5, is not an MCISTG, since  $l$  individuals at  $t$  cannot receive treatment  $H$  at that time.

Lemma F1 gives a sufficient condition for the Stage 0 reduction of an SCISTG to be an SCISTG. Lemma F2 gives a sufficient condition for the Stage 0 reduction of an R SCISTG to be an R SCISTG. (A necessary condition cannot be given because of insufficient structure in our formal definition of CISTG. In a separate paper, we modify the definition of CISTG so as to provide the additional structure necessary to determine whether or not a particular Stage 0 reduction of a CISTG is itself a CISTG.) By Lemmas F1 and F2, the MPISTG in Fig. 7.1b will be an R MSCISTG if the nonidentifiable temporal assumptions hold. (Note that Lemma F2 is a generalization of the discrete time versions of Lemma B1 and Lemma 2.1.)

### B. The isomorphism between the approaches in Sections 2 and 3

When we are interested in estimating the  $G$ -causal parameters of a given FR MCISTG  $B$ , we only require the knowledge that it is an R SCISTG  $B$  (see Sec. 4). We can view any R SCISTG  $B$  (whether from an observational study or from an alternative designed randomized trial) as being the  $AB$ -complete Stage 0 PL-sufficient reduction of a hypothetical ordinary designed double-blind randomized trial in which the treatment protocol assigned at  $t_1$  specifies the treatment to be received at  $t_s$ , possibly conditional on the value of time-dependent covariates measured after start of follow-up. This is proved in Theorem 7.1 below. The idea is that we can view the set of generalized treatments of the R MSCISTG,  $G^B = \{G^B = \{j_s(\cdot i_s), j_s \neq 0\}\}$ , as the possible treatment protocols of a hypothetical ordinary designed randomized trial, one of which is assigned at random at  $t_1$  to each study subject. A treatment protocol  $G_i^B$ , assigned in the hypothetical trial at  $t_1$  gives, at each time  $t_s$ , the planned treatment  $j_s$  for an individual surviving to  $t_s$  with covariate history  $\cdot i_s^B$ . We can suppose treatment  $G_i^B$  is assigned at  $t_1$  in the hypothetical trial with probability  $p_i = w_i^B(m^B, t_s)$  for the function  $m^B(\cdot i_s j_s^B) = \gamma(\cdot i_s j_s^B)$  (see Sec. 4E). Individuals in a particular  $\cdot i_s^B$  in MPISTG  $B$  from which only one internodal line originates can be viewed as all having been assigned (in the hypothetical trial) to receive the same treatment at  $t_s$ . Alternately, such individuals can be seen as being off protocol at  $t_s$ . In that case, if two or more internodal lines arise from  $\cdot i_s j_s i_{s+1}^B$  in MPISTG  $B$ , the individuals in  $\cdot i_s j_s i_{s+1}^B$  must be viewed as having returned to protocol at  $t_{s+1}$ .

**EXAMPLE.** R MSCISTG 3.4 can be viewed as the  $AB$ -complete Stage 0 PL-sufficient reduction of a double-blind ordinary designed randomized trial in which individuals off work are off protocol and  $G^{3.4}$  is their planned exposure. If they return to work later they must be assumed to have returned to protocol. Equivalently, we may view R MSCISTG 3.4 as the  $AB$ -complete Stage 0 PL-sufficient reduction of a double-blind ordinary randomized trial in which, as part of protocol, all individuals were assigned at  $t_1$  to received zero exposure at  $t_s$  if off work at  $t_s$ .

**THEOREM 7.1.** Suppose it were the case that each individual in the study represented by an MPISTG  $B$  had been assigned a treatment protocol  $G_k^B$  at random at  $t_1$  with the probabilities described in the paragraph above. Suppose that  $G_k^B$  were not recorded for data analysis. Let R OSCISTG  $A$  of this ordinary designed randomized trial be defined by the intra- and internodal lines  $\cdot i_s^A$  and  $\cdot i_s j_s^A$  where each  $[\cdot i_s^A] = [\cdot i_s^B] \cap [G_k^B]$  for some  $\cdot i_s^B, G_k^B$  (where  $[G_k^B]$  is the set of individuals assigned  $G_k^B$ ). Similarly,  $[\cdot i_s j_s^A] = [\cdot i_s j_s^B] \cap [G_k^B]$ . Suppose further that, because of double blinding, the generalized nonidentifiable

temporal assumption holds for R OSCISTG A, that is,

$$p[\cdot i_s^B \mid \cdot i_{s-1} j_{s-1}^B(t_s), G_k^B, i] \text{ and} \quad (7.1)$$

$$p[D > t_{s+1} \mid \cdot i_s j_s^B, i, G_k^B] \text{ do not depend on } G_k^B. \quad (7.2)$$

Then

- (1) MPISTG B is an R MSCISTG;
- (2) The generalized identifiable temporal assumption holds for R OSCISTG A (i.e. the lack of dependence on  $G_k^B$  holds in Eqs. (7.1) and (7.2) without conditioning on  $i$ ).
- (3) MPISTG B is the AB-complete Stage 0 PL-sufficient reduction of R OSCISTG A under the generalized identifiable temporal assumption.

*Proof.* (1) and (2) follow directly from Lemmas F1 and F2. (3) follows by noting that (2) implies that the suppositions of Theorem F1 in Appendix F are met. It is easy to check that AB-completeness holds.

If some node of MPISTG B has (at least) two right circumference points from which (at least) two internodal lines arise, it is not necessary (although it is sufficient) that the  $G_i^B$  were assigned with the probabilities  $p_i = w_i^B(m^B, t_s)$  given above in order for Theorem 7.1 to hold. Nevertheless, certain randomization schemes, i.e.  $p_i$ s, will be incompatible with the observed MPISTG B. In fact, one can suggest a possible randomization scheme and then empirically test whether the observed MPISTG B is compatible. In Sec. 8D.3 we use this device to advantage.

## 8. CIRCUMSTANCES UNDER WHICH STANDARD ANALYSES ARE VALID

### A. Circumstances in which data on time-dependent covariates may be ignored (lack of confounding)

We have noted that, in standard analyses of occupational mortality studies, employment history is generally ignored and the population parameter  $RR[E(t)] = \gamma_D(t + \Delta t \mid E(t))/\gamma_D(t + \Delta t \mid E(t) = 0)$  estimated. In point exposure studies, conditions under which one may ignore data on a covariate have been thoroughly examined under the rubric of confounding[15]. We consider the appropriate generalization of such conditions to studies with time-dependent exposures and covariates. First we will review the concept of confounding in point exposure studies.

#### A.1 Confounding in point exposure studies

In a point exposure study with a fixed follow-up period, a time-independent covariate  $L$ , measured at start of follow-up, is defined to be a causal confounder if the crude population risk difference (ignoring  $L$ ) differs from the  $G$ -causal parameter that compares the proportion of the population who would have died by end of follow-up had the entire study population been exposed to the proportion dying had it been unexposed. (To be consistent with the epidemiologic literature on confounding we have, for the moment, expressed our comparisons in terms of the probability of death rather than survival.) This definition assumes (and is vacuous unless) the stratum-specific risk differences are the  $G$ -causal parameters for their respective strata (where strata are defined by  $L$ -status). If so, a necessary condition for  $L$  to be a causal confounder is that (1) there is an exposure-covariate association in the population and (2)  $L$  is a risk factor for disease in either the exposed or unexposed population. This result differs slightly from the standard definition

in the epidemiologic literature which defines  $L$  to be a confounder when it is both associated with exposure and a risk factor in the unexposed. This difference reflects the fact that, in the standard epidemiologic literature on confounding, the implicit causal parameter of interest is the  $G$ -causal parameter representing the effect of exposure on the subset of the population that was exposed in the observed study rather than on the whole population[16]. If there is no residual confounding within levels of  $L$ , the  $G$ -causal parameter associated with the exposed population is the internally standardized risk difference with weights taken from the exposed population, while the  $G$ -causal parameter associated with the entire population is the standardized risk difference with weights taken from the entire population. When exposure is time-dependent and risk factors (e.g. employment status) determine future exposure, the implicit causal parameter of interest in the epidemiologic literature on confounding is not usefully generalized in a straightforward manner. For example, suppose MPISTG 3.4 was an FR MCISTG; and consider the subgroup of the study population who, in the observed study, received high exposure at work until end of follow-up. Suppose one wished to compare this highly exposed subgroup's survival probability when unexposed (throughout the study) to its observed survival probability (which is one). This parameter is not identifiable from FR MCISTG 3.4. Rather, it would seem that the appropriate identifiable generalization of the implicit causal parameter of interest would be a comparison of the observed survival (or mortality) curve of the entire study population with the survival curve that would have been observed had all study subjects received zero exposure throughout (i.e. had been treated with "if at work, receive zero exposure"). Such a comparison measures the potential impact of intervention. In a point exposure trial, in Miettinen and Cook's notation, this parameter (when expressed in terms of mortality) is  $(O-E)$  divided by the total number of study subjects while the internally standardized risk difference is  $(O-E)$  divided by the number of exposed study subjects.

#### A.2 Confounding in survival studies with time-dependent exposures and covariates

We shall consider the appropriate generalizations of the above conditions for confounding in studies with sustained exposure periods. We first give a number of definitions.

*Definition.*  $L$  is an (independent) population risk factor for death controlling for exposure if

$$\gamma_D[t + \Delta t | L_1(t), E(t)] \neq \gamma_D[t + \Delta t | E(t), L_2(t)] \quad (8.1)$$

for some  $L_1(t)$ ,  $L_2(t)$ ,  $E(t)$ .

*Definition.*  $L$  is a population predictor of future exposure if

$$p[E(t + \Delta t) | L_1(t), E(t), D > t + \Delta t] \neq p[E(t + \Delta t) | L_2(t), E(t), D > t + \Delta t] \quad (8.2)$$

for some  $L_1(t)$ ,  $L_2(t)$ ,  $E(t)$ .

*Definition.*  $L$  is a population predictor of exposure if

$$\begin{aligned} p[E(t + \Delta t) | L_1(t + \Delta t), E(t), D > t + \Delta t] \\ \neq p[E(t + \Delta t) | L_2(t + \Delta t), E(t), D > t + \Delta t] \end{aligned} \quad (8.3)$$

for some  $L_1(t + \Delta t)$ ,  $L_2(t + \Delta t)$ ,  $E(t)$ .

*Definition.* Current  $l$ -status is a predictor of exposure if

$$\begin{aligned} p[E(t + \Delta t) | L(t), E(t), l_1(t + \Delta t), D > t + \Delta t] \\ \neq p[E(t + \Delta t) | L(t), E(t), l_2(t + \Delta t), D > t + \Delta t] \end{aligned} \quad (8.4)$$

for some  $l_1(t + \Delta t)$ ,  $l_2(t + \Delta t)$ ,  $E(t)$ ,  $L(t)$ .

*Definition.*  $L$  is a causal risk factor for death controlling for exposure if

$$\gamma_D[t + \Delta t | E(t), L_1(t), i] \neq \gamma_D[t + \Delta t | E(t), L_2(t), i] \quad (8.5)$$

for some individual  $i$  and some  $E(t)$ ,  $L_1(t)$ ,  $L_2(t)$ .

*Definition.*  $L$  is a causal risk factor for exposure if

$$p[E(t + \Delta t) | L_1(t), E(t), i] \neq p[E(t + \Delta t) | L_2(t), E(t), i] \quad (8.6)$$

for some  $L_2(t)$ ,  $L_1(t)$ ,  $E(t)$ ,  $i$ .

Note that Eqs. (8.5) and (8.6) do not require that every individual could receive both  $L_1(t)$  and  $L_2(t)$ . For example, Eqs. (8.5) and (8.6) might hold even if MCISTG 3.3 were not an MCISTG. Equations (8.5) and (8.6) are Eqs. (8.1) and (8.2), except at the individual level (and thus nonidentifiable). In Eq. (8.6) since we do not condition on the event  $D > t + \Delta t$ , we have assumed (consistent with our definition of a CISTG in Sec. 4) that an individual's "covariate history" is well defined even after death. The above definitions have symmetric versions with the role of  $L$  and  $E$  interchanged. We shall see that it is Defs. (8.1) and (8.3) that are the natural generalizations of the concepts (developed for point exposure studies) of a risk factor and of a covariate-exposure association.

**EXAMPLE.** Table 5 in Appendix D shows that, in our arsenic data,  $L$  is an independent population risk factor for death (provided our statistical model was correctly specified). Also,  $L$  is tautologically a predictor of future exposure in our data, since individuals off work receive no measured exposure.

A natural generalization of a point exposure study "with no residual confounding within levels of the covariate" is FR MCISTG 8.1 where in Fig. 8.1 the covariate is  $L$  history. An even more natural generalization would have been to allow both levels of  $l$  to occur at  $t_1$  in FR MCISTG 8.1, which we shall call the  $t_1$ -modified FR MCISTG 8.1. In that case, a point exposure trial would be precisely equivalent to a study represented by the  $t_1$ -modified FR MCISTG 8.1, when follow-up ended at  $t_2$ ; and  $S(t_2, G_H^{8.1}, G_0^{8.1})$  with the two generalized treatments being "if alive receive high exposure" and "if alive receive zero exposure" would be the population  $G$ -causal parameter of the point exposure trial. All our results will apply to FR MCISTG 8.1 as is and when  $t_1$ -modified. In contrast, FR MCISTG 8.2 represents a study in which there exists residual confounding within levels of  $L$ , in that individuals with a given work and exposure history through  $t_{s-1}$  who are  $l$  at  $t_s$  do not receive exposure at random at that time. Note also that the finer MCISTG derived from MCISTG 8.1, in which, for all nodes, all four internodal lines leaving the node arise from a single right circumference point, would not be the appropriate gener-

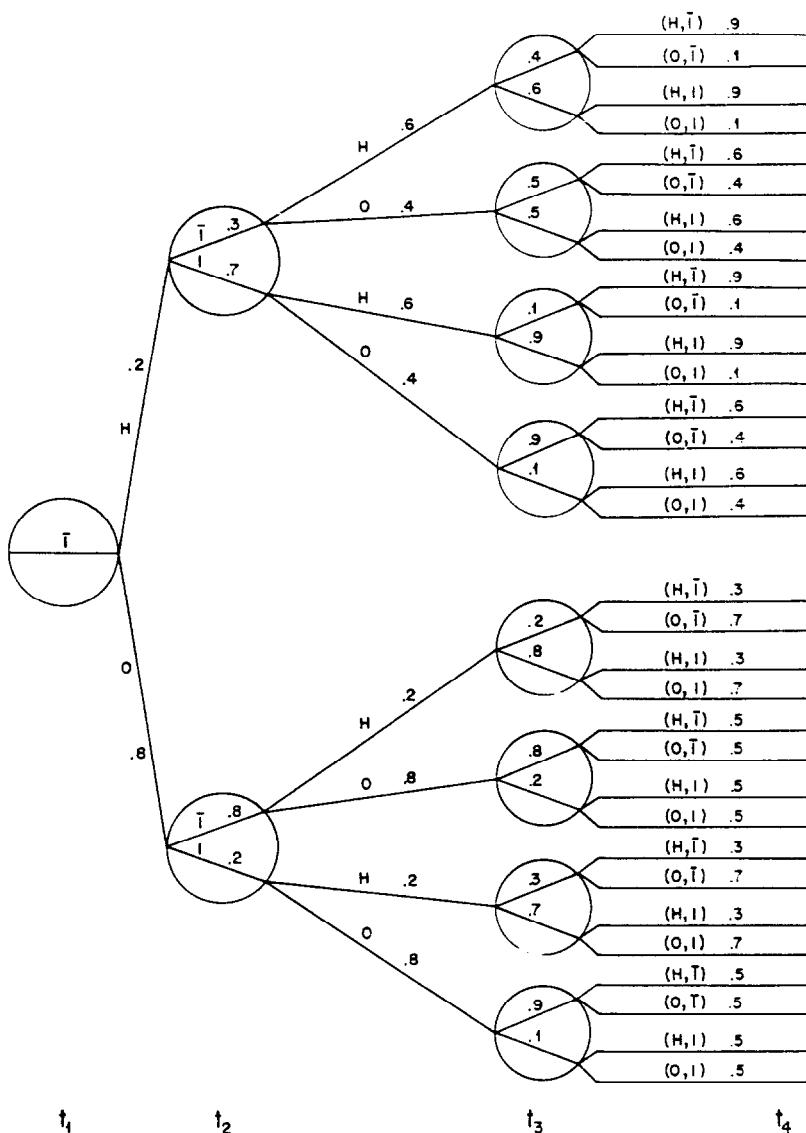


Fig. 8.1. An MPISTG with  $\gamma(\cdot, i_s)$  and  $\gamma(\cdot, i_s j_s)$  displayed on intra- and internodal lines, respectively.

alization for several reasons, one among which is that in our investigation of confounding in point exposure studies we do not require the covariate  $L$  to be a treatment.

Under what circumstances can (some)  $G$ -causal parameters of FR MCISTG 8.1 be estimated in the absence of data on  $L$ , that is, from the data available from the MPISTG 3.5? Lemmas 8.1 and 8.2 demonstrate that two sufficient circumstances are (1)  $L$  is not a predictor of exposure and (2)  $L$  is not an independent population risk factor for death.

**LEMMA 8.1.** If  $L$  is not an independent population risk factor for death, then: (1) MPISTG 3.5 is a  $B$ -complete Stage 0 PL-sufficient reduction of FR MCISTG 8.1; (2) For any  $G_1^{8.1}$  that assigns to each individual alive at  $t$  an exposure level  $e_1(t)$  (irrespective of their  $l$  status at  $t$ ),  $S(t | G_1^{8.1}) = S(t | "G_1^{3.5}") \equiv S(t | E_1(t - \Delta t))$  [as defined in Eq. (6.3)]

where “ $G_1^{3.5}$ ” is the stage 0 counterpart of “ $G_1^{8.1}$ ”. (3) For any  $G^{8.1}$  that assigns, at some time  $t$ , exposures to individuals on the basis of their  $l$  status,  $S(t | G^{8.1})$  is, in general, nonidentifiable in the absence of data on  $L$ . (4)  $S(t, “G_1^{3.5}”, “G_2^{3.5}”) \equiv 0 \Leftrightarrow S(t, “G_1^{8.1}”, “G_2^{8.1}”) \equiv 0$ .

*Proof.* If Eq. (8.1) is false, then Assumptions  $R$  of Theorem F1 hold for MPISTG 8.1 and its Stage 0 reduction MPISTG 3.5. (1) follows from the fact that MPISTG 3.5 is  $B$ -complete and Corollary F1. Furthermore, the  $G^{8.1}$  described under (2) are precisely those  $G^{8.1}$  with Stage 0 counterparts in  $G^{3.5}$ . (3) follows by noting that the functions  $m^{3.5}$  for which  $S(t | m^{3.5}) = S(t | G^{8.1})$  will depend on the distribution of  $L$  for those  $G^{8.1}$  without a Stage 0 counterpart. Finally, (4) follows from PL-sufficiency.

*Remark.* Except for  $B$ -completeness, Lemma 8.1 also holds for FR MCISTG 3.4 replacing FR MCISTG 8.1.

LEMMA 8.2. If  $L$  is not a population predictor of exposure and MPISTG 8.1 is an FR MCISTG then (A) conclusion (2) of Lemma 8.1 holds; but (B) if data on  $L$  is available, then, even when  $L$  is known *a priori* not to be a predictor of exposure, the NPMLE of  $S(t | E(t - \Delta t))$  will depend on the data through  $L$ ; (C) conclusion (4) of Lemma 8.1 is false; (D) for any  $G_2^{8.1}$  that assigns at time  $t' < t$  different exposures to  $\bar{l}$  individuals than to  $l$  individuals, it is not necessary that  $S(t | “G_2^{8.1}”) = S(t | “m^{3.5}”)$  for some  $m^{3.5}$ ; (E) conclusion (4) of Lemma 8.1 modified so the implication arrow points only toward the  $G$ -causal parameter of MPISTG 3.5 still holds (i.e. valid but not consistent tests of the  $G$ -null hypothesis of FR MCISTG 8.1 are possible without data on  $L$ ).

*Proof.* (A) and (E) follow by noting that if Eq. (8.3) is false, the suppositions of Theorem F2 are satisfied for all  $G^{8.1}$  with Stage 0 counterparts  $G^{3.5}$ . (B)–(D) are true even in the special case of a point exposure trial in which the dichotomous covariate and exposure are measured once at start of follow-up. For example, in a point exposure study, (B) can be rephrased as “if there is an exposure-covariate association in the data, but not in the population, the NPMLE of the standardized risk difference (with weights chosen from the entire population) differs from the empirical crude risk difference.” For proof see Refs. [15] and [17]. [A key idea is that knowledge that Eq. (8.3) is false is a prior restriction that crosses a cut, as described in Appendix A.] Points (C) and (D) will hold in a point exposure trial when the two stratum-specific risk differences are of opposite sign and the crude risk difference is zero.

(A) of Lemma 8.2 is false unless (1)  $L$  is not a population predictor of future exposure (as would be the case in a point-exposure study), and (2) current  $l$ -status is not a predictor of exposure. (1) and (2) hold if and only if  $L$  is not a predictor of exposure. We use Fig. 8.1 to clarify the distinction between Eqs. (8.3) and (8.4).

The fractions written over the internodal lines in Fig. 8.1 are  $p(\cdot i_s j_s | \cdot i_s)$ . In Fig. 8.1,  $L$  is not a population predictor of exposure since, e.g. the ratio of individuals receiving high to zero exposure at  $t_3$  is 0.9/0.1 at the first, second, fifth and sixth right circumference points when reading from the top. If the ratio at the first and second points differed from that at the fifth and sixth, then  $L$  would be a population predictor of exposure, but current  $l$ -status would not be a predictor of exposure.

If there is residual confounding within levels of  $L$  (e.g. if Fig. 8.2 was our finest FR MCISTG), then, even when  $L$  is neither a population risk factor nor a predictor of exposure,  $S(t + \Delta t | E_1(t)) - S(t + \Delta t | E_2(t))$  may not be the  $G$ -causal parameter of any MCISTG.

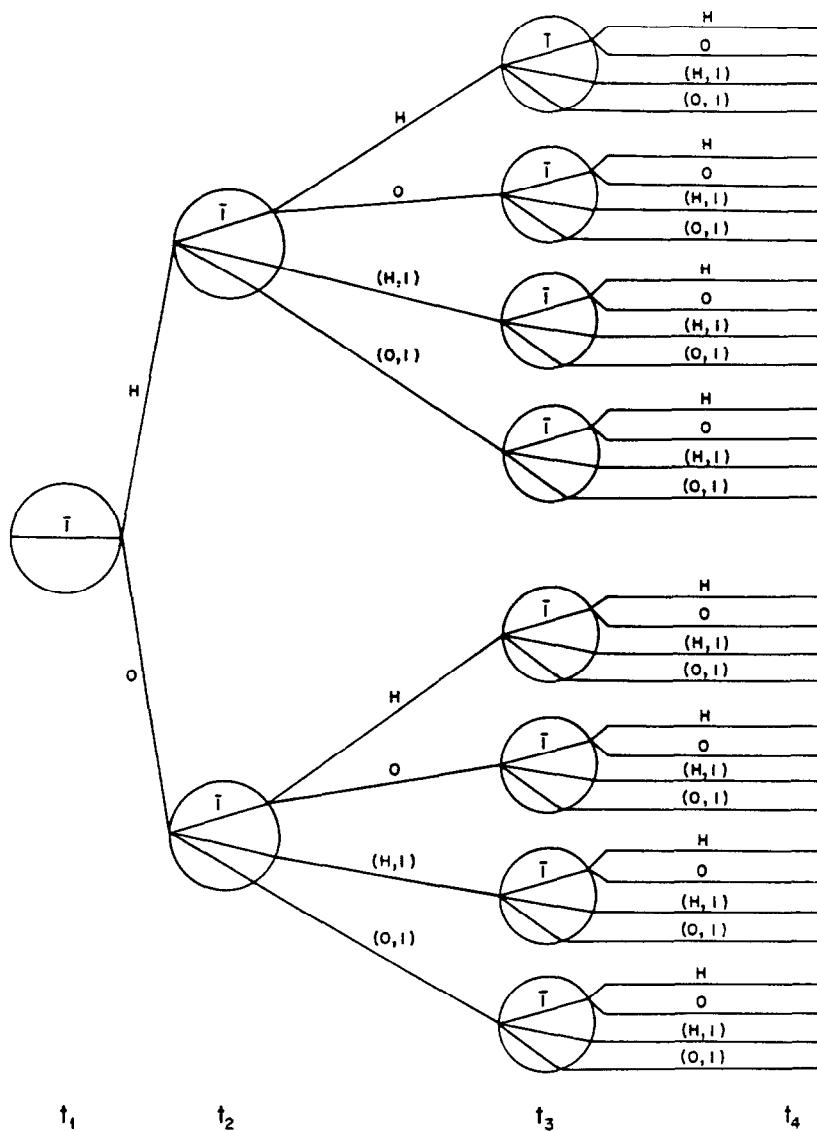


Fig. 8.2. An MPISTG.

Finally, we might wonder, if the conditions of Lemma 8.1 or 8.2 are satisfied and MPISTG 8.1 is an FR MCISTG, will MPISTG 3.5 be an FR MCISTG? In general, it will be only if MPISTG 3.5 is a causal melded reduction of FR MCISTG 8.1 as defined in Appendix F. See Theorems F3 and F4. In the next section we discuss some additional difficulties in causal interpretation that occur when FR MCISTG 3.4 is our finest FR MCISTG.

### A.3 A difference between FR MCISTGs 3.4 and 8.1

In occupational mortality studies individuals are (presumed) to be unexposed while off work. As such, FR MCISTG 3.4 rather than 8.1 would represent the data. If  $L$  is not a risk factor in the unexposed, it follows from Lemma 8.1 that we can test the  $G$ -null hypothesis for FR MCISTG 3.4 based on MPISTG 3.5. Nonetheless (in contrast to the FR

MCISTG 8.1),

$$S(t | E_1(t)) - S(t | E_2(t)) \quad (8.7)$$

will not necessarily have a causal interpretation (when nonzero) since  $S(t | E_1(t))$  will, in general, not equal any  $S(t | G^{3.4})$  if  $E_1(t) \neq 0$ . Given that exposure is a nonrisk factor in the unexposed and MCISTG 3.4 is an FR MCISTG, who do we tend to believe that Eq. (8.7) has a causal interpretation? One justification for such a belief is a conjunction of the following theorem with the following subjective belief.

**THEOREM 8.1.** If MPISTG 3.3 is an FR MCISTG and  $L$  is not an independent causal risk factor, then MPISTG 3.5 is an FR MSCISTG and  $L$  is a population nonrisk factor controlling for  $E(t) \equiv 0$ .

*Proof.* A direct application of Lemma F2.

*Subjective belief of many investigators.* If  $L$  was a causal risk factor and or MPISTG 3.3 was not an FR MCISTG, it is unlikely (though not impossible) for  $L$  to be a population nonrisk factor controlling for  $E(t) \equiv 0$ . This belief reflects the fact that it would take a precise balancing of the healthy worker selection effect by the direct causal effects of  $L$ , for  $L$  to be a population nonrisk factor. Therefore, one might have low subjective probability that such a “balancing act occurred.” Armed with this belief, we would consider that Eq. (8.7) represented a  $G$ -causal parameter of FR MSCISTG 3.5 when  $L$  is an empirical nonrisk factor.

We now ask, In what circumstances might we believe Eq. (8.7) is causal if either MPISTG 3.3 is not believed to be an MCISTG (because disabled individuals cannot, even conceptually, return to work), or  $L$  is a causal risk factor and the unusual balancing act described above occurred? Suppose one believed that, at least conceptually, individuals off work could have received high exposure (e.g. by having arsenic dust pumped into their homes). Then one could view FR MCISTG 3.4 as the special case of FR MCISTG 8.1 in which the probability of receiving high exposure while off work was zero. The survival curve of the controlled trial defined by  $G^{8.1}$ , “if alive, receive high exposure”, is well defined, although the curve will not be identifiable without further nonidentifiable assumptions when no individuals actually received high exposure off work.

One might be willing to assume if, in the observed study,  $L$  is not a risk factor among the unexposed,  $L$  would not be a risk factor in the hypothetical study represented by  $G^{8.1}$ , “if alive, receive high exposure”. If so, Eq. (8.7) would be a  $G$ -causal parameter of  $G^{8.1}$ . But unless MPISTG 3.3 is an FR MCISTG, and  $L$  is a causal nonrisk factor, Eq. (8.7) would not represent the effect of exposure controlling for employment history.

#### B. Circumstances under which time-dependent covariates may be adjusted for in standard fashion

Suppose in the study represented by MPISTG 3.3,  $L$  is an independent population risk factor. If so, a common practice is to adjust or stratify on  $L$ -history (i.e. to determine whether exposure is an independent population risk factor controlling for  $L$ ). This provides a valid test of whether exposure is a causal risk factor, controlling for the intermediate variable employment history, provided MCISTG 3.3 is an FR MCISTG. But suppose now one believes that  $L$  is not an independent causal risk factor. Then, by Theorem 8.1, MCISTG 3.3 cannot be an FR MCISTG. Nevertheless, the test of the  $G$ -null hypothesis of FR MCISTG 3.4 would be a test of whether exposure is an independent causal risk

factor (since its  $G$ -null hypothesis is implied by the sharp null hypothesis for MCISTG 3.3). In this setting, under what circumstances does testing whether exposure is a population risk factor controlling for  $L$  constitute a valid test of the  $G$ -null hypothesis for FR MCISTG 3.4? Lemma 8.3 shows that it constitutes a valid test when exposure is not a population predictor of future  $L$ -history where we use future  $L$ -history to mean future  $L$ -experience (equivalently,  $L$ -status).

**LEMMA 8.3.** If exposure is not a predictor of future  $L$ -status [i.e. Eq. (8.2) is false with the roles of  $E$  and  $L$  interchanged], then for a given “ $G_1^{3,4}$ ”:

$$(1) S(t_s + \Delta t | "G_1^{3,4}") = \sum_{L(t_s)} S((t_s + \Delta t) | E(t_s), L(t_s)) p^\Delta(L(t_s))$$

where

$$\begin{aligned} S(t_s + \Delta t | E(t_s), L(t_s)) &\equiv \prod_{t_k=t_1}^{t_s} [1 - \gamma_D(t_k + \Delta t | E(t_k), L(t_k))] \\ p^\Delta[L(t_s)] &\equiv \prod_{t_k=t_1}^{t_s} p[L(t_k) | L(t_k - \Delta t), D > t_k] \end{aligned} \quad (8.7a)$$

and  $E(t)$  is uniquely determined by  $L(t)$  and the chosen  $G_1^{3,4}$ ; the sum is over the  $2^{s-1}$  possible paths  $L(t_s)$ ; and  $L(t_k)$  is the initial part of  $L(t_s)$ .

(2)  $\gamma_D(t + \Delta t | E(t), L(t)) = \gamma_D(t + \Delta t | L(t))$  for all  $E(t)$  if and only if  $S(t, "G_1^{3,4}", "G_2^{3,4}") \equiv 0$  (i.e. valid and consistent tests of the  $G$ -null hypothesis for FR MCISTG 3.4 can be based on standard methods of assessing the effect of exposure controlling for  $L$ ).

*Proof.* Direct computation using the  $G$ -computation algorithm and the fact that exposure is not a predictor of future  $L$ -status.

**EXAMPLE.** In our arsenic data, exposure is a determinant of future  $L$ -status (see Table 5 in Appendix D), and so valid tests of the  $G$ -null hypothesis of FR MCISTG 3.4 had to be based on the  $G$ -null test.

Since  $S(t_s + \Delta t | G_1^{3,4})$  can be written as a weighted average of  $S((t_s + \Delta t | E(t_s), L(t_s))$  when MPISTG 3.4 is an FR MCISTG and exposure is not a predictor of future  $L$ -status, one might wonder whether differences in the  $S((t_s + \Delta t) | E(t_s), L(t_s))$  for fixed  $L(t_s)$  and varying  $E(t_s)$  may themselves have a causal interpretation (even when MPISTG 3.3 is not an FR MCISTG or even an MCISTG). [Note that since  $S((t_s + \Delta t) | E(t_s), L(t_s)) \equiv S(t_s | "G^{3,3}")$  for any  $G^{3,3}$  that gives treatments  $E(t_s), L(t_s)$ , such differences between the  $S((t_s + \Delta t) | E(t_s), L(t_s))$  would have a causal interpretation as the effect of exposure controlling for employment history if MPISTG 3.3 were an FR MCISTG. But this would be true even if exposure were a determinant of future  $L$ -status.] The following theorem gives an answer to our question (although it will not be appreciated by those who do not believe in the life we would have led had we not died).

**THEOREM 8.2.** If MCISTG 3.4 is FR MCISTG and (1) exposure is not a causal risk factor for  $L$  [i.e. Eq. (8.6) is false with the roles of  $E$  and  $L$  interchanged]. [Note this implies that each individual in the population has associated an  $L$ -history  $L(t_s)$  that does not depend on the exposures they received, or whether they die before  $t_s$ . This would be the case if there was an unmeasured covariate at  $t_1$  that determined  $L(t_s)$ .]

$$(2) \quad \gamma_D(t + \Delta t | E(t), L(t), D > t, L(t_S)) = \gamma_D(t + \Delta t | E(t), L(t), D > t)$$

[i.e. those dying at  $t + \Delta t$  with history  $E(t)$ ,  $L(t)$  are representative with respect to  $L(t_S)$  history]. Then, (a) exposure is not a population predictor of future  $L$ -history and (b) the  $G$ -causal parameter through  $t_s$  of the subset of the population with a given  $L(t_S)$  [i.e.  $S(t_s, G_1^{3,4}, G_2^{3,4})$ ,  $i \in L(t_S)$ ] equals  $S[t_s | E_1(t_{s-1}), L(t_{s-1})] - S[t_s | E_2(t_{s-1}), L(t_{s-1})]$  where  $L(t_{s-1})$  is the initial part of  $L(t_S)$  through  $t_{s-1}$  and  $E_1(t_{s-1})$  and  $E_2(t_{s-1})$  are uniquely determined by  $L(t_{s-1})$ ,  $G_1^{3,4}$ , and  $G_2^{3,4}$ .

Use of this theorem is much like that of Theorem 8.1. In Theorem 8.2 only conclusion (a) is identifiable. One might believe, that given MPISTG 3.4 is an FR MCISTG, it is likely that if exposure is not a population predictor of future  $L$ -history, the suppositions of the theorem would be true. If so, one would believe that whenever (a) is empirically true that conclusion (b) is likely to be true.

*Proof of Theorem 8.2.* We use the device of showing the theorem is true for the hypothetical ordinary randomized trial whose  $AB$ -complete PL-sufficient reduction is R SCISTG 3.4. In that trial,  $(\bar{D}_{G^{3,4}}, \bar{L}_{G^{3,4}}) \amalg G^{3,4}$  (using the independence notation of Dawid[18]—see proof of Theorem E1 of Appendix E) where  $\bar{D}_{G^{3,4}}$ ,  $\bar{L}_{G^{3,4}}$  are the vectors of death times and  $L$ -histories through  $t_S$  as  $G^{3,4}$  runs through  $G^{3,4}$ . Therefore,  $\bar{D}_{G^{3,4}} \amalg G^{3,4} | \bar{L}_{G^{3,4}}$ . This implies  $\bar{D}_{G^{3,4}} \amalg G^{3,4} | L(t_S)$  by supposition (1). This, in turn implies, using the  $G$  computation algorithm, that  $S(t_s | G_1^{3,4}, L(t_S)) = S[t_s | E_1(t_{s-1}), L(t_{s-1}), L(t_S)]$  which does not depend on  $L(t_S)$  by supposition (2), proving (b).

To prove (a), we note that by Bayes theorem

$$\begin{aligned} p[L(t_s + \Delta t) | L(t_s), E(t_s), D > t_s + \Delta t] \\ = \left[ \frac{p[D > t_s + \Delta t | E(t_s), L(t_s), L(t_s + \Delta t)]}{p[D > t_s + \Delta t | E(t_s), L(t_s)]} \right] p[L(t_s + \Delta t) | L(t_s), E(t_s)]. \end{aligned}$$

Now by (2) the ratio term is unity, and by (1) and the fact MCISTG 3.4 is an FR MCISTG, the second term does not depend on  $E(t_s)$ . This proves (a). [Note that  $L(t_s + \Delta t)$  in the numerator of the ratio term is the event that an individual would have had history  $L(t_s + \Delta t)$  at  $t_1$ , irrespective of whether they were alive at  $t_s + \Delta t$ .]

Rosenbaum[19] established an essentially identical theorem in the case of a point exposure and time-dependent covariate except he effectively assumed there are no censoring or covariate history by death (i.e. that all deaths occurred only at end of follow-up), so he did not consider the necessity for (2). We notice that even if (a) is empirically false, (1) may be true for a subset of the population. If the equation in (2) is still true (whenever the left side is well defined), then (b) still holds by the previous proof. Note that the proof of (a) does not go through when (1) holds only for a subset. To better understand the implications of this observation, consider the simple example of a point exposure randomized trial in which all deaths occur at  $t_3$ , all individuals have identical values at  $t_1$  on a (time-dependent) dichotomous covariate. The covariate is measured again at  $t_2$ , and exposure is found to be a predictor of future covariate history. Nonetheless, it is always possible although never identifiable, that the empirical differences in survival probability, controlling for the level of the covariate measured at  $t_2$ , represent  $G$ -causal parameters for the effect of exposure on subsets of the population for whom exposure is not a causal risk factor for the covariate. If all deaths occurred between  $t_S$  and  $t_{S+1}$  so that there was no “censoring” of  $L$ -history by death, examination of our proof (see also Ref. [19]) shows that (a) and (b) hold without requiring (2). But when (1) applies to only a subset of pop-

ulation, (2) is necessary for (b) to hold, even when, as in our example, there is no censoring of  $L$ -history by death.

*C. A circumstance under which the causal effect of exposure controlling for  $L$  may be estimated in the absence of data on  $L$*

Suppose the finer PISTG formed from PISTG 8.1 by having the four internodal lines leaving each node arise from a single point on the right circumference, which we will call PISTG F8.1, were an FR MCISTG. Suppose also that  $L$  were an independent population risk factor for death. Then, a test of the null hypothesis of no direct exposure effect for FR MCISTG F8.1 is available by testing whether exposure is a population risk factor for death controlling for  $L$ -history. Under what circumstances may we validly test the above null hypothesis in the absence of data on employment history. The following theorem shows that we may do so when  $L$  is not a predictor of exposure history and exposure is not a predictor of future  $L$ -history.

**THEOREM.** If (1)  $L$  is not a predictor of exposure history, (2) exposure is not a predictor of future  $L$ -history, and (3) exposure is not a population risk factor for death controlling for  $L$ -history, then  $S(t, "G_1^{3.5}", "G_2^{3.5}") \equiv 0$ , i.e.

$$S(t + \Delta t | E(t)) \text{ does not depend on } E(t). \quad (8.8)$$

*Proof.* From Lemma 8.3 we know that when exposure is not a predictor of future  $L$ -history and exposure is not an independent population risk factor for death, then  $S(t, "G_1^{3.4}", "G_2^{3.4}") \equiv 0$ . But, using Lemma 8.2, if  $L$  is not a predictor of exposure,  $S(t, "G_1^{3.5}", "G_2^{3.5}") \equiv 0$ . Therefore, a valid test of (3) may be based on Eq. (8.8).

*Remark.* The theorem is true if  $L$  is interchanged with exposure in the statements of (1) and (2).

*D. Estimating exposure effects given data on cigarette smoking history*

*D.1 Circumstances under which standard analyses are valid*

In occupational mortality studies, data on cigarette smoking history  $C(t)$  may be obtained. In such instances, the investigator will usually report an estimate of the parameters

$$\gamma_D(t + \Delta t | E(t), C_1(t))/\gamma_D(t + \Delta t | E(t) = 0, C_1(t)) \quad (8.9)$$

for various  $C_1(t)$  in an attempt to assess the direct effect of exposure controlling for cigarette smoking history. When is this attempt successful?

In general, when the healthy worker survivor effect is operative, parameters that represent the effect of exposure controlling for smoking history will only be identifiable when exposure and cigarette smoking at  $t_s$  are received at random conditional on cigarette smoking and exposure history through  $t_{s-1}$ , and employment history through  $t_s$  (i.e. Fig. 8.3 is an FR MCISTG). (In Fig. 8.3, for simplicity, we have supposed that smoking status at  $t_s$  is recorded only as currently smoking or not currently smoking. In actual practice we would normally record at least three to four smoking levels.)

Any  $G$ -causal parameters of MCISTG 8.3 that represent the effect of exposure controlling for cigarette smoking history can be written

$$S(t; G_1^{8.3} = ["G_1^{3.4}", C(t_s)], G_2^{8.3} = ["G_2^{3.4}", C(t_s)]), \quad (8.10)$$

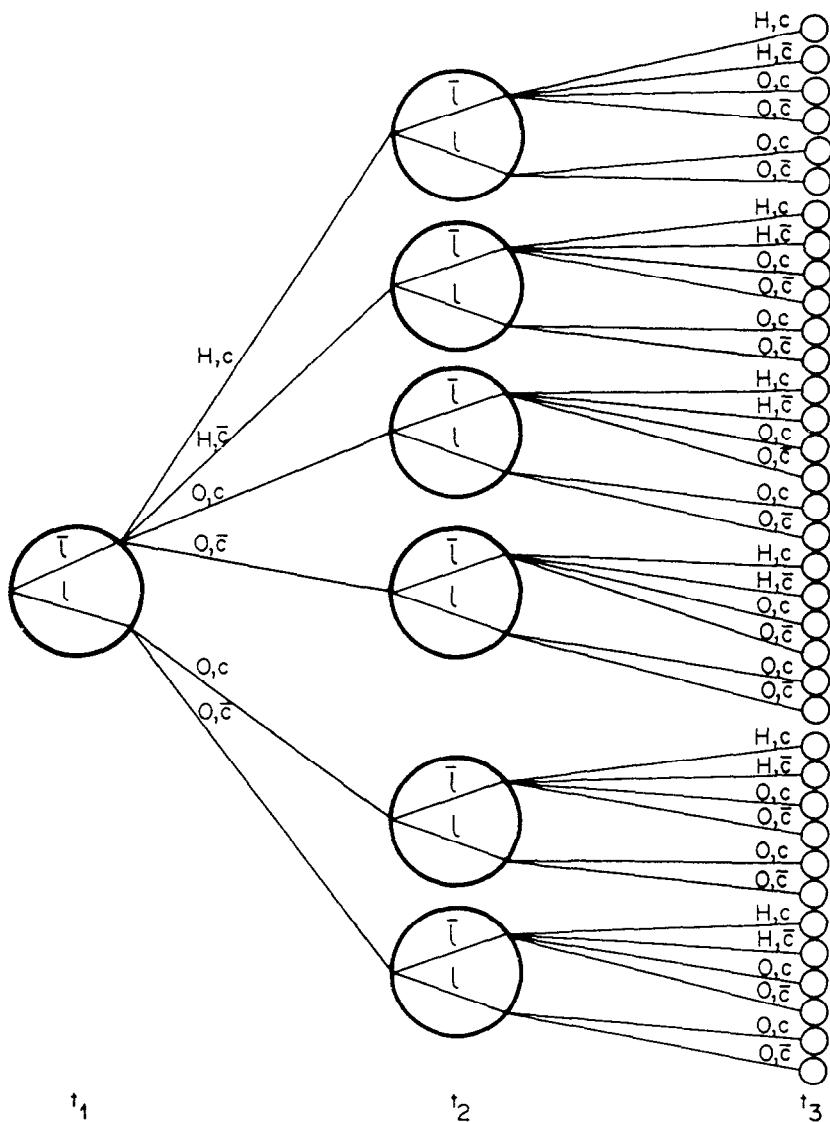


Fig. 8.3. An FR MCISTG:  $H$  = high exposure concentration,  $0$  = unexposed,  $\bar{l}$  = at work,  $l$  = off work,  $c$  = current smoker,  $\bar{c}$  = current nonsmoker.

MPISTG C8.3 is the coarser MPISTG formed from MPISTG 8.3 in which internodal lines that differ in  $c$ -status are given separate right circumference points.

MPISTG F8.3 is the finer MPISTG formed from MPISTG 8.3 in which all six internodal lines leaving each node arise from a single right circumference point.

MPISTG CF8.3 is the coarser MPISTG formed from MPISTG F8.3 in which individuals with different  $c$ -status (but neither  $l$ -status nor exposure status) are given separate right circumference points.

MPISTG SO8.3 is the Stage 0 reduction of MPISTG 8.3 in which data on  $L$ -status is absent and each node has a single intranodal line from which four internodal lines arise representing joint levels of  $c$ -status and exposure status.

where  $G_1^{8,3}$  is the generalized treatment that assigns to each individual at  $t_s$  a particular cigarette smoking status obtained from the value of the curve  $C(t_s)$  at that time (irrespective of employment or exposure history), and furthermore assigns to an individual at work a particular exposure at  $t_s$  which may depend on past exposure and employment history. Such an exposure assignment function can be characterized by a unique generalized treatment of PISTG 3.4. Note that  $G_1^{8,3}$  and  $G_2^{8,3}$  have identical cigarette smoking history assignments but different " $G^{3,4}$ 's. To motivate our choice of the parameters characterized by Eq. (8.10), consider the finer PISTG, say PISTG F8.3, formed from PISTG 8.3 by having all six internodal lines leaving each node arise from a single right circumference point. Suppose F8.3 is an MCISTG (but not an FR MCISTG because of the healthy worker survivor effect) and  $L$ -history is not a causal risk factor controlling for  $C$ - and  $E$ -history. If, in addition, exposure is not a causal risk factor controlling for  $C$ - (and, now trivially,  $L$ -) history in MCISTG F8.3, then the identifiable  $G$ -causal parameters of FR MCISTG 8.3 defined by Eq. (8.10) will be identically zero (as they must if they are to represent the effect of exposure controlling for cigarette smoking history). Note that the  $G$ -causal parameters of Eq. (8.10) would not be null when exposure had no effect controlling for cigarette smoking history if we allowed  $G_1^{8,3}$  and  $G_2^{8,3}$  in Eq. (8.10) to represent generalized treatments of MCISTG 8.3 that assign cigarette smoking behavior based on employment history. In that case, if exposure history determines  $L$ -status, a given individual could have different cigarette smoking histories in the controlled trials represented by  $G_1^{8,3}$  and  $G_2^{8,3}$ . If so, the  $G$ -causal parameters represented in Eq. (8.10) would not be null even when exposure had no effect controlling for cigarette smoking history.

We now consider the circumstances under which the parameters represented in Eq. (8.9) can be used to test the null hypothesis that all  $G$ -causal parameters of FR MCISTG 8.3 of the form given in Eq. (8.10) are identically zero (which we shall call the hypothesis of no exposure effect controlling for cigarette smoking). In Lemma 8.4, we show that valid tests can be based on Eq. (8.9) either when  $L$  is not a population risk factor controlling for cigarette smoking and exposure history, i.e.

$$\gamma_D(t + \Delta t | E(t), C(t), L(t)) = \gamma_D(t + \Delta t | E(t), C(t)) \quad \text{for all } L(t) \quad (8.11)$$

or when employment history is not a predictor of the joint distribution of exposure and cigarette smoking, i.e.

$$\begin{aligned} p[C(t + \Delta t), E(t + \Delta t) | E(t), C(t), L(t + \Delta t), D > t + \Delta t] \\ = p[C(t + \Delta t), E(t + \Delta t) | E(t), C(t), D > t + \Delta t] \end{aligned} \quad (8.12)$$

for all  $L(t + \Delta t)$ . Note that for MPISTG 8.3 (as with MPISTG 3.4) employment history is trivially a determinant of future exposure history. Therefore, for Eq. (8.12) to be other than trivially false we need to consider the modified version of MCISTG 8.3, called MCISTG EM8.3, which differs from MCISTG 8.3 only in that individuals off work can receive high (as well as zero) exposure. EM8.3 stands in the same relation to Fig. 8.3 as does Fig. 8.1 to Fig. 3.4. Therefore, when considering EM8.3 we must replace the  $G^{3,4}$  by  $G^{8,1}$  in Eq. (8.10).

Let MPISTG S08.3 be the MPISTG that is the Stage 0 reduction of MPISTG EM8.3 (and MPISTG 8.3) obtained by deleting data on employment history. MPISTG S08.3 has one intranodal line per node from which four internodal lines arise. The internodal lines reflect possible combinations of exposure and cigarette smoking status.

**LEMMA 8.4.** If Eq. (8.11) holds, then MPISTG S08.3 is the Stage 0 PL-sufficient

reduction of MPISTG EM8.3 (or of MPISTG 8.3). In addition, the null hypothesis of no exposure effect controlling for cigarette smoking holds if and only if Eq. (8.9) is always unity. If Eq. (8.12) holds, then the null hypothesis of no exposure effect controlling for cigarette smoking history implies that Eq. (8.9) is unity. Finally, if either Eq. (8.11) or (8.12) holds, then  $S(t | "G^{EM8.3}") = S[t | E_1(t - \Delta t), C_1(t - \Delta t)] \equiv S(t | "G^{S08.3}")$  for any  $G^{EM8.3}$  that assigns individuals at  $t_s$  to treatment  $[e_1(t_s), c_1(t_s)]$  [irrespective of  $L(t_s)$ ], and  $G^{S08.3}$  is its Stage 0 counterpart.

*Proof.* We use Theorems F1 and F2 exactly as we did in Lemmas 8.1 and 8.2.

## D.2 Circumstances under which cigarette smoking history is a nonconfounder

Note that it would not be uncommon for an investigator to believe that MCISTG 8.3 was not fully randomized. For example, consider a group of current smokers at  $t_{s-1}$  who are out of work at  $t_s$  with identical exposure, smoking, and employment history through  $t_{s-1}$ . It is reasonable to believe that the subgroup who left work due to (unrecorded) ill health are more likely to give up smoking at  $t_s$  than those who left work for purely socioeconomic reasons. Even among a group of smokers at  $t_{s-1}$  who remain at work at  $t_s$ , one might believe that those who continue smoking at  $t_s$  differ on unmeasured risk factors from those who choose to give up smoking. In that case, our finest FR MCISTG might be the coarser MPISTG, formed from MCISTG 8.3, that has four intranodal lines per node defined by joint levels of present smoking and employment status. Note this FR MCISTG, which we call FR MCISTG C8.3, assumes that exposure is received at random at work at  $t_s$  conditional on past exposure through  $t_{s-1}$  and on cigarette smoking and employment history through  $t_s$ . As such, in the absence of further assumptions, one cannot identifiably estimate the direct effect of exposure controlling for cigarette smoking. One can estimate the overall (direct and indirect) effect of exposure. (See Sec. 8D.3 for further discussion.)

Occupational mortality studies often fail to collect data on cigarette smoking history. In Lemma 8.5, we examine the circumstances under which various  $G$ -causal parameters of FR OCISTG C8.3 can be estimated from its Stage 0 reduction, MPISTG 3.4, when data on cigarette smoking is unavailable. Technically, the Stage 0 reduction of C8.3 is the  $t_1$ -modified version of MPISTG 3.4 in which both levels of  $l$  occur at  $t_1$ . Henceforth, for convenience, we also refer to the  $t_1$ -modified version of MPISTG 3.4 as MPISTG 3.4.

**LEMMA 8.5.** If cigarette smoking history is not a population risk factor for death controlling for exposure history and employment history, i.e.

$$\gamma_D(t + \Delta t | E(t), L(t), C(t)) = \gamma_D(t + \Delta t | E(t), L(t)) \quad \text{for all } C(t) \quad (8.13)$$

and cigarette history is not a predictor of future  $L$ -history controlling for exposure history, i.e.

$$\begin{aligned} p(L(t + \Delta t) | E(t), L(t), C(t), D > t + \Delta t) \\ = p[L(t + \Delta t) | E(t), L(t), D > t + \Delta t] \quad \text{for all } C(t) \end{aligned} \quad (8.14)$$

then MPISTG 3.4 (technically the  $t_1$ -modified MPISTG 3.4) is the  $B$ -complete Stage 0 PL-sufficient reduction of OPISTG C8.3. As such, the “ $G$ -null” hypothesis holds for MPISTG 3.4 if and only if it holds for OPISTG C8.3.

*Proof.* Equations (8.13) and (8.14) constitute Assumptions  $R$  of Theorem F1. Note that Eq. (8.13) is not sufficient in itself to prove the Lemma.

LEMMA 8.6. If cigarette smoking history is not a predictor of exposure history controlling for employment history, i.e.

$$\begin{aligned} p[E(t + \Delta t) | E(t), C(t + \Delta t), L(t + \Delta t), D > t + \Delta t] \\ = p[E(t + \Delta t) | E(t), L(t + \Delta t), D > t + \Delta t] \end{aligned} \quad (8.15)$$

for all  $C(t + \Delta t)$ , then for any  $G^{C8.3}$  which assigns exposure at work at  $t_s$  conditional on past work and exposure history without regard to cigarette smoking history,  $S(t | "G^{C8.3}") = S(t | "G^{3.4}")$  where " $G^{3.4}$ " is the Stage 0 counterpart of " $G^{C8.3}$ ." Thus valid tests of the " $G$ "-null hypothesis of MPISTG 3.4 constitute valid tests of the " $G$ "-null hypothesis of OPISTG C8.3.

*Proof.* Equation (8.15) satisfies the supposition of Theorem F2 for all such  $G^{C8.3}$ .

Given that MCISTG C8.3 is an FR MCISTG, we say that, given data on  $L$ -history, smoking is not a confounder for the overall effect of exposure on mortality, whenever either Eqs. (8.13) and (8.14) hold or Eq. (8.15) holds. Note that, in analyzing our arsenic data, we have considered MPISTG 3.4 to be an FR MCISTG, even though we know that cigarette smoking is a population risk factor for death. Thus, we must have been implicitly assuming that OPISTG C8.3 is an FR OCISTG, that Eq. (8.15) holds, and finally that MPISTG 3.4 is a causal melded Stage 0 reduction of FR MCISTG C8.3 so that we are able to apply Theorem F3 to conclude that MPISTG 3.4 is an FR MCISTG.

But scientific interest would likely center on the question of whether there is a causal effect of exposure controlling for cigarette smoking history. Given that we believe PISTG C8.3 is our finest FR CISTG, what further assumptions would allow us to test for a causal effect of exposure controlling for cigarette smoking when data on cigarette smoking are available? are unavailable? We answer these questions in the next subsection.

### D.3 Circumstances under which the causal effect of exposure controlling for smoking may be tested in the presence and absence of smoking data

Given MCISTG C8.3 is our finest FR MCISTG, we define circumstances under which we may obtain valid tests of the hypothesis that "exposure has no effect controlling for cigarette smoking." We begin by presenting two parallel lemmas, one for FR MCISTG 3.4 and the other for FR MCISTG C8.3.

LEMMA 8.7. If

- (1) exposure is not a causal risk factor for death controlling for  $L$ -history;
- (2) exposure is not a causal risk factor for  $L$ ; and
- (3) MCISTG 3.4 is an FR MCISTG, then
  - (a) exposure is not a predictor of future  $L$ -history;
  - (b) exposure is not a population risk factor controlling for  $L$ ; and
  - (c) the " $G$ "-null hypothesis holds for MPISTG 3.4 (since the sharp null hypothesis holds for MCISTG 3.4).

*Proof.* See proof of Lemma 8.8.

*Remark 8.1.* Given MPISTG 3.3 is an MCISTG, supposition (1) of Lemma 8.7 can be written as, for all individuals  $i$ ,

$$S(t_s, G_1^{3.3} = ["G_1^{3.4}", L(t_s)], G_2^{3.3} = ["G_2^{3.4}", L(t_s)], i) \equiv 0. \quad (8.16)$$

The  $G_k^{3.4}$  could be replaced by  $E_k(t_s)$  since the pair [ $\langle G_k^{3.4} \rangle$ ,  $L(t_s)$ ], determines a unique  $E_k(t_s)$ .

Supposition (2) can be written as  $p[L(t_s + \Delta t) | L(t_s), G^{3.4}, i]$  does not depend on  $G^{3.4}$ . (a) and (b) can be written as  $p[L(t_s + \Delta t) | L(t_s), \langle G^{3.4} \rangle]$  and  $\gamma_D(t_s + \Delta t | \langle G_1^{3.4} \rangle, L(t_s))$  do not depend on  $\langle G^{3.4} \rangle$  (as defined in definition G1 of Appendix G with  $L(t_s) \equiv \cdot i_s^B$ .)

*Remark 8.2.* If  $L$  is not a causal risk factor controlling for  $E$ , then Conclusion (c) follows from (1) and (3) alone.

*Remark 8.3.* (a) implies that (b) holds  $\Leftrightarrow$  (c) holds by Lemma 8.3.

*Remark 8.4.* The suppositions of the Lemma are all nonidentifiable. Suppositions (a) and (b) refer to individual causal effects. In contrast, the conclusions are all identifiable population relationships.

*Remark 8.5.* Supposition (2) of Lemma 8.7 implies each individual  $i$  has a unique employment history, say  $L^i(t_s)$ , when treated with any  $G^{3.4}$ . If Suppositions (2) and (3) hold, and for each individual  $i$ , Eq. 8.16 holds only for  $L^i(t_s)$ , then Conclusions (a), (b), and (c) hold [even if Eq. (8.16) is either false or undefined for various other  $L(t_s)$ ]. Note if, for some individuals, Eq. (8.16) is undefined for certain  $L(t_s)$ , PISTG 3.3 is not a CISTG.

*Remark 8.6.* Lemma 8.7 holds (as will nearly all results in this subsection) even if an individual's "covariate" history is undefined after his death. In particular, Lemma 8.7 would hold if in the restatement of Supposition (2) in Remark 8.1 we had conditioned on the event  $D > t + \Delta t$ .

**LEMMA 8.8.** Given MPISTG 8.3 is an MCISTG, if

- (1)  $G^{C8.3}$  is not a causal risk factor for death controlling for cigarette smoking, which, by definition, is,

$$S(t_s, G_1^{8.3} = [\langle G_1^{C8.3} \rangle, C(t_s)], G_2^{8.3} = [\langle G_2^{C8.3} \rangle, C(t_s)], i) \equiv 0 \quad (8.17)$$

[ $\langle G_k^{3.4} \rangle$  could be substituted for  $\langle G_k^{C8.3} \rangle$  in Eq. (8.17)—see Remark 8.1];

- (2)  $G^{C8.3}$  is not a causal risk factor for cigarette smoking, i.e.

$$p[C(t + \Delta t) | C(t), G^{C8.3}, i] \text{ does not depend on } G^{C8.3}; \quad (8.18)$$

- (3) MCISTG C8.3 is an FR MCISTG, then

- (a)  $G^{C8.3}$  is not a predictor of future  $C$ -history, that is,

$$p[C(t_s + \Delta t) | C(t_s), \langle G_1^{C8.3} \rangle] \text{ does not depend on } G_1^{C8.3} \quad (8.19)$$

(as in Definition G1 with  $C(t_s) \equiv \cdot i_s^B$ );

- (b)

$$\gamma_D(t_s + \Delta t | \langle G_1^{C8.3} \rangle, C(t_s)) \text{ does not depend on } \langle G_1^{C8.3} \rangle \quad (8.20)$$

(again see Definition G1); and

- (c)

$$\text{the } \langle G \rangle\text{-null hypothesis holds for MPISTG C8.3} \quad (8.21)$$

(since the sharp null hypothesis holds for MCISTG C8.3).

*Sketch of Proof of Lemma 8.8.* Since MCISTG C8.3 is, by supposition, an FR MCISTG, we can proceed, as in the proof of Theorem 8.2, to view the observed data as having arisen from a randomized trial in which each individual was assigned a treatment  $G^{C8.3}$  (see Theorem 7.1). It is clear that (1) and (2) imply the sharp null hypothesis for MCISTG C8.3 and that furthermore each individual has a history  $(C(t_s), D)$  that does not depend on the  $G^{C8.3}$  to which they are assigned (where  $D$  is the individual's time of death). Thus,  $(C(t_s), D) \perp\!\!\!\perp G^{C8.3}$ . Conclusions (a), (b), and (c) follow immediately.

*Remark 8.7.* The statement of Lemma 8.8 can be seen to be formally equivalent to that of Lemma 8.7 when we rewrite Lemma 8.7 by making the substitutions described in Remark 8.1 and then identify PISTG 3.3 with PISTG 8.3, and PISTG 3.4 with C8.3.

*Remark 8.8.* Both Lemmas 8.7 and 8.8 are special cases of Theorem G5.

*Remark 8.9.* Paralleling Remark 8.3, Equation (8.19) implies that Equation (8.20) holds  $\Leftrightarrow$  Equation (8.21) holds. The proof of this Result is isomorphic to the proof of Lemma 8.3, and is a special case of Theorem G4 in Appendix G.

*Remark 8.10.* Lemma 8.8 remains true if, for each individual  $i$ , Eq. (8.17) holds only for the individual's unique smoking history  $C^i(t_s)$ . (See Remark 8.5.)

We now discuss how we can use Lemma 8.8 to help draw inferences concerning the null hypothesis Eq. (8.17). Since all the suppositions of Lemma 8.8 are nonidentifiable, it is clear that to make use of the Lemma, we shall need to merge our prior beliefs concerning these suppositions with the data evidence. Here we shall give an informal Bayesian analysis in which uncertainty is expressed qualitatively rather than quantitatively. For pedagogical purposes (only), we shall suppose there are three different actions we would take depending on whether we (1) strongly believe that Eq. (8.17) holds [then accept Eq. (8.17)], (2) strongly believe that Eq. (8.17) is false [then reject Eq. (8.17)], or (3) do not have strong beliefs either that Eq. 8.17 is true or false [i.e. do not accept or reject Eq. (8.17)]. Furthermore, we suppose that prior to observing the data, we do not have strong beliefs one way or the other about the truth of Eq. (8.17). This exercise is pedagogic in the sense that it is meant to sharply raise the issues one must face in analyzing real data. It is not meant, though, as a prescription for actual analysis.

We first consider the case in which data on cigarette smoking are not available. Suppose we have strong subjective beliefs that (1) OPISTG C8.3 is an FR OCISTG, (2) Eq. (8.18) holds, and (3) Eq. (8.15) holds. Then, as a direct consequence of Lemma 8.8 and Lemma 8.6, we would reject the null hypothesis Eq. (8.17) when the " $G$ "-null test for MPISTG 3.4 rejects. On the other hand, even when both Eq. 8.15 and the " $G$ "-null hypothesis for MPISTG 3.4 hold, the " $G$ "-null hypothesis need not hold for OPISTG C8.3. Nevertheless, we would consider it unlikely that this latter situation would occur. Therefore, if the " $G$ "-null test for MPISTG 3.4 accepts, we would accept Eq. (8.17).

If our prior belief that Equation 8.18 held was weak, we would neither accept nor reject Eq. (8.17) irrespective of whether " $G$ "-null test for MPISTG 3.4 accepted or rejected, as we could not invoke Lemma 8.8.

If our beliefs in Eq. (8.18) were moderately strong [and our beliefs that C8.3 was an FR OCISTG and that Eq. (8.15) held remained strong] we might accept Eq. (8.17) if the " $G$ "-null test for MPISTG 3.4 accepted, but neither accept nor reject Eq. (8.17) if the " $G$ "-null test for MPISTG 3.4 rejects. To see why, consider the following argument. For both Eq. (8.18) and (8.17) to be false and yet for MPISTG 3.4 not to reject, a somewhat unusual balancing act must occur (for example, a direct adverse effect of exposure on mortality controlling for cigarette smoking is perfectly balanced by an indirect beneficial

effect on mortality operating through the effect of exposure on smoking behavior). Since we have moderately strong beliefs that Eq. (8.18) holds, we would be unlikely to believe both that Eq. (8.18) is false and the unusual balancing act described above occurred. Thus, we would accept Eq. (8.17). On the other hand, if the “ $G$ ”-null test for MPISTG 3.4 rejected, we would be justified in rejecting (8.17) only if we were near certain that Eq. (8.18) held (which we are not).

We now consider the case in which data on cigarette smoking is available. If we make the following rather reasonable assumption, then somewhat sharper results can be obtained. Suppose we assume that any individual affected by a treatment  $G^{C8.3}$ , when controlling for  $C(t_s)$ , is adversely affected. Formally, we shall only require the following.

*Adverse effect assumption.* If Eq. (8.17) is false and Eq. (8.18) is true, then, when  $G_2^{C8.3}$  is the treatment “always receives zero exposure”, we assume that: (1) for all  $G_1^{C8.3}$ , Eq. (8.17) is true when the  $\equiv$  is replaced by  $\leq$  and (2) for some individual  $i$ , some  $t_s$  and some  $G_i^{C8.3}$ , the inequality described in (1) is strict for  $C^i(t_s)$  (as defined in Remark 8.10).

The following two lemmas follow from the adverse effect assumption.

LEMMA 8.9. Given that the adverse effect assumption holds, if MCISTG C8.3 is an FR MCISTG and Eq. (8.18) holds, and furthermore, either (8.20) or (8.21) holds, then Eq. (8.17) holds.

*Proof.* By contradiction. Assume Eq. (8.17) is false and Eq. (8.21) holds. For the  $G_1^{C8.3}$ ,  $t_s$  described in (2) of the adverse effect assumption, and for the  $G_2^{C8.3}$  “always receive zero exposure”  $S(t_s, G_1^{C8.3}, G_2^{C8.3}) < 0$  under the suppositions of the Lemma. But this contradicts the fact that Eq. (8.21) holds. If we assume Eq. (8.20) rather than (8.21) holds, a similar proof can be used.

LEMMA 8.10. If MCISTG C8.3 is an FR MCISTG, Eq. (8.20) or (8.21) holds, Eq. (8.19) is false, and the adverse effect assumption holds, then Eq. (8.18) is false.

*Proof.* By contradiction. Assume Eq. (8.18) is true. Then Eq. (8.17) must hold by the previous Lemma. But if Eq. (8.17) holds, then, by Lemma 8.8, Eq. (8.19) is true.

Suppose now that we have strong beliefs that (1) MPISTG C8.3 is our finest FR MCISTG, (2) Eq. (8.18) holds, and (3) the adverse effect assumptions holds. We will call this set of beliefs—Belief Structure 1. Then if Eq. (8.20) or (8.21) holds, we would accept Eq. (8.17) by Lemma 8.9. But, if Eq. (8.19) is false and Eq. (8.20) or (8.21) holds, then by Lemma 8.10 our beliefs must have been mistaken. [Here we are supposing that we have sufficient data to determine, without error, whether Eq. (8.19) and Eq. (8.20) empirically hold.] Therefore, we would neither accept nor reject Eq. (8.17). If Eq. (8.20) and (8.21) were both false, then we would reject Eq. (8.17) irrespective of whether Eq. (8.19) were true or false. [Note, without further assumptions, if Eq. (8.18) is true and Eq. (8.17) is false, Eq. (8.19) will, in general, be expected to be false.]

Now suppose our belief that Eq. (8.18) is true is only moderately strong and our other beliefs remain as in Belief Structure 1. We call this set of beliefs Belief Structure 2. Then, if Eq. (8.19) and Eq. (8.20) [and thus Eq. (8.21) by Remark 8.9] hold, we might accept Eq. (8.17), since, again, it would take an unusual balancing act for Eqs. (8.18) and (8.17) to be false and yet for Eqs. (8.19) and (8.20) to hold. We now consider the three other possible empirical outcomes.

*Case 1.* If Eq. (8.20) or (8.21) holds and Eq. (8.19) is false, in general, we neither accept nor reject Eq. (8.17). This follows because Eq. (8.18) is false by Lemma 8.10, and thus Lemma 8.8 cannot be invoked.

*Case 2.* If Eqs. (8.20), (8.21), and (8.19) were all false, we would, in general, neither accept nor reject Eq. (8.17). This reflects the fact that we would not be able to come to a definitive conclusion as to whether it was Eq. (8.17) or (8.18) (or both) that was false.

*Remark 8.11.* If one but not both of Eqs. (8.17) and (8.18) is true, then, without further assumptions, Eqs. (8.20), (8.21), and (8.19) will, in general, all be false.

*Case 3.* If Eqs. (8.20) and (8.21) were false and Eq. (8.19) true, we shall assume we would usually reject Eq. (8.17), although it would not be unreasonable to neither accept nor reject Eq. (8.17). To see why, we first show that it is possible for Eq. (8.17) to be true and Eq. (8.18) to be false. As a rather extreme example, suppose a population was composed of equal numbers of two homogeneous groups of individuals *A* and *B*. Individuals in group *A* are nonsmokers at  $t_1$ , become smokers at  $t_2$  when treated with high exposure at  $t_1$ , and remain nonsmokers at  $t_2$  when unexposed at  $t_1$ . Group *A* individuals die at  $t_3$  if and only if they are smokers at  $t_2$ . On the other hand, individuals in group *B* are smokers at  $t_1$ , remain smokers at  $t_2$  if treated with zero exposure at  $t_1$ , and become nonsmokers at  $t_2$  if treated with high exposure at  $t_1$ . Furthermore, no group *B* individuals die at  $t_3$  irrespective of smoking status at  $t_2$ . Then, Eqs. (8.20) and (8.21) will be false and (8.19) true, although Eq. (8.17) is true and Eq. (8.18) is false. This example required that the exposure influence the smoking status of type *A* and *B* individuals in opposite directions. For the outcome of smoking status (in contrast with mortality), it does not seem reasonable to exclude such a possibility *a priori*. Even so, the "balancing act" described in this example seems quite unlikely to actually occur. Therefore, we shall reject Eq. (8.17) when Eqs. (8.20) and (8.21) are false and Eq. (8.19) true. In light of Remark 8.11, we should not expect Case 3 to often occur.

In the absence of moderately strong beliefs that Eq. (8.18) holds, we would neither accept nor reject Eq. (8.17) even when Eqs. (8.19) and (8.20) held.

Henceforth, we shall assume that we hold beliefs compatible with Belief Structure 2. It follows that when data on cigarette smoking history are available, we shall need to test whether Eqs. (8.19) and (8.21) [equivalently Eqs. (8.19) and (8.20)] simultaneously hold. It may appear that any test of Eq. (8.19) requires that we use statistical models to estimate the probabilities  $S(\cdot|s, j_s)$  and  $\gamma(\cdot|s)$  of MPISTG C8.3 in order to estimate  $p(C(t + \Delta t) | C(t), "G^{C8.3}")$  for several times  $t$ , smoking histories  $C(t)$ , and generalized treatments " $G^{C8.3}$ ". Such an approach would be quite sensitive to model misspecification. Fortunately, we can construct a nonparametric test of the joint null hypothesis that Eqs. (8.19) and (8.20) hold based on the following lemma (which is a special case of Theorem G4).

**LEMMA 8.11.** Equation (8.19) and Eq. (8.20) hold if and only if for  $1 \leq k \leq s$ ,

$$p(C(t_s + \Delta t) | C(t_s), L(t_k), E(t_{k-1}), e(t_k), D > t_s + \Delta t) \text{ does not depend on } e(t_k) \quad (8.22)$$

and

$$\gamma_D(t_s + \Delta t) | C(t_s), L(t_k), E(t_{k-1}), e(t_k)) \text{ does not depend on } e(t_k), \quad (8.23)$$

where  $e(t_k)$  is the exposure concentration at  $t_k$ . In Sec. 9 we explicitly construct a nonparametric test of the null hypothesis that Eq. (8.19) and (8.20) hold based on Eqs. (8.22) and (8.23).

Suppose now our beliefs conform with Belief Structure 2. Then, if our nonparametric test of Eqs. (8.19) and (8.20) accepts, we accept Eq. (8.17). If the test rejects, we would

neither accept nor reject Eq. (8.17), unless Eq. (8.19) held and Eqs. (8.20) and (8.21) both did not (Case 3) in which case we would reject Eq. (8.17). But, unfortunately, if Eqs. (8.19) and (8.20) do not both hold, there exists no valid nonparametric test of either Eq. (8.19) or (8.20). In particular, Eq. (8.19) alone does not imply Eq. (8.22) and Eq. (8.20) alone does not imply Eq. (8.23). Thus, in theory to test whether Eq. (8.19) holds, we must resort to modelling the conditional probabilities of MCISTG C8.3.

We have seen that acceptance of the joint null hypothesis that Eq. (8.19) and (8.20) hold would lead us to accept Eq. (8.17) only if we have moderately strong beliefs that Eq. (8.18) holds. But our beliefs about Eq. (8.18) may be hard to assess because the meaning of Eq. (8.18) may be rather nonintuitive. Therefore, we now give two sufficient conditions for Eq. (8.18) to hold.

LEMMA 8.12. If

$$p[C(t + \Delta t) | C(t), L(t), E(t), i] = p[C(t + \Delta t) | C(t), L(t), i] \quad (8.24)$$

holds and either

$$p[L(t + \Delta t) | L(t), E(t), C(t), i] = p[L(t + \Delta t) | L(t), C(t), i] \quad (8.25)$$

or

$$p[C(t + \Delta t) | E(t), C(t), L(t), i] = p[C(t + \Delta t) | E(t), C(t), i] \quad (8.26)$$

then Eq. (8.18) holds. In words, if  $E$  is not a causal risk factor for  $L$  and  $C$  jointly [Eqs. (8.24) and (8.25)], or neither  $E$  nor  $L$  is a causal risk factor for  $C$  [Eqs. (8.24) and (8.26)], then Eq. (8.18) holds.

*Proof.* Left to the reader.

It is our conjecture that the meanings of these two sufficient conditions are more intuitive than that of Eq. (8.18) in the sense that any investigator having moderately strong belief in Eq. (8.18) would have this belief only as a consequence of holding a moderately strong belief in Eq. (8.24) and in either Eq. (8.25) or (8.26).

Our conjecture, if correct, has implications for our inference concerning the null hypothesis Eq. (8.17). To deduce these implications, we shall need the following two lemmas.

LEMMA 8.13. If Eqs. (8.24), (8.25), and (8.17) hold, and MCISTG C8.3 is an FR MCISTG then

$$\gamma_D(t + \Delta t | E(t), L(t), C(t)) = \gamma_D(t + \Delta t | L(t), C(t)) \quad (8.27)$$

$$\begin{aligned} p[C(t + \Delta t), L(t + \Delta t) | C(t), L(t), E(t), D > t + \Delta t] \\ = p[C(t + \Delta t), L(t + \Delta t) | C(t), L(t), D > t + \Delta t] \end{aligned} \quad (8.28)$$

*Proof.* The lemma is a special case of Theorem G5 on letting  $\cdot i_s^B \equiv (L(t_s), C(t_s))$ , CISTG A be CISTG C8.3, and  $\tau^A \equiv G^{C8.3}$ . Supposition (1) of Theorem G5 is then equivalent to Eqs. (8.24) and (8.25). Supposition (2) follows from Lemmas 8.12 and 8.8. Finally, Eq. (G1) becomes Eq. (8.28) and Eq. (G5) is Eq. (8.27).

LEMMA 8.14. If Eq. (8.28) holds then we have Eq. (8.27) holds if and only if Eqs.

(8.19) and (8.20) hold [and thus Eq. (8.21) holds]. Furthermore, if Eqs. (8.27), (8.28), and (8.15) holds, then, when ignoring data on  $C$ , exposure is neither a predictor of future  $L$ -history nor a population risk factor for death controlling for  $L$ .

*Proof.* Let  $F(t_s) \equiv (L(t_s), C(t_s))$  replace  $L(t_s)$  in Lemma 8.3. Let MPISTG C8.3 replace MPISTG 3.4 in that Lemma. Then Lemma 8.3 states that if Eq. (8.28) holds, then Eq. (8.27) holds if and only if Eq. (8.21) holds. It remains to show that Eq. (8.27) plus Eq. (8.28) imply Eq. (8.19). This is left as an exercise. The second part of the Lemma is proved in Ref. [7].

It follows that if we held Belief Structure 2 on the basis of a moderately strong belief in Eqs. (8.24) and (8.25), then, if Eq. (8.27) and (8.28) hold, we would accept that Eq. (8.17) holds. Furthermore, by reasoning similar to that we have used above, if Eq. (8.28) is false, we would neither accept nor reject Eq. (8.17), and if Eq. (8.28) held and Eq. (8.27) were false, we would reject Eq. (8.17). In Sec. 9, we briefly discuss empirical tests of Eqs. (8.27) and (8.28). The second part of Lemma 8.14 has implications for our inferences about Eq. (8.17) in the absence of data on cigarette smoking. Previously, we suggested that when our beliefs in Eq. (8.18) were moderately strong and the “ $G$ ”-null test for MPISTG 3.4 accepts, we accept Eq. (8.17). It is now clear that if the basis of our *a priori* belief in Eq. (8.18) is our beliefs in Eq. (8.24) and (8.25), then even when the “ $G$ ”-null test for MPISTG 3.4 accepted, if exposure was an empirical predictor of future  $L$ -history, we should neither accept nor reject Eq. (8.17).

*Remark 8.12.* In contrast to Lemma 8.13, the knowledge that MPISTG C8.3 is an FR MCISTG and that Eqs. (8.24), (8.26), and (8.17) hold implies no further empirical relationships beyond Eqs. (8.19), (8.20), and (8.21).

We next suppose that one was willing to assume *a priori* that MPISTG 8.3 was FR MCISTG. Then Eq. (8.17) implies

$$S(t, "G_1^{8.3}") = ["G_1^{C8.3}", C(t_s)], "G_2^{8.3}" = ["G_2^{C8.3}", C(t_s)] \equiv 0 \quad (8.29)$$

for all  $t$ ,  $G_1^{8.3}$ ,  $G_2^{8.3}$  of the above form. Obviously if Eq. (8.29) is false, Eq. (8.17) is false even when Eqs. (8.19) and (8.20) are true. (Thus, in this circumstance, it follows from Lemma 8.9 that, if the adverse effect assumption holds, Eq. (8.18) would be false as well.) Unfortunately, in general, no nonparametric test of Eq. (8.29) exists. Nevertheless, we would like to determine particular (empirical) conditions under which a nonparametric test of Eq. (8.29) exists. For instance, we would like to know the (empirical) circumstances under which the “ $G$ ”-null test for MPISTG C8.3 is a valid test of Eq. (8.29). These circumstances are described in the following lemmas.

**LEMMA 8.15.** If Eq. (8.29) holds and if for all  $t_k$ ,  $2 \leq t_k \leq t_{s+1}$

$$\begin{aligned} p[C(t_{s+1}) | C(t_s), D > t_{s-1}, L(t_{k-1}), E(t_{k-1}), l(t_k)] \\ = p[C(t_{s+1}) | C(t_s), D > t_{s-1}, L(t_{k-1}), E(t_{k-1}), \bar{l}(t_k)] \end{aligned} \quad (8.30)$$

then Eq. (8.23) holds.

LEMMA 8.16. Equations (8.19) and (8.30) hold if and only if for  $s \geq 1$

$$p[C(t_{s+1}) | C(t_s), E(t_s), D > t_{s+1}, L(t_{s+1})] = p[C(t_{s+1}) | C(t_s), D > t_{s+1}]. \quad (8.31)$$

Furthermore, if Eq. (8.19) and (8.30) hold, then we have Eq. (8.29) holds  $\Leftrightarrow$  Eq. (8.21) holds  $\Leftrightarrow$  Eqs. (8.20), (8.22), and (8.23) hold.

*Corollary 8.16.* If Eq. (8.31) holds and Eq. (8.15) holds, then Eq. (8.29) implies that the “ $G$ ”-null hypothesis holds for MPISTG 3.4.

If we strengthen Eq. (8.31) slightly, we have the following.

LEMMA 8.17. If Eq. (8.31) holds for all  $s \geq 0$ , then

$$\gamma_D(t_s + \Delta t | "G_1^{C8.3}", C(t_s)) = \gamma_D(t_s + \Delta t | "G_1^{8.3}" = ("G_1^{C8.3}", C(t_s)))$$

and

$$S(t_{s+1} | "G_1^{C8.3}") = \sum_i S[t_{s+1} | "G_{1i}^{8.3}"] = ("G_{1i}^{C8.3}", C_i(t_s)) p^\Delta[C_i(t_s)], \quad (8.32)$$

where  $i$  indexes the  $2^s$  possible paths  $C_i(t_s)$ ;  $C_i(t_s)$  is only required to have initial segment  $C_i(t_s)$ , and  $p^\Delta[C(t_s)] = \prod_{k=0}^{s-1} p[C(t_{k+1}) | C(t_k), D > t_{k+1}]$  with  $C(t_k)$  the initial part of  $C(t_s)$ .

*Proofs.* Lemmas 8.15–8.17 are special cases of Theorem G1–G3.

As promised, Lemma 8.16 gives a sufficient (empirical) condition [i.e. that Eq. (8.31) holds] for the equivalence of Eqs. (8.21) and (8.29). In Sec. 9, we show that Eq. (8.31) can itself be tested nonparametrically. Lemma 8.15 shows that if Eq. (8.30) holds [even though Eq. (8.19) does not], then Eq. (8.29) implies Eq. (8.23). In Sec. 9 we consider whether Lemma 8.15 can serve as the basis of a valid nonparametric test of Eq. (8.29).

*Remark 8.13.* Lemmas 8.15, 8.16, and Corollary 8.16 were not initially obvious to us. We guessed them by imagining various double blind randomized trials that could have led to R SCISTG 8.3 as their Stage 0 PL-sufficient reduction. The informal power of this kind of reasoning to guess theorems about PISTGs is great enough that we describe the process by which we guessed the above lemmas. If an investigator had conducted a double blind ordinary designed randomized trial in which treatments  $G^{3.4}$  and  $C(t_s)$  were assigned independently at  $t_1$ , and if no individual left their treatment protocol  $C(t_s)$  or  $G^{3.4}$  [e.g. individuals at work always received their assigned exposure, irrespective of their  $C(t_s)$  history, and treatment  $G^{3.4}$  had no effect on mortality controlling for  $C(t_s)$ ], then the crude survival- $G^{3.4}$  association would be null. This would have the “empirical consequence” that the “ $G$ ”-null hypothesis would hold for MPISTG 3.4 even in the absence of data on treatment protocol. As such, we were interested in determining whether the observed study data were consistent with being the Stage 0 PL-sufficient reduction of such an ordinary designed randomized trial. To do so, we built up such a trial in three steps corresponding to Lemmas 8.15, 8.16, and Corollary 8.16, respectively. First, we only supposed that in the ordinary designed double blind randomized trial, protocols  $C(t_s)$  and  $G^{C8.3}$  were assigned at random [i.e. individuals were assigned a protocol such that, if at

work at  $t_s$ , their assigned exposure might depend on their value of  $C(t_s)$ . no individual left protocol, but the probability of assignment to a particular  $G^{C8.3}$  might depend on the assigned  $C(t_s)$ . In such a trial, Eqs. (8.29), (8.30) and (8.23) follow from (8.17). If Eq. (8.17) were false, neither (8.30) nor (8.23) need be true. We leave the proof as an exercise. This approach allows us to guess Lemma 8.15. It is not a proof of the Lemma since we would still need to show that Eqs. (8.30) and (8.29) imply (8.23). In the second stage we asked what the further observable consequences would be if  $C(t_s)$  and  $G^{C8.3}$  had been assigned independently (and no one left protocol). It follows from the nonidentifiable temporal assumption that Eqs. (8.30), (8.19), and (8.31) must all hold even when Eq. (8.17) is false. In addition, Eq. (8.21) must hold since if the stratum-specific  $G^{C8.3}$  effects are null [Eq. (8.17) and (8.29) hold] and there is no stratum- $G^{C8.3}$  association [with strata defined by  $C(t_s)$ ], then empirical crude effect [as represented by Eq. (8.21)] must hold. In the last stage, we asked what would be the additional empirical consequences of having all the  $G^{C8.3}$  assigned in the above trial be of the form  $G^{3.4}$  (and no individual left protocol). It is obvious that Eq. (8.15) and the conclusion to Corollary 8.16 would then hold.

We now give a further (empirical) condition under which Eqs. (8.29) and (8.21) are equivalent.

**LEMMA 8.18.** If Eq. (8.28) holds, then we have Eq. (8.29) holds  $\Leftrightarrow$  Eq. (8.21) holds  $\Leftrightarrow$  Eq. (8.27) holds.

*Proof.* From Lemma 8.14 we know that if Eq. (8.21) holds, then Eq. (8.27) holds, which immediately implies Eq. (8.29) holds. Conversely, one can show that if Eq. (8.29) holds, then Eq. (8.27) holds, which implies that Eq. (8.21) holds.

*Remark 8.14.* Suppose MCISTG C8.3 were an FR MCISTG. Consider an investigator who held moderately strong *a priori* beliefs that Eqs. (8.24) and (8.25) held and who subsequently accepted Eq. (8.17) upon empirically confirming that Eqs. (8.27) and (8.28) held. Lemma 8.18 shows that this investigator's inferences concerning Eq. (8.17) could not be changed by further *a priori* knowledge that MCISTG 8.3 was an FR MCISTG.

Finally, we have the following.

**LEMMA 8.19.** If the “ $G$ ”-null hypothesis holds for MPISTG 8.3 then Eq. (8.29) holds and Eq. (8.21) holds. Of course, in general, cigarette smoking will be a causal risk factor for death controlling for C8.3 and so

$$S(t; "G_1^{8.3}" = ["G^{C8.3}", C_1(t_s)]; "G_2^{8.3}" = ["G^{C8.3}", C_2(t_s)]) \equiv 0 \quad (8.33)$$

will be false if MCISTG 8.3 is an FR MCISTG.

Now, it is possible that one's interest is not in testing whether there is any effect of exposure controlling only for cigarette smoking history, [i.e. whether Eq. (8.17) holds], but rather in whether there is an effect of exposure controlling for cigarette and employment history, i.e. whether for MCISTG F8.3

$$\gamma_D(t + \Delta t | E(t), L(t), C(t), i) = \gamma_D(t + \Delta t | L(t), C(t), i). \quad (8.34)$$

If so, in order to utilize Lemma 8.8, we need to describe conditions under which Eq. (8.34) implies Eq. (8.17). We have the following.

LEMMA 8.20. If Eq. (8.34) holds and either Eq. (8.25) holds or

$$\gamma_D(t + \Delta t | E(t), L(t), C(t), i) = \gamma_D(t + \Delta t | E(t), C(t), i), \quad (8.35)$$

then Eq. (8.17) holds.

*Remark 8.15.* Given that MCISTG C8.3 is an FR MCISTG, knowledge that Eqs. (8.34) and (8.25) hold or that (8.34) and (8.35) hold does not imply any empirical restrictions. Knowledge that Eqs. (8.34), (8.35), (8.24), and (8.26) hold imply no empirical restrictions beyond those implied by Eqs. (8.18) and (8.17) in Lemma 8.8. Finally, knowledge that Eqs. (8.34), (8.25), and (8.24) hold implies no empirical restrictions beyond those implied by Eqs. (8.24), (8.25), and (8.17) in Lemmas 8.13, 8.14, and 8.18. These last remarks are unchanged if it is known that MCISTG 8.3 is an FR MCISTG.

We now summarize the conditions under which one may test for a causal effect of exposure controlling for employment and cigarette smoking history [i.e. Eq. (8.34)] both in the presence and absence of data on cigarette smoking. First we consider the case in which only exposure is received at random (i.e. MCISTG C8.3 is our finest FR MCISTG). From Lemmas 8.8, 8.12, 8.13, 8.14, 8.20, and 8.6 it follows that

LEMMA 8.21. If MCISTG C8.3 is an FR MCISTG, Eq. (8.24) holds, and either (a) Eqs. (8.26) and (8.35) hold, or (b) Eq. (8.25) holds, then Eq. (8.34) implies (1) the “ $G$ ”-null hypothesis for MPISTG C8.3 holds, and, if Eq. (8.15) holds as well, the “ $G$ ”-null hypothesis holds for MPISTG 3.4; and (2) if (b) holds, Eqs. (8.27) and (8.28) hold and, if Eq. (8.15) holds as well, exposure is not a predictor of future  $L$ -history and exposure is not a population risk factor for death controlling for  $L$ .

Next we consider the case in which both exposure and cigarette smoking are given at random (i.e. MCISTG 8.3 is our finest FR MCISTG). Then, if either Eq. (8.25) or (8.35) holds, Eq. (8.34) implies Eq. (8.29). On the other hand, if data on  $C$  is unavailable, the additional knowledge that MCISTG 8.3 is an FR MCISTG does not appear to be useful.

Next we consider the case in which  $E$  and  $L$  are given at random, i.e. CF8.3 is our finest FR MCISTG (see legend to Fig. 8.3).

LEMMA 8.22. If Eqs. (8.24) and (8.34) hold then for all individuals  $i$ :

$$S[t, G_1^{\text{CF8.3}} = (E_1(t_S), L(t_S)), G_2^{\text{CF8.3}} = (E_2(t_S), L(t_S)), i] \equiv 0. \quad (8.36)$$

*Proof.* This is a special case of Theorem G5 with MCISTG F8.3 as  $Q$ , MCISTG CF8.3 as  $A$  and for each  $L(t_S)$ ,  $\tau^{\text{CF8.3}}(L(t_S))$  is determined by the condition  $G_1^{\text{CF8.3}} \in \tau^{\text{CF8.3}}(L(t_S)) \Leftrightarrow G_1^{\text{CF8.3}} = [E_1(t_S), L(t_S)]$  for some  $E_1(t_S)$ .

Thus, if MCISTG CF8.3 is an FR MCISTG and we have data on  $C$  we only require the additional information that Eq. (8.24) holds in order to validly test Eq. (8.34). We now give further consideration to inference in the absence of data on cigarette smoking.

If (as is natural) we assume that MCISTG 3.3 is a causal melded reduction of MCISTG CF8.3, we have

*Corollary 8.22.* If Eqs. (8.24) and (8.34) hold, then Eq. (8.36) holds with CF8.3 replaced by 3.3 [equivalently, Eq. (8.16) holds]. In words, this corollary says that if exposure is

not a causal risk factor for  $C$  controlling for  $L$ , and if exposure is not a causal risk factor for death controlling for  $L$  and  $C$ , exposure is not a causal risk factor for death controlling only for  $L$ .

*Remark 8.16.* Now from Lemma 8.7 and Remark 8.2 we know that if Eq. (8.16) holds, the  $G$ -null hypothesis will hold for MCISTG 3.4 if  $E$  is not a causal risk factor for  $L$  or if  $L$  is not a causal risk factor for death controlling for  $E$ . But,

LEMMA 8.23. If Eqs. (8.24) and (8.25) hold, and MCISTG 3.4 is a causal melded reduction of MCISTG C8.3, then, in MCISTG 3.4,  $E$  is not a causal risk factor for  $L$ .

*Proof.* Straightforward.

LEMMA 8.24. If Eqs. (8.26) and (8.35) hold and MCISTG 3.3 is a causal melded reduction of MCISTG CF8.3, then  $L$  is not a causal risk factor for death controlling for  $E$ .

*Proof.* Isomorphic to Corollary 8.22. Note that Eq. (8.26) implies  $L$  is not a causal risk factor for  $C$  controlling for  $E$  and Eq. (8.35) implies  $L$  is not a causal risk factor for death controlling for  $C$  and  $E$ .

Note, the last two lemmas really just reprove part of Lemma 8.21. This follows by noting that if Eq. (8.15) holds, if MCISTG C8.3 is an FR MCISTG, and if MCISTG 3.4 is the causal melded reduction of MCISTG C8.3, then MCISTG 3.4 is an FR MCISTG by Theorem F3. Now apply the empirical conclusions of Lemma 8.7 and Remark 8.2. The results beginning with Corollary 8.22 did not require that MCISTG CF8.3 was an FR MCISTG. When MCISTG CF8.3 is an FR MCISTG we have

LEMMA 8.25. If Eq. (8.24) holds, MCISTG CF8.3 is an FR MCISTG, Eq. (8.15) holds, MCISTG 3.3 is the causal melded reduction of FR MCISTG CF8.3 and

$$\begin{aligned} p[L(t + \Delta t) | E(t), L(t), C(t + \Delta t), D > t + \Delta t] \\ = p[L(t + \Delta t) | E(t), L(t), D > t + \Delta t] \end{aligned} \quad (8.37)$$

then Eq. (8.34) implies

$$\gamma_D(t + \Delta t | E(t), L(t)) = \gamma_D(t + \Delta t | L(t)). \quad (8.38)$$

*Proof.* By Corollary 8.22, Eqs. (8.24) and (8.34) together imply Eq. (8.16) holds. Thus, Eq. (8.38) will follow if MCISTG 3.3 is an FR MCISTG. This in turn will follow from Theorem F3 if CF8.3 is an FR MCISTG, and Eqs. (8.15) and (8.37) hold.

*Remark 8.17.* Since, in this subsection, we have, where possible, tried to clarify the basis of any empirical conditions in terms of the underlying individual causal effects, we now attempt to do so for Eq. (8.37). To do so, we first consider the conditions under which  $C$  is not an empirical predictor of future  $L$ -history controlling for  $E$ -history [i.e. the conditions under which Eq. (8.37) would hold when “modified” such that  $C(t + \Delta t)$  is replaced by  $C(t)$ ]. If  $C$  were a causal risk factor for  $L$ -history controlling for  $E$ -history (as would be the case if cigarette smoking caused an (unrecorded) illness which, in turn, caused individuals to leave work), then the “modified” Eq. (8.37) would, in general, not hold. Similarly, if  $C$  and  $E$  were not received at random with respect to risk factors for

future  $L$ -history conditional on past  $E$ -,  $C$ -, and  $L$ -history, the "modified" Eq. (8.37) would in general be false. This would be the case if, for example, socially maladjusted individuals both tend to smoke at a greater rate than other individuals and to leave work early (because of inability to hold a stable job). Finally, even if  $C$  and  $E$  are received at random with respect to risk factors for future mortality and  $L$ -history, and  $C$  is not a causal risk factor for  $L$  controlling for  $E$ , nonetheless, if  $C$  is an independent causal risk factor for death, the "modified" Eq. (8.37) still may not hold (compare Lemmas 8.7 and 8.8). Rather, an additional condition sufficient for the "modified" Eq. (8.37) to hold, would be that deaths occur independently of future  $L$ -history, i.e. the second supposition of Theorem 8.2 would hold when  $C(t)$  is added to the conditioning event on each side of the equation. For Eq. (8.37) to hold given that the "modified" version of Eq. (8.37) holds, we would first require that, at  $t_1$ ,  $L$  and  $C$  were unassociated. This would be trivially true if  $t_1$  is the date of hire since all individuals would be at work. Secondly, we would require that the time intervals  $\Delta t$  on our causal tree graph be very short in order to approximate the infinitesimal intervals needed to insure that a change in a cause is not followed by a change in its effect in the same interval. (If  $\Delta t$  were not short then, even if  $C$  did not predict future  $L$ , there could exist a cigarette smoking-employment status association in a given node if it was the case that  $L$  is an empirical predictor of future  $C$ -history when the data are recorded at intervals of  $\Delta t/2$ . If we take  $\Delta t$  sufficiently small, no individual's  $L$ - and  $C$ -history would simultaneously change in any interval  $\Delta t$  so that there would be no possibility of association with a given node after  $t_1$ . A formal treatment would require the restatement of our problem in terms of multivariate counting processes in continuous time.)

We must now determine whether our actual beliefs concerning our arsenic-exposed cohort are such that we are able to empirically test for an effect of exposure controlling for employment and cigarette smoking history. (We have no data on smoking history.) Presumably if disabled individuals tend to leave the workforce MCISTG CF8.3 will not be an FR MCISTG and Eq. (8.37) will be false (see above discussion). Therefore, any valid test of Eq. (8.34) must be based on the results of Lemma 8.21. We now examine the plausibility of the various suppositions of Lemma 8.21.

To begin, it seems reasonable to assume that MCISTG C8.3 is an FR MCISTG provided we are not studying an industry in which job assignment is related to unmeasured health status (e.g. the mining industry). Furthermore, it seems reasonable to assume that Eq. (8.15) holds, i.e. job assignment conditional on past work and exposure history is unrelated to cigarette smoking history. [In the asbestos industry, Eq. (8.15) might not hold since smokers are now often preferentially transferred out of jobs with high exposure to asbestos.]

We next consider the plausibility of Eq. (8.18) [i.e. by Lemma 8.12, the plausibility of Eq. (8.24) and either Eq. (8.25) or (8.26)]. If we accept that an individual is more likely to give up cigarette smoking when ill (e.g. when suffering from chronic heart or lung disease) than when healthy, Eq. (8.18) can, in general, hold only when Eq. (8.17) is true. This follows because if Eq. (8.17) were false and the deaths caused by exposure (i.e.  $G^{C8.3}$ ) are preceded by a period of illness, then the treatments  $G^{C8.3}$  would influence cigarette smoking history through the intermediate variable of ill health. Fortunately, for testing a null hypothesis that implies Eq. (8.17), it is immaterial that Eq. (8.18) would be false under the alternative.

But, as argued above, we require not only that Eq. (8.18) be plausible, but that we have moderately strong beliefs that it be true [either as a consequence of our beliefs in Eqs. (8.24) and (8.26), or in Eqs. (8.24) and (8.25)]. It seems reasonable that Eq. (8.24) will hold under the null hypothesis Eq. (8.34) since, if exposure does not cause a disabling illness leading to death, exposure may well not affect cigarette smoking history. An ex-

ception would be when the exposure was a respiratory irritant (for example, an asthmagen). Exposure would then influence an individual's cigarette smoking behavior through the intermediate variable, asthma (even though it did not influence mortality.) Fortunately, arsenic is not a known respiratory irritant.

In contrast, there is a good chance that, even under the null hypothesis, Eq. (8.25) would be false. For example, high exposure jobs may pay more than low exposure jobs. Therefore, individuals may tend to remain employed longer in high exposure jobs for purely economic reasons. Thus, it would not seem warranted to put much faith in Eq. (8.25). [By Lemmas 8.14 and 8.20 it follows that, given we have accepted Eq. (8.15) and (8.24) as true, Eq. (8.34) and (8.25) together imply  $E$  is not a predictor of future  $L$ -history. But in Table 5 of Appendix D, we observe that  $E$  is a predictor of future  $L$ -history. Thus, if our prior beliefs in Eq. (8.18) were based on beliefs concerning Eq. (8.25) then, even if the “ $G$ ”-null test for MPISTG 3.4 accepted, we would not accept the null hypothesis Eq. (8.34). See the discussion following Lemma 8.14.]

From the above discussion it follows that we would consider the “ $G$ ”-null test for MPISTG 3.4 to be a valid test of Eq. (8.34) only if we had moderately strong prior beliefs that Eq. (8.26) and (8.35) held. (See Lemma 8.21) Now Eq. (8.35) will be false, under the null hypothesis, when  $L$ -history is a causal risk factor controlling for  $C$ -history (mediated, e.g. through the adverse effects of the loss of health insurance and the increased stress associated with unemployment). We believe that various investigators would hold quite different degrees of belief as to whether Eq. (8.35) was false to a biologically significant degree. For the time being, we proceed as if we were willing to give high credibility to Eq. (8.35).

Equation (8.26) would be false (i.e.  $L$  would be an independent causal risk factor for  $C$ -history) if individuals respond to the increased stress of unemployment by increasing their cigarette consumption. Again, we believe investigators would disagree as to the likelihood that Eq. (8.26) would be false to a biologically significant degree. Note also that if individuals, when ill, tend to give up smoking, Eq. (8.26) would be false if  $L$  were a causal risk factor for death controlling for  $C$ -history [i.e. if Eq. (8.35) were false].

One might hope that if data on cigarette smoking were available, one could test Eq. (8.26) directly by testing, for example, whether  $L$ -history was an independent population predictor of future  $C$ -history. But, this would not be a valid test even under the null since, when the healthy worker survivor effect is operating, even were Eq. (8.26) true,  $L$ -history may be a population predictor of future  $C$ -history (because, for example, illness status predicts future cigarette smoking history controlling for past cigarette smoking history, and  $L$ -history is correlated with illness status because sick individuals leave work).

## 9. CAUSAL INFERENCE FROM CASE-CONTROL DATA

In cohort mortality studies, nested case-control sampling designs are often used to save computing and/or data acquisition costs. In the usual case-control design[11, 12] one samples controls from individuals at risk at the death age of the case (possibly matched to the cases on levels of various covariates). As we did in Sec. 5, one can allow other types of failure to also represent cases (e.g. “leaving work” and “returning to work”) and sample controls from individuals at risk for these failures as well. As such, one may simultaneously have data from nested case-control studies of different failure types. In this section, we consider the estimation of and tests for the  $G$ -causal parameters of an R MCISTG using case-control data. For the present, we define an analysis in which information on the control sampling fractions is either unknown or not used to be a case-control analysis. We define a *cohort analysis* of case-control data to be one in which

information on the control sampling fraction is used to estimate absolute hazards as well. We choose to slightly modify these definitions later.

*A. The null hypothesis of no exposure effect controlling for cigarette smoking may be untestable from case-control data when the healthy worker selection effect is operative*

In the following example, we show that even when Fig. 8.3 is an FR MCISTG, we cannot test the null hypothesis that exposure has no effect on mortality controlling for smoking history on the basis of a case-control analysis.

*Example 1.* We suppose that (1) the observed data are at their expected values (i.e. we can ignore sampling variability and the associated problems of small sample size); (2) data from a single case-control study (in which deaths constitute the cases) has been obtained; and (3) no data is available concerning the sampling fraction.

The STG in Fig. 9.1b represents data from a case-control study in which all deaths and a random sample of noncases were sampled. Individuals who are small  $c$  ( $\bar{c}$ ) are (not) current smokers at  $t_2$ . Subjects who are  $\bar{l}(l)$  are (not) at work at  $t_2$ . All deaths occurred at  $t_3$ . If the control sampling fraction were .1, Fig. 9.1a would represent the full cohort data. If the sampling fraction were 1, Fig. 9.1b would represent the full cohort (both 9.1A and 9.1B can be viewed as special cases of Fig. 8.3). Obviously, without knowledge of the sampling fraction, we cannot empirically determine which graph represents the full cohort data. Whichever it is, we will assume that it is an FR MCISTG. If Fig. 9.1a were the FR MCISTG of the cohort, then exposure has no effect on controlling for smoking history [i.e. using the G-computation algorithm,  $p(D > t_3 | G_{H,c}^{9,1a}) = \frac{3000}{4000} \frac{8000}{16000} + \frac{4000}{8000} \frac{8000}{16000} = \frac{100000}{160000}$  and  $p(D > t_3 | G_{0,c}^{9,1a}) = \frac{10000}{16000}$ .  $G_{H,c}$  is the hypothetical study in which all individuals received high exposure at  $t_1$  and do not smoke at  $t_2$ .  $G_{0,c}$  is similar, except all individuals received zero exposure at  $t_1$ ]. If Figure 9.1b is the FR MCISTG, then exposure has an adverse effect on mortality, controlling for smoking history, since  $p(D > t_3 | G_{H,c}^{9,1b}) = \frac{300}{1300} \frac{1100}{8800} + \frac{400}{4400} \frac{1100}{8800} = .08$  and  $p(D > t_3 | G_{0,c}^{9,1b}) = \frac{1000}{8000} = 0.14$ .

Allowing for sampling variability, we conclude that for FR MCISTG 8.3 there can be no consistent test of the null hypothesis of no effect of exposure controlling for smoking based on our case-control data even under large sample limiting model 1. That is, there is no test which can, even asymptotically, distinguish between the null hypothesis and all possible alternatives. On the other hand, a consistent test can be constructed from full cohort data. Simply compute the NPMLE of  $S(t, G_{H,c}, G_{0,c})$ , which under limiting model 1 is asymptotically unbiased. As an obvious corollary, we cannot estimate the G-causal parameters of FR MCISTG 9.1a or 9.1b from our case-control data (not even the ratio of the odds of disease when treated with  $G_{H,c}$  to the odds of disease when treated with  $G_{0,c}$  is estimable). In a standard point exposure study in which exposure and smoking data are measured only at start of follow-up, we can test the null hypothesis of no exposure effect controlling for smoking from case-control data. We can also estimate the prospective odds ratio of death (or survival) controlling for smoking behavior.

*Example 2.* Example 2 is the same as Example 1, except we now add to our previous data, data from a case-control study of "leaving work" in which at  $t_2$  all individuals who left work and a random sample of controls remaining at work at that time are selected. Data on exposure status at  $t_1$  and on cigarette smoking behavior at  $t_2$  are obtained for each of the cases and controls. The control sampling fraction is unknown. The somewhat surprising result is that although in neither case-control study is the control sampling fraction known, we can use the data from the two studies together to calculate both control

sampling fractions. This result can be seen as a special case of more general results found in Hsieh *et al.*[20]. We leave it to the reader to demonstrate this result for our special case. But if we know the control sampling fractions, we know the population proportions in the FR MCISTG representing the full cohort data. Therefore, we can compute the same population parameters from the combined data in our two case-control studies as we can from the full cohort data.

The conclusion is that, now allowing for sampling variability, we can obtain from our combined case-control data an asymptotically normal and unbiased estimator of the  $G$ -causal parameter representing the effect of exposure controlling for smoking history (under large sample limiting model 1). We do not examine the efficiency of this estimator here.

The estimation of the  $G$ -causal parameter representing the effect of exposure controlling for smoking on the basis of our combined case-control data we would prefer to call a cohort analysis (rather than a case-control analysis) of case-control data, since the analysis inherently (even if indirectly) relied on the fact that the control sampling fractions were estimable. The following definition provides what we shall treat as the fundamental definition of a case-control analysis.

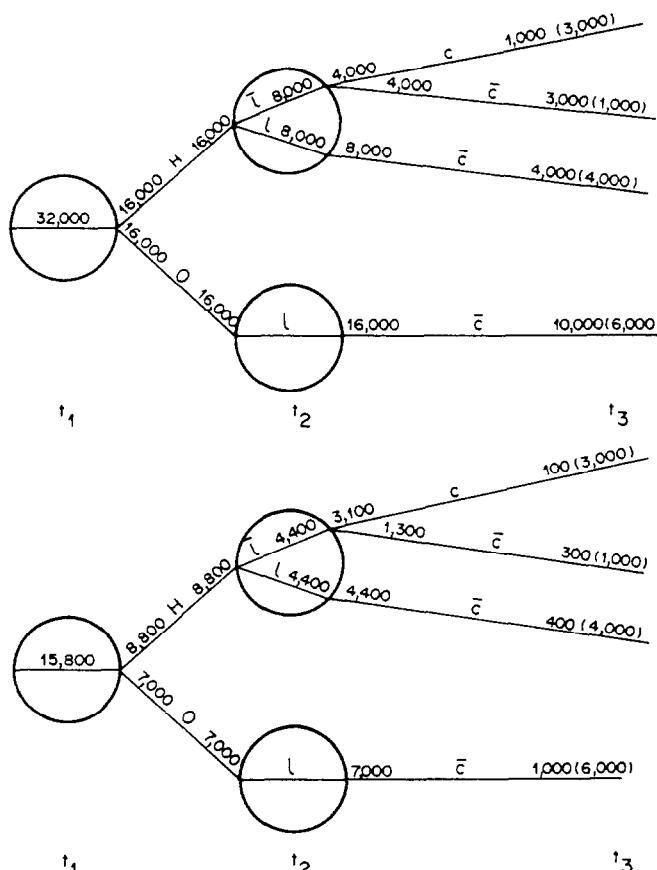


Fig. 9.1a. Population from which the data in Fig. 9.1b was sampled if control sampling fraction = .1. Numbers in parentheses at  $t_3$  = deaths at  $t_3$ , numbers not in parentheses at  $t_s$  = survivors at  $t_s$ . b. Data from a case-control study of deaths. Numbers in parentheses at  $t_3$  = deaths at  $t_3$ , numbers not in parentheses at  $t_s$  = survivors at  $t_s$ .

*Definition.* *Case-control analysis.* When the goal of an analysis is the estimation of a population parameter, we will say that the parameter is identifiable from a case-control analysis if and only if it can be expressed as an odds ratio function of the (joint distribution of the) observed data. Tests of the null hypothesis can be incorporated into this framework by regarding such tests as attempts to estimate the null parameter.

*Example 3.* All corresponding odds ratios in Figs. 9.1a and 9.1b are identical. The reader may check this for himself. Thus, the  $G$ -causal parameters of FR MCISTG 8.3 representing the effect of exposure controlling for smoking (even when all are identically zero) may not be identifiable from a case-control analysis. This is true regardless of the number of case-control studies whose data are combined.

Note, for any FR MCISTG, a consistent test of the  $G$ -null hypothesis under large sample limiting model 1 can be obtained from a case-control study in which the deaths are the cases and the sampling fraction is unknown. [Simply compute the standard omnibus chi-squared test statistic for independence separately for each table contributing to the  $G$ -null test algorithm. Add the test statistics from each table and compare the value to that expected under the null. Since under limiting model 1 the number of observations per cell increases without bound, this test will be consistent. The  $G$ -causal parameters of any FR MCISTG are therefore identifiable from a case-control analysis when they are all identically zero (as would be the case for FR MCISTG 8.3, if there was no effect of cigarette smoking controlling for exposure as well as no effect of exposure controlling for cigarette smoking).]

The above results concerning FR MCISTG 8.3 become quite disturbing when we recognize that often our interest will be in detecting whether or not there exists a small exposure effect controlling for smoking, and the data will be sparse—disturbing, because, in general, the only potentially valid tests of the null hypothesis, Eq. (8.29), require that we first estimate the conditional probabilities of MPISTG 8.3 and then apply the  $G$ -computation algorithm. In sparse data, the validity of these tests will be compromised by unavoidable model misspecification.

One potential way out is to recognize that the null hypothesis of no effect of exposure controlling for cigarette smoking for FR MCISTG 8.3, although not testable from a case-control analysis for the state of nature (parameter values) represented in the above example, is testable based on a case-control analysis for certain other states of nature. For example, if (1)  $L$  is not a predictor of the joint distribution of exposure and cigarette smoking history or (2)  $L$  is a population nonrisk factor controlling for exposure and cigarette history, then, by Lemma 8.4, Eq. (8.9) is the basis of a valid test of Eq. (8.29). The hypothesis that Eq. (8.9) is always unity can be tested using a case control analysis. Of course, in sparse data, if few individuals have the same smoking history, even in testing Eq. (8.9), one must resort to modelling assumptions and sacrifice the nonparametric nature of the test. Since individuals start smoking at similar ages and since giving up smoking is a relatively infrequent event, matching rather precisely on cigarette smoking history in large occupational cohorts may be feasible. (1) and (2) above can themselves, in principle, be tested by case-control analysis. Of course, in occupational mortality studies (1) is obviously false and (2) is false when the healthy worker survivor effect is operative. (If we were interested in the effect of exposure controlling for employment and cigarette smoking history, and if MCISTG F8.3 were an FR MCISTG, we could test the null hypothesis using a case-control analysis by an empirical test of whether exposure was a population risk factor for death controlling for  $L$ - and  $C$ -history. But, of course, if the healthy worker survivor effect is operative, F8.3 will not be an FR MCISTG.)

### B. An algorithm for testing for an effect of exposure controlling for cigarette smoking

A potentially more fruitful approach to testing whether the null hypothesis of Eq. (8.17) holds for FR MCISTG 8.3 is to use the results of Lemmas 8.15 and 8.16 and 8.18. For example, from Lemma 8.16, it follows that if Eq. (8.31) holds, the “*G*-null test” for MPISTG C8.3 provides a valid nonparametric test of the “null hypothesis of no exposure effect controlling for cigarette smoking history”.

Further a nonparametric test of Eq. (8.31), itself, can be constructed from case control data on the “outcome”  $C(t_s)$ . One possible test, among many, is as follows. First, we ample “cases” and controls. Since “change in smoking behavior” between  $t_s$  and  $t_s + \Delta t$  is, in most cohorts, a low probability event, we define any change in cigarette smoking behavior to be a “case”. Each case is identified by (1) the time at which the change in smoking rate occurred (the “failure” time), (2) the case’s past smoking history, and (3) whether the change in smoking rate was an increase or decrease. (We could choose to assign a score for the magnitude of the change but we do not consider that possibility here.) A given individual may be a case many times. In the first pass through the data set, we determine all cases and keep a record of their smoking histories. In a second pass, we select “controls” for a case whose change occurred at  $t_s + \Delta t$  from among individuals, alive at  $t_s + \Delta t$ , with smoking histories identical to that of the case through  $t_s$ , but who did not change smoking behavior at  $t_s + \Delta t$ . For simplicity of exposition we are assuming there is only one case failing at a particular time with a given cigarette smoking history.

Next, in order to have power against the alternatives to Eq. (8.31) such as “leaving work is associated with a decrease in smoking rate”, we relabel as controls, cases whose “change in their cigarette smoking rate” was a decrease; and we relabel their matched controls as cases. Our data now appears as a standard matched case-control study with a mixture of 1-*M* and *M*-1 matching. We now fit, by conditional logistic regression, a model such as  $\ln[p[\text{case}(t_s) | E(t_{s-1}), L(t_s), k]/(1 - p[\text{case}(t_s) | E(t_{s-1}), L(t_s), k])] = B_{0,k} + B_1 ce(t_{s-1}) + B_2 cl(t_s)$ , where  $B_{0,k}$  are nuisance parameters for the matched sets indexed by  $k$ ;  $t_s$  describes the “failure time of the matched set” and  $ce(t)$  and  $cl(t)$  are cumulative exposure and cumulative years off work, respectively. A two-degree-of-freedom score test of  $B_1 = 0$  and  $B_2 = 0$  is a (nonparametric) test of Eq. (8.31). (It can be shown that the contribution to the “likelihood score” are uncorrelated. Therefore standard errors based on the information matrix are valid.) Terms such as an interaction between pack years of cigarettes through  $t_{s-1}$  and cumulative exposure through  $t_{s-1}$  could be added to the right side of the equation. We would then have a three-degree-of-freedom test of Eq. (8.31) that would have greater power against certain alternatives.

Now, if our test of Eq. (8.31) fails to reject, and our “*G*”-null test for MPISTG C8.3 rejects, we “reject” the null hypothesis, Eq. (8.29). If both tests fail to reject we “accept” the null hypothesis. [Even if Eqs. (8.15) and (8.31) both hold, the power of the *G*-null test for C8.3 should be greater than the power of the (now valid) *G*-null test of MPISTG 3.4. This reflects the fact that MPISTG 3.4, even when Eq. (8.15) holds, is not the Stage 0 PL-sufficient reduction of C8.3. In fact, if Eq. (8.15) were known to hold in the population, but, due to sampling variability, it was not empirically true in the data, the *G*-null test based on MPISTG 3.4 would be an invalid test of the “*G*”-null hypothesis for C8.3, conditional upon approximate ancillary statistics that measure the degree to which the empirical version of Eq. (8.15) was false in the data.]

If our test of Eq. (8.31) rejects, we could attempt to take advantage of the observation that, by Lemma 8.15, even if Eq. (8.31) is false, if Eq. (8.30) holds, then Eq. (8.29) implies Eq. (8.23). We can construct a test of Eq. (8.30) using the mixture of 1-*M* and *M*-1 matched sets of cases and controls selected to test Eq. (8.31).

Given a single case (single control) and its *M*-matched controls (cases) with the failure

time at  $t_{s-1}$ , we construct at each  $t_k$ ,  $t_k \leq t_{s+1}$  a  $2 \times 2$  table comparing the employment status of the case (control) at  $t_k$  to that of the subset of his matched controls (cases) whose exposure and employment history agree with that of the case (control) through  $t_{k-1}$ . We combine the information over tables and matched sets by performing a standard Mantel-Haenszel summary test of the null hypothesis of no effect of "employment status" on case-control status.

We can proceed to construct a test of Eq. (8.23) as follows. One matches cases dying at  $t_s + \Delta t$  to controls alive at that time with cigarette smoking histories through  $t_s$  identical to that of the case. To the set of matched sets, one applies the  $G$ -null test algorithm for MPISTG C8.3. Unfortunately, even when both Eqs. (8.30) and (8.23) hold, the table-specific contributions to the numerator of the above test of Eq. (8.23) will usually be correlated. Similarly for the test of Eq. (8.30). Therefore neither test statistic may have a standard normal distribution, and thus the two statistics cannot be recommended for use together as a test of Eq. (8.29) without further modifications which are yet to be developed.

We therefore proceed to test whether Eq. (8.28) holds using a case-control analysis with the joint outcomes  $C$ - and  $L$ -history. We omit the details. If our test of Eq. (8.28) accepts, then by Lemma 8.18, the  $G$ -null test for MPISTG C8.3 and a test of Eq. (8.27) would both be valid tests of Eq. (8.29). For example, if the " $G$ "-null test for MPISTG C8.3 accepted (rejected) we would accept (reject) the null hypothesis Eq. (8.29). If our test of Eq. (8.28) rejects, we have no recourse but to test Eq. (8.29), in a cohort analysis, by modelling the probabilities in MPISTG 8.3 and then applying the  $G$ -computation algorithm. Here we are assuming *a priori* that  $C$  is a causal risk factor for death controlling for exposure history. If not, then a " $G$ "-null test for MPISTG 8.3 would be a valid test of Eq. (8.29).

If we believe that MCISTG C8.3 (rather than MCISTG 8.3) were our finest FR MCISTG, we might accept the null hypothesis, Eq. (8.17), if both Eq. (8.22) and (8.23) hold (see the discussion following Lemmas 8.11 and 8.14). We have previously described a test of Eq. (8.23). A test of Eq. (8.22), similar to the test of Eq. (8.30), can be based on the mixture of  $1-M$  and  $M-1$  matched sets of cases and controls previously selected to test Eq. (8.31). In particular, given a single case (single control) and its  $M$  matched controls (cases) with the failure time at  $t_{s+1}$ , we construct at each  $t_k$ ,  $t_k \leq t_s$ , a  $2 \times 2$  table comparing the exposure status of the case (control) at  $t_k$  to that of the subset of its matched controls (cases) whose exposure status through  $t_{k-1}$  and employment status through  $t_k$  agree with that of the case (control). Corollary E1 can be generalized to show that when both Eqs. (8.22) and (8.23) hold, the table-specific contributions to the numerators of each test are uncorrelated. It follows that both test statistics will have an (asymptotic) standard normal distribution. The test statistics for the tests of Eq. (8.22) and Eq. (8.23) can be shown to be uncorrelated with one another, as well. If either or both the tests of Eqs. (8.22) and (8.23) reject, we know only that Eq. (8.19) or (8.20) or both are false (ignoring sampling error).

In the next sub-section we return to the simpler setting of an occupational mortality study in which only data on exposure and employment history are available. We discuss whether previously proposed designs for nested case-control studies are valid when the healthy worker survivor effect is operating.

### *C. The bias of previously suggested design and analysis strategies for nested case-control studies*

We have seen that when MCISTG 3.4 is our finest FR MCISTG, the  $G$ -null test algorithm provides a valid test of the null hypothesis of no exposure effect on any person's

mortality when controlling for employment history (provided employment history is not an independent causal risk factor) from case-control data. We now discuss case-control design and analysis strategies commonly used in occupational epidemiology and show that all are potentially biased for testing the  $G$ -null hypothesis of FR MCISTG 3.4.

In occupational epidemiology, three commonly suggested case-control designs and analysis strategies (hereafter DA strategies) are to randomly sample controls from those individuals at risk at the failure age of the case (preferably, matched to the case on date of birth and age at hire) who: (DA1) were sampled without regard to time of termination of employment; or (DA2) terminated employment at the same time as the case; or (DA3) were employed at least as long as the case. In a matched analysis, a summary measure of each case's lifetime exposure history (e.g. lifetime cumulative exposure or lifetime cumulative exposure lagged some number of years) is compared to that of his matched controls. Under DA 1 and 2 the control's exposure history up to the death age of the case is summarized, while under DA 3 the control's exposure history subsequent to the termination age of the case is ignored. DA 2 and 3, by comparing the exposure history of a case and its matched controls only while both are at work, insure, albeit by different strategies, that the case and their controls have an "equal opportunity for exposure" [21].

We also consider a design and analysis strategy, DA 4, in which controls are sampled at random as in DA 1, and, then, at each time  $t$  at which the case is at work (i.e. has an opportunity for exposure), the exposure concentration of the case at that time is compared with that of the subset of his matched controls who are also at work at  $t$ . This design and analysis strategy has been proposed independently by the author and by Allan Smith.

Lemmas 8.1 and 8.2 indicate that DA 1 will in general be biased under the  $G$ -null hypothesis of FR MCISTG 3.4 unless  $L$  is not an independent risk factor for death or  $L$  is not a predictor of exposure. Thus, DA 1 will be biased when the healthy worker survivor effect is operative. From Lemma 8.3, we see that DA 2 will be biased unless  $L$  is not an independent risk factor for death or exposure is not a predictor of future  $L$ -history. Thus, if the healthy worker survivor effect is operative and exposure is, for example, an irritant that increases the termination rate in the exposed (or if high exposure jobs are more or less socioeconomically desirable than low exposure jobs; or if high exposure and low exposure jobs have different layoff rates), DA 2 will be biased.

DA 3 can be biased even when  $L$  is not an independent risk factor for death (i.e. even when there is no healthy worker survivor effect) if exposure is a predictor of future  $L$ -history. To see this, consider a "point-exposure" occupational study in which workers are randomly assigned to receive high or zero exposure at start of follow-up and are never exposed thereafter. Suppose a random sample of 50% of the highly exposed workers leave work early due to irritation effects or for socioeconomic reasons. Suppose no unexposed workers leave work. Then, under the null, employment history will not be a risk factor. Any individual who dies while off work must have been highly exposed. Yet his controls under DA 3 consist of a random sample of highly exposed and unexposed individuals who survived at work at least as long as he did. Thus, under the null, there will be an apparent excess of exposure among the cases dying while off work. For individuals dying at work, there will be no bias. Thus, a summary Mantel-Haenszel test will falsely reject the null.

Finally, from its definition, DA 4 will be valid only if, among individuals at work at time  $t$ , exposure is received at random irrespective of past exposure history (since past exposure history is not matched on in DA 4). But FR MCISTG 3.4 only assumes that exposure is received at random among individuals at work at time  $t$  with identical past exposure and work histories. Now suppose exposure was an upper airway irritant that produced cough (but had no effect on mortality). Then workers with chronic lung disease might be less likely than healthy workers to be able to tolerate the irritating effects of exposure, and thus would preferentially terminate employment. If so, individuals with

high exposure at work at time  $t$  would be healthier than those with lower exposure and thus less likely to be cases at  $t + \Delta t$ . Thus, DA 4 will be biased.

## 10. ARTIFACTUAL "EMPIRICAL HEALTHY WORKER SURVIVOR EFFECTS"

### A. An artifactual "empirical healthy worker survivor effect" due to model misspecification

An investigator may base his causal inferences on FR MCISTG 3.4 if  $L$  is an independent population risk factor and on FR MCISTG 3.5 if it is not (see Sec. 8A.3). As such, an investigator's first task is to empirically test whether  $L$  is a population risk factor. Due to limitations of sample size, this can involve modelling  $\gamma_D[t + \Delta t | E(t), L(t)]$ . Model misspecification can lead one to conclude that  $L$  is a strong population risk factor when, in fact, it is not. For example, consider the model of Eq. (5.1) with  $\beta_D \cdot X_D$  as defined in Appendix D. Suppose, in reality, that Eq. (5.1) is correctly specified when, in  $\beta_D \cdot X_D$ , cumulative exposure  $ce(t)$  is replaced by cumulative exposure up to 9 years previously, i.e.  $ce(t - 9)$  (due, for example, to a biological latent period). Suppose, in reality,  $L$  is a population nonrisk factor, i.e.  $\beta_{2,D}$  and  $\beta_{3,D}$  are 0 and  $\beta_{1,D} = \beta_{4,D}$  in the correctly specified model. If such is the case, inference based on the misspecified  $\beta_D \cdot X_D$  of Appendix D will suggest  $L$  is a population risk factor. To see this, consider two groups of workers at risk at age  $t$  with identical cumulative exposures up to  $t$ . Suppose all members of one group left work many years ago. The members of the other left work only one year previously. On average, individuals who has been out of work for many years will have attained a higher level of cumulative exposure at  $t - 9$  years. Thus, the group which terminated employment many years previously will be at greater risk for death at  $t$ . It follows that the misspecified model will incorrectly estimate  $\beta_{3,D}$  to be positive in order to account for the increasing risk with years off work when controlling for cumulative exposure. One might then discover an artifactual empirical healthy worker survivor effect upon performing the computations outlined in Appendix D using the misspecified model. In our arsenic data set, the estimates of  $\beta_{2,D}$  and  $\beta_{3,D}$  were little effected by using cumulative exposure lagged 5, 15, or even 20 years.

### B. An artifactual healthy worker survivor effect due to "Lay-Off"

Investigators who have attempted to demonstrate the existence of an empirical healthy worker survivor effect have often (1) defined date of termination as date of last employment and (2) ignored all other information on employment history. In studies of an industry in which workers frequently leave and then return to work (for example, due to economic layoff), such a practice will artifactually create the appearance that the healthy worker survivor effect is operating even when it is not. Furthermore, incorrect causal inferences

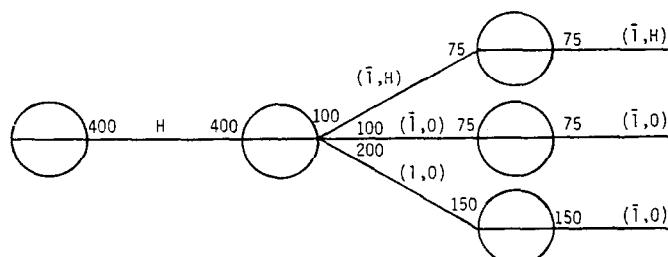


Fig. 10.1. An FR MCISTG in which all surviving laid off workers return at  $t_3$ .

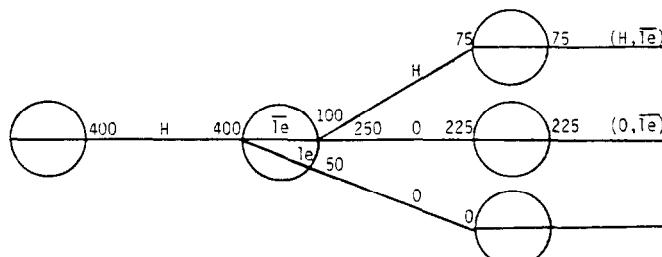


Fig. 10.2. An MPISTG derived from FR MCISTG 10.1 where an individual is *le* after leaving work for the last time.

will be made if one identifies such an occupational mortality study with a randomized trial in which 'date of last employment' is 'date of termination of protocol'.

*Example.* Figure 10.1 is an FR MCISTG representing the results of an occupational mortality study. Since we have supposed that the true state of nature is that Fig. 10.1 is an FR MCISTG, there is no actual healthy worker survivor effect operative. Also note employment history is not a cause of death on controlling for exposure history. Likewise, exposure has no effect on mortality either overall or when controlling for employment history. Of the 200 individuals who left work at  $t_2$ , all 150 surviving at  $t_3$  return to work.

An investigator who defined date of termination as date of last employment would create the MPISTG shown in Fig. 10.2 where an individual at  $t$  is *le* if he has terminated employment at a time less than or equal to  $t$  and is  $\bar{le}$  otherwise. *le* status is an enormously strong population risk factor for death on controlling for exposure history. Many investigators have incorrectly interpreted this observation as evidence that "unhealthy individuals tend to leave work". Of course, in our example *le* status is a population risk factor only because, when we define date of termination as date of last employment, the future event, death, determines the previous event "date of termination". If an investigator assumed that MPISTG 10.2 was an FR MCISTG, he would falsely discover an adverse effect of high exposure. The problem, of course, is that MPISTG 10.2 is not an FR MCISTG. The 100 individuals with covariates (*le*, *H*) measured at  $t_2$  are not randomized with respect to the 250 ( $\bar{le}$ , 0) individuals, since the 150 ( $\bar{le}$ , 0) individuals who had been off work at  $t_2$  are guaranteed to survive to  $t_3$  by the very fact that they are  $\bar{le}$  rather than *le*.

## 11. GILBERT'S PROPOSED METHOD FOR CONTROLLING FOR THE HEALTHY WORKER SURVIVOR EFFECT AND THE CONCEPT OF MINIMUM LATENT PERIOD

As mentioned in the introduction, Gilbert[1] noted that when the healthy worker survivor effect is operative, the mortality-cumulative exposure association would be biased, whether or not one adjusted for employment history. Gilbert assumed for most cancers that (G.1) the minimum latent period exceeds 10 years and that (G.2) the healthy worker survivor effect is operative for less than 10 years after leaving work. (She assumed that G.2 followed from the empirical observation that most cancers will cause death within 10 years of the onset of clinical disease if they are to cause death at all. We will show in Secs. 11C-11E that this empirical observation is necessary but not sufficient for G.2 to hold.) She suggested that if these assumptions held, the association of mortality with observed exposure history lagged ten years would be causal. She therefore proposed that

the investigator compute the ratio

$$\frac{\gamma_{D_1}(t + \Delta t | E(t - 10))}{\gamma_{D_1}(t + \Delta t | E(t - 10) = 0)}$$

where  $D_1$  is death from the cancer of interest and  $E(t - 10)$  is observed exposure history up until 10 years prior to  $t$ .

In this section we provide formal definitions of Assumptions  $\mathcal{G}.1$  and  $\mathcal{G}.2$  above. We then demonstrate that for death from all causes Gilbert's proposed parameter, modified to  $S(t + \Delta t | E(t - x)) - S(t + \Delta t | E(t - x) = 0)$  [as defined in Eq. (11.3c)], is indeed causal under Assumptions  $\mathcal{G}.1$  and  $\mathcal{G}.2$ , provided employment history has no causal effect on mortality controlling for exposure history, which we shall call Assumption  $\mathcal{G}.3$ . This latter assumption was implicit in Gilbert[1]. In Sec. 12, we consider the conditions necessary for our results to extend to the competing risk case. Many recent analyses of occupational lung cancer mortality have been based on Gilbert's proposed parameter (e.g. Refs. [22, 23]), and therefore have implicitly assumed that Assumptions  $\mathcal{G}.1$ ,  $\mathcal{G}.2$ , and  $\mathcal{G}.3$  held for lung cancer. We show in Sec. 11E that, in our arsenic data set, Assumption  $\mathcal{G}.2$  apparently fails to hold for lung cancer.

#### A. Formal definition of Assumption $\mathcal{G}.2$

Since we have considered outcomes to be deterministic, at each time  $t_s$  there must exist a well-defined set of "x-doomed" individuals who will be dead within  $x$  years from  $t_s$  regardless of their future exposure or work history. *We say that the healthy worker survivor effect is operative for less than  $x$  years*, if and only if, at each time  $t_s$  employment and exposure status are received by the subset of individuals who are not x-doomed at random conditional on past employment and exposure history. Let  $ls$  be a time-dependent covariate such that  $ls(t) = 1$  if an individual is doomed to die before  $t + x$  and  $ls(t) = 0$  (also written  $\bar{ls}$ ) otherwise. Then, the healthy worker effect is operative for less than  $x$  years if Fig. 11.1 is an FR OCISTG. In Fig. 11.1 we have assumed  $x = 2\Delta t$ . This is shown by vertical bars that block internodal lines. Such bars serve to indicate that no individual in the subset represented by a particular internodal line survives past the time at which the bar is blocking the line. Since data on  $ls$  status is not available, we cannot estimate the  $G$ -causal parameters of FR OCISTG 11.1 without further assumptions. (Figure 11.1 is the appropriate graph when  $x/\Delta t = 2$ . The situation of interest considered by Gilbert has  $\Delta t \rightarrow 0$ ,  $x = 10$  years. The corresponding graph will be exactly like Fig. 11.1 except with vertical bars drawn  $10/\Delta t$  generations after a subgroup first becomes  $ls$ . Due to lack of space we obviously cannot draw this graph.)

Throughout the remainder of this paper, we shall assume that  $\Delta t = 1$  year and that  $x$  is measured in years.

*Remark.* If we assume, for convenience, that MPISTG 3.3 is an MCISTG and that the death time for any individual who is  $x$ -doomed at  $t_s$  is uninfluenced by exposure or employment status experienced at later times, then  $S(t | G_1^{3.3}, i) = S(t | G_1^{11.1}, i)$  whenever  $G_1^{3.3}$  and  $G_1^{11.1}$  are characterized by the same  $E_1(t_s), L_1(t_s)$ .

#### B. Definition of the minimum latent period

At an individual level, one might wish to express the statement that  $x$  is less than the minimal biologic latent period (Assumption  $\mathcal{G}.1$ ) by the nonidentifiable relationship

$$\gamma_D(t + \Delta t | E(t), i) = \gamma_D(t + \Delta t | E(t - x), i) \quad (11.1)$$

for each individual  $i$ . One might expect that if an increment of exposure given at  $t$  could have no direct biological effect on (recorded) mortality for at least  $x + \Delta t$  years, Eq. (11.1) would hold. This is not the case because it is possible for exposure to have no biological effect on outcome whatsoever, but Eq. (11.1) to be false, if, for example, high exposure were an irritant that causes people to terminate employment and then the effects of termination per se (i.e. loss of health insurance, poverty, etc.) cause death within  $x$  years. Thus a better definition of “ $x$  is less than the ‘biologic minimum latent period’” is that

$$\gamma_D(t + \Delta t | E(t), L(t), i) = \gamma_D(t + \Delta t | E(t - x), L(t), i) \quad (11.2)$$

(i.e. that exposure has no effect on any individual’s mortality for  $x + \Delta t$  years when controlling for employment history). Gilbert did not consider the distinction between Eqs. (11.1) and (11.2) because she implicitly assumed that  $L(t)$  was not an independent causal risk factor controlling for exposure. At the moment we will not completely exclude the possibility that  $L$  is an independent causal risk factor, although we will suppose, following Gilbert, that it too has a minimum latent period of at least  $x$  years. That is, we will express Assumption G.1 as

*Assumption G.1.*

$$\gamma_D(t + \Delta t | E(t), L(t), i) = \gamma_D(t + \Delta t | E(t - x), L(t - x), i). \quad (11.3)$$

We now generalize the concept of the minimum latent period so that it applies to any CISTG.

*Definition.* For a given  $t_s$ , “ $G_1^A(t_s)$ ” is the portion of the highlighted subgraph of “ $G_1^A$ ” from  $t_1$  to  $t_s$ .

*Definition.* Given a CISTG A, we say that  $x$  is less than the generalized minimum latent period if and only if for all  $t$  and all individuals  $i$ ,  $\gamma_D(t + \Delta t | i, G_1^A) = \gamma_D(t + \Delta t | i, G_2^A)$  if “ $G_1^A(t - x)$ ” = “ $G_2^A(t - x)$ ” (that is, the highlighted subgraphs up to  $t - x$  are the same).

**LEMMA 11.1.** If Eq. (11.3) holds, then  $x$  is less than the generalized minimum latent period of OCISTG 11.1 (and MCISTG 3.4). (Note that Fig. 11.1 is not required to be an FR CISTG for the lemma to hold.)

*Proof.* Obvious.

**LEMMA 11.2.** If  $x$  is less than the generalized minimum latent period of an R CISTG A, then

$$\gamma_D(t + \Delta t | "G_1^A") = \gamma_D(t + \Delta t | "G_2^A") \quad \text{when } "G_1^A(t - x)" = "G_2^A(t - x)". \quad (11.3a)$$

*Proof.* See Ref. [7].

*Definition.* When Eq. (11.3a) holds for a PISTG A, we say that  $x$  is less than the *identifiable* generalized minimum latent period.

Given a PISTG A we construct a Stage 0 reduction PISTG,  $A(x)$ , for each positive integer  $x$  as follows. (Figures 11.1 and 11.2 are examples of A and  $A(x)$  for  $x = 2$ .)

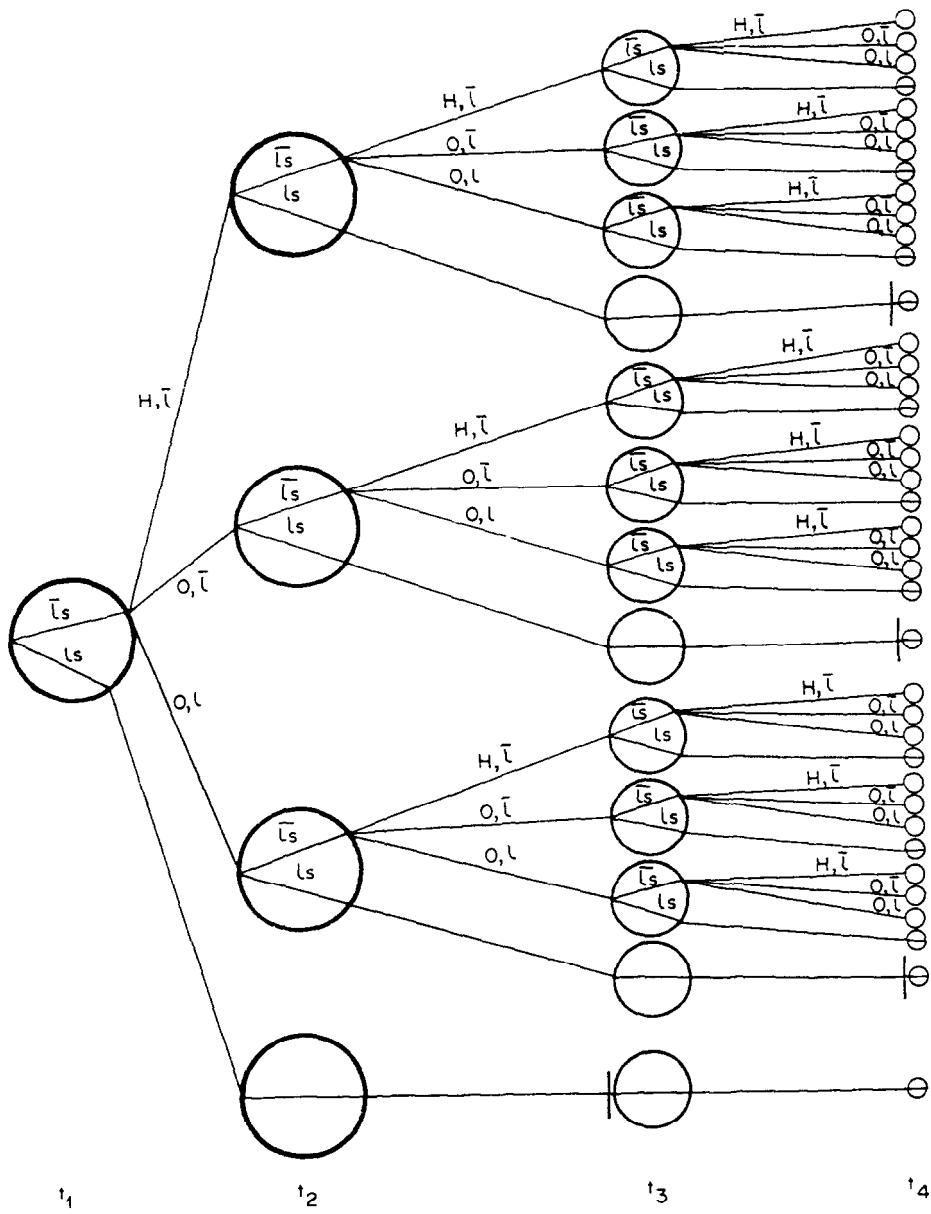


Fig. 11.1. An FR OCISTG.  $H$  = high exposure concentration,  $O$  = unexposed,  $\bar{l}$  = at work,  $l$  = off work,  $ls$  = doomed to die within  $x = 2\Delta t$  years,  $\bar{ls}$  = not doomed to die within  $x = 2\Delta t$  years.

*Step 1.* STG  $A(x)$  is STG  $A$  with  $x$  single nodes (each with a single inter- and intranodal line) added onto the left of the graph. The generations of  $A(x)$  are labelled beginning with  $t_1$ . There is a one-to-one correspondence between  $A(x)$  at  $s$ , for  $s \geq x + 1$ , and  $A$  at  $t_{s-x}$ . The corresponding intra- and internodal lines in  $A(x)$ ,  $s \geq x + 1$ , and  $A$  are said to be images of one another.

*Step 2.* If  $s \leq x$ ,  $[i_s^{A(x)}]$  [in the standard labelling of  $A(x)$ ] is the subset of individuals in the entire study population who survived to  $t_s$ . If  $s > x$ ,  $[i_s^{A(x)}]$  and  $[i_s j_s^{A(x)}]$  are those individuals in the image sets  $[i_{s-x}^A]$  and  $[i_{s-x} j_{s-x}^A]$  who survived to  $t_s$ .

There is a natural many-to-one map from the set of "generalized treatments" of PISTG

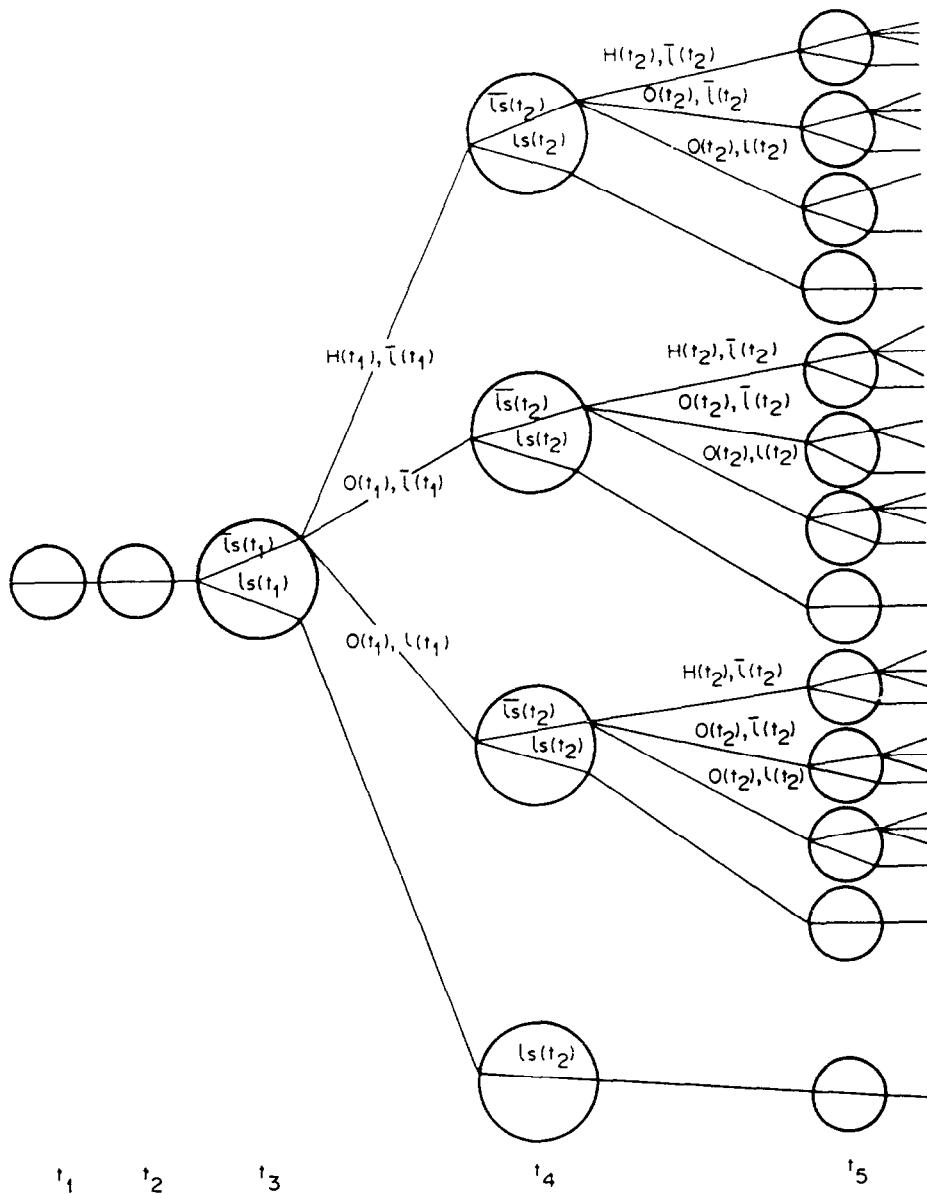


Fig. 11.2. An MPISTG.

$A$  onto those of  $A(x)$ . For any “ $G^A$ ”, consider the highlighted subgraph “ $G^A$ ”. On  $A(x)$ , highlight the subgraph representing the images of “ $G^A(t_s - x)$ ” plus the single intranodal and internodal lines from  $t_1$  to  $t_x$ . The result will be a unique generalized treatment of  $A(x)$ , the Stage 0 counterpart of “ $G^A$ ”. Thus,  $A(x)$  is a  $AB$ -complete Stage 0 reduction of  $A$ , although it is not unique  $B$ -complete.

**THEOREM 11.1.** If  $x$  is less than the identifiable generalized minimum latent period of PISTG  $A$ , then  $S(t | "G^A") = S(t | "G^{A(x)}")$  where “ $G^{A(x)}$ ” is the Stage 0 counterpart of “ $G^A$ ”. Furthermore, MPISTG  $A(x)$  is the  $AB$ -complete PL-sufficient Stage 0 reduction of PISTG  $A$  under Eq. (11.3a).

*Proof.* See Ref. [7].

**THEOREM 11.2.** If for OPISTG 11.1 the following three assumptions hold—(G.1)  $x$  is less than the minimum latent period (i.e. Eq. (11.3) holds], (G.2) the healthy worker survivor effect is operative for less than  $x$  years, and (3)

$$\gamma_D[t + \Delta t | E(t - x), L(t - x)] = \gamma_D[t + \Delta t | E(t - x)] \quad (11.3b)$$

—then, for any  $G_1^{11.1}$  characterized by “stay at work and receive, at  $t_s$ , exposure  $e_1(t_s)$ ”

$$S(t_{s+1} | G_1^{11.1}) = S[t_{s+1} | E_1(t_s - x)] = \prod_{k=1}^s [1 - \gamma_D[t_k + \Delta t | E_1(t_k - x)]]. \quad (11.3c)$$

*Proof.* Assumption G.1 implies by Lemma 11.1 that  $x$  is less than the generalized minimum latent period of OCISTG 11.1. By Assumption G.2, OCISTG 11.1 is an FR OCISTG. By Lemma 11.2, Eq. (11.3a) holds for FR OCISTG 11.1. Thus, by Theorem 11.1, OPISTG 11.2 is the AB-complete PL-sufficient Stage 0 reduction of FR OCISTG 11.1. The subsets on OPISTG 11.2 at  $t_s$  represented on the graph by  $ls(t_s - x)$  are empty, since by Assumption G.2,  $ls$  individuals do not survive  $x$  years. Thus, OPISTG 11.2 is actually an MPISTG and can be represented by MPISTG 3.3(x). Furthermore, given Eq. (11.3b) holds, Theorem F1 implies that MPISTG 3.5(x) is the AB-complete PL-sufficient Stage 0 reduction of MPISTG 3.3(x) under Eq. (11.3b) [and the AB-complete PL-sufficient Stage 0 reduction of FR OCISTG 11.1 under G.2, Eq. (11.3a), and Eq. (11.3b)]. Applying the  $G$ -computation algorithm [equivalently, Eq. (4.7)] to MPISTG 3.5(x) proves the theorem. (Note Theorem 11.2 holds even if PISTG 3.4 is not an FR MCISTG, as, for example, if  $x$ -doomed individuals alive at  $t_s$  preferentially transfer from highly exposed to unexposed jobs, as in the mining industry.)

In Sec. 11F we show that, given Assumptions G.1 and G.2 hold, Eq. (11.3b) is implied by

*Assumption G.3.*  $L$  is not a causal risk factor for death controlling for exposure (as defined in Sec. 8A.2).

Thus, if Assumptions G.1–G.3 hold, the NPMLE of the  $G$ -causal parameters of FR OCISTG 11.1 do not depend on the data through  $L(t)$  or  $ls$ -status. We would therefore only wish to collect data on exposure and vital status and estimate Gilbert's parameter. Of course, we do not usually have *a priori* knowledge that Assumptions G.1–G.3 hold. We cannot empirically prove these assumptions hold, since they are nonidentifiable even if data on  $ls$ -status is included. But we can empirically rule them out. It is obvious that tests of Assumptions G.1–G.3 must be based on MPISTG 3.4 when data on  $ls$ -status is missing.

### C. Tests of Assumptions G.1 and G.2

Tests of whether Assumptions G.1 and G.2 hold are based on the following theorems.

**THEOREM 11.4.** If MPISTG 3.4 is an MCISTG and Assumption G.1 [i.e. Eq. (11.3)] holds, then for all  $\cdot i_s$ ,  $G_1^{3.4(\cdot i_s)}$ ,  $G_2^{3.4(\cdot i_s)}$ ,

$$\gamma_D(t | G_1^{3.4(\cdot i_s)}) = \gamma_D(t | G_2^{3.4(\cdot i_s)}) \quad \forall \Delta t + t_s + x \geq t \geq t_s. \quad (11.4)$$

**THEOREM 11.5.** If Assumption G.2 holds and the minimum latent period for employment history controlling for exposure history is greater than  $x$  [i.e. Eq. (11.2) holds with

the roles of  $L$  and  $E$  reversed], then

$$\begin{aligned}\gamma_D(t_s + x + \Delta t | "G_1^{3.4[\cdot i_{s-1}t_{s-1}(t_s)]}" , [\cdot i_s^{3.4}]) \\ = \gamma_D(t_s + x + \Delta t | "G_1^{3.4[\cdot i_{s-1}t_{s-1}(t_s)]}" , [\cdot i_s^{3.4'}])\end{aligned}\quad (11.5)$$

for all nodes  $\cdot i_{s-1}t_{s-1}(t_s)$ , but only for the generalized treatment  $G_1$  of FR MCISTG 3.4( $\cdot i_{s-1}t_{s-1}(t_s)$ ) “if at work, receive zero exposure” where  $[\cdot i_s^{3.4}]$  are those individuals at work at  $t_s$  and  $[\cdot i_s^{3.4'}]$  are those individuals off work at  $t_s$ . That is, consider any node at  $t_s$  and compare the incidence of death at  $t_s + x + \Delta t$  of the subset of individuals who were off work at  $t_s$  with that of the subset of individuals who were on work at  $t_s$  when subsequent to  $t_s$  both are treated with the generalized treatment “if at work receive zero exposure”. If Assumption G.3 holds, the theorem will be true with  $t_s + x + \Delta t$  replaced by any  $t$  such that  $t \geq t_s + x + \Delta t$ .

**THEOREM 11.6.** If Assumptions G.1 and G.2 both hold, then

$$\begin{aligned}\gamma_D[t + \Delta t | L(t - x - \Delta t), E(t - x - \Delta t), l_1(t - x), e_1(t - x)] \\ = \gamma_D[t + \Delta t | L(t - x - \Delta t), E(t - x - \Delta t), l_2(t - x), e_2(t - x)].\end{aligned}\quad (11.6)$$

#### *Proofs of Theorems 11.4–11.6 in a discursive manner*

*Proof of Theorem 11.4.* Theorem 11.4 follows at once from the fact that Assumption G.1 implies that the sharp null hypothesis holds for MCISTG 3.4( $\cdot i_s$ ) for  $x + \Delta t$  years of follow-up (since the sharp null hypothesis implies the  $G$ -null hypothesis expressed in the theorem).

*Proof of Theorem 11.5.*  $\bar{l}_s$  individuals off work in any given node at  $t_s$  are comparable with respect to mortality to  $\bar{l}_s$  individuals on work, since by Assumption G.2, Fig. 11.1 is an FR OCISTG. Therefore, if the  $\bar{l}_s$  individuals off work at  $t_s$  receive exactly the same subsequent exposure history as  $\bar{l}_s$  individuals on work, they will still be comparable at  $t_s + x$  since by assumption  $L$  has a minimum latent period of  $x$  years. Furthermore, the  $\bar{l}_s$  individuals off work and on work at  $t_s$  surviving to  $t_s + x$  are, respectively, the set of all individuals off work and on work at  $t_s$  surviving to  $t_s + x$  by Assumption G.2. Thus, the incidence of death at  $t_s + x + \Delta t$  will be exactly the same for individuals off work at  $t_s$  as individuals on work at  $t_s$  [by Eq. (11.2) with  $L$  and  $E$  interchanged]. The only generalized treatment which is guaranteed to give individuals off work and on work at  $t_s$  the same subsequent exposure history is “if at work, receive zero exposure”.

*Proof of Theorem 11.6.* If G.1 and G.2 hold, then the survival of subsets of  $\bar{l}_s$  individuals at any given node at  $t_s - x$  who differ in  $l(t_s - x)$  and  $e(t_s - x)$  will be identical through  $t_s + \Delta t$  [since the groups were randomized at  $t_s - x$  and Eq. (11.3) holds]. By G.2, the set of individuals at  $t_s - x$  who survive past  $t_s$  is exactly the set of  $\bar{l}_s$  individuals at  $t_s - x$  who survived past  $t_s$ . This proves the theorem.

We now discuss how one can, in theory, use Theorems 11.4–11.6 to test whether Assumptions G.1 and G.2 hold. We first suppose that MPISTG 3.4 is an FR MCISTG.

Consider the following lagged exposure test algorithm. (1) Test, using the method given in (2) below, whether (at some prespecified  $\alpha$ -level) the minimum latent period is greater than  $x$  years for  $x = \Delta t, x = 2\Delta t, x = 3\Delta t$ , etc. until one finds an  $x = n\Delta t$  at which the test rejects. Declare  $x = (n - 1)\Delta t$  to be the greatest  $x$  less than the minimum latent

period. (2) To test whether  $x$  is less than the minimum latent period use any  $G$ -null test for FR MCISTG 3.4 such that  $w(\cdot i_s, t_m) = 0$  if  $t_s < t_m - x$  (i.e. ignore all tables more than  $x$  years before a death). It is straightforward to check that such a  $G$ -null test is a test of Eq. (11.4). (3) Test whether the healthy worker effect lasts for a period of time less than  $x$  [declared in (1) above] as follows: Apply the  $G$ -null test algorithm to MPISTG 3.3 (not FR MCISTG 3.4), using a weight function such that  $w(\cdot i_s, t_m) = 0$  if  $t_s \neq t_m - x - \Delta t$ , except modify the  $G$ -null test so that we construct for a given table  $(\cdot i_{t_m-x-\Delta t}, t_m)$  a  $2 \times 2$  table of cases and controls by at work and not at work status (i.e. we deviate from the usual  $G$ -null test algorithm in combining the treatments "high and low exposure at work"). Since each  $\cdot i_s$  of MPISTG 3.3 represents a unique  $E(t_s)$ ,  $L(t_s)$ , and since we are presumably at the minimum latent period, this is a test of whether Eq. (11.6) holds. If the test does not reject, we assume that  $\mathcal{G}.1$  and  $\mathcal{G}.2$  hold simultaneously for less than  $x$ .

If, as in the mining industry, we do not wish to assume MCISTG 3.4 is an FR MCISTG, the lagged exposure algorithm must be modified [since the  $G$ -null test statistic described in Step (2) above will not have expectation 0 for  $x = k\Delta t$  for all  $k\Delta t$  less than the minimum latent period]. Rather, we can use the following modified algorithm.

(1) For each successive value of  $x$  ( $x = 0, x = \Delta t, x = 2\Delta t, x = 3\Delta t$ , etc.) construct the test of "work status" described in Step (3) of the lagged exposure algorithm above.

(2) For the same successive values of  $x$  construct a similar test for the "effect of exposure at work" based on MPISTG 3.3 in which for each table of the form  $(\cdot i_{t_m-x-\Delta t}, t_m)$  we construct a  $2 \times K$  table of cases and controls by the  $K$  levels of exposure received at work. Individuals off work do not contribute to the test statistic. We then construct a one degree-of-freedom test statistic based on the sum of the  $2 \times K$  table-specific Mantel-Haenszel test for trend statistics.

(3) Let  $n\Delta t$  be the first value of  $x$  such that neither of the tests described in (1) or (2) above rejects. Let  $n'\Delta t$  be the last value of  $x$  at which both tests fail to reject. Declare  $n'\Delta t$  to be the latest time at which Assumptions  $\mathcal{G}.1$  and  $\mathcal{G}.2$  both hold. If there is no  $x$  for which neither test rejects, assume that there is no  $x$  for which  $\mathcal{G}.1$  and  $\mathcal{G}.2$  both hold. This algorithm is simply a test of Eq. (11.6) broken into two pieces, one piece testing for the work status effect and the other for an exposure effect at work. If Eq. (11.6) is true, then for any fixed value of  $x$ , the two test statistics are uncorrelated.

If we have a greatest  $x$  for which  $\mathcal{G}.1$  and  $\mathcal{G}.2$  both hold, then we can test Assumption  $\mathcal{G}.3$  by testing whether

$$\gamma_D[t + \Delta t | E(t - x), L(t - x)] = \gamma_D[t | E(t - x)] \quad \forall L(t - x). \quad (11.7)$$

Tests of Eq. (11.7) will require modelling in sparse data.

The lagged exposure test algorithms, although theoretically pure, have serious weaknesses. The power to detect the true minimum latent period and/or the duration of the healthy worker survivor effect, when short, may be very poor in sparse data. This is because few cases will have "controls" with identical employment and exposure history until a few years before the case fails. The cases that will have such controls tend to be cases that died soon after hire. They probably are a quite unrepresentative sample of all cases. Further problems with the algorithm are that tests for successive years are correlated. Also, it is not clear that we have optimally exploited the fact that if Assumption  $\mathcal{G}.1$  is false for  $x$ , it is false for  $x + \Delta t$ .

Often we would be willing to assume *a priori* that Assumption  $\mathcal{G}.3$  holds. In Sec. 11E, we show that we can use this assumption to develop tests with reasonable power that are able to demonstrate that the healthy worker survivor effect lasts for at least  $x$  years even when  $x$  is greater than the minimum latent period.

#### D. A substantive interpretation of Assumption G.2

Since the  $\alpha$ -level chosen for a test of Assumption G.2 should, in an informal Bayesian sense, depend on one's prior degree of belief as to the likelihood that the assumption is true, we need to better understand the substantive meaning of G.2. Gilbert implied that the healthy worker effect for lung cancer is operative for less than 10 years because individuals with symptomatic lung cancer (almost) never survive 10 years. For the purposes of this discussion assume lung cancer is the only cause of death (see Sec. 12 for discussion of this assumption). Implicit in her logic (and explicit in our Assumption G.2) is the assumption that, among individuals at work at  $t - \Delta t$  who are not doomed to die within 10 years (and are thus without symptomatic lung cancer), those leaving work at  $t$  are on average at no higher risk for lung cancer mortality at times greater than  $t + 10$  than individuals remaining at work at  $t$  (when controlling for exposure history up to time at risk). (Here we are supposing, as did Gilbert, that employment history is not an independent causal risk factor.) Assumption G.2 is almost certainly incorrect even for lung cancer if, as in most occupational health studies, data on cigarette smoking history has not been obtained. To see this, note that individuals terminating employment at time  $t$  will have, in general, a greater history of exposure to cigarettes when compared to similar individuals continuing employment (since terminees will include disabled individuals with nonmalignant lung disease and heart disease induced by cigarettes). (In fact, socioeconomic factors alone may also lead to terminees having a greater history of cigarette smoking. For instance, early terminees are more likely to be socially maladjusted individuals who cannot hold a job. Such social maladjustment is probably highly associated with the cigarette smoking habit.) In the absence of the ability to control for cigarette smoking history, terminees (even without lung cancer) will be at an increased risk of death from lung cancer for many years to come compared to individuals surviving them at work (since heart disease, nonmalignant respiratory disease, and social maladjustment, unlike lung cancer, are not rapidly fatal). Thus, even when exposure has a minimum latent period of  $x$  years and the disease of interest is uniformly fatal within  $x$  years of clinical onset, one must empirically check, using the test described above, whether causal parameters can be defined in terms of the association of mortality with observed exposure history lagged some number of years. Furthermore, even if data on cigarette smoking had been obtained, the healthy worker effect for lung cancer might still be operative for more than 10 years when controlling for imperfectly measured cigarette smoking history. This reflects the fact that among individuals with identical past measured covariate histories, including measured cigarette history, those who leave work may be at greater risk than those who remain at work because their true cumulative exposure to cigarettes is greater than that of individuals remaining at work.

#### E. A worked example

We analyzed lung cancer and all-cause mortality in our cohort of arsenic-exposed copper smelter workers with an eye to investigating the duration of the minimum latent period and the duration of the healthy worker survivor effect. The results are summarized in Table 4. The entries in Table 4 are explained as follows. Columns 1–5 are tests based on modifications of the  $G$ -null test algorithm for a modified MPISTG 3.3 (modified so that four internodal lines arise from the single right circumference point in each node, representing three possible levels of exposure at work—L, M, H—and being off work, l). (Note L in Table 4 refers to low exposure at work and not to employment history.) In implementing the  $G$ -null test algorithm (see Sec. 6C), 50 controls per case were selected. In the analysis of lung cancer mortality, deaths from other causes were treated as censored

Table 4. Analysis of mortality in a cohort of arsenic-exposed copper smelter workers (see Sec. 11E)

Column	Outline	Lag	Modified G-null tests of modified MPISTG 3.3			Type of test	7
			L vs M vs H	L vs (M, H)	l vs (L, M, H)		
Type of comparison							
Lung cancer	0	+ (18)	1.9 (11)	.89 (.82)	.81 (2.3)	1.6 (3.3)	-.003 (.5) .006 (.64)
	9	+ (17)	1.9 (9.8)	.90 (.6)	.80 (2.0)	1.6 (3.5)	-.006 (.2) .02 (1.2)
	15	+ (15)	1.8 (8)	.86 (1.18)	.77 (2.9)	1.5 (2.2)	-.005 (.1) .035 (2.2)
	-5	+ (8)	1.8 (.6)	.462 (3.4)	.390 (4.6)	1.8 (.5)	—
	-9	+ (2.7)	1.9 (2.5)	.849 (.257)	.793 (.481)	1.8 (.7)	—
All causes of death	0	+ (9.06)	1.16 (7.9)	.935 (4.7)	.914 (7.6)	1.06 (.72)	.006 (.0) .002 (1.6)
	9	+ (6.2)	1.13 (4.4)	.964 (1.8)	.939 (2.9)	1.01 (.019)	.003 (.4) .007 (2.2)
	15	+ (5.5)	1.12 (4.2)	.955 (2.6)	.928 (4.1)	1.01 (.018)	.008 (.8) .012 (2.6)
	-5	+ (8.8)	1.6 (9.2)	.865 (2.1)	.834 (3.07)	1.6 (4.1)	—
	-9	+ (5.5)	1.3 (5.0)	.867 (4.9)	.834 (7.4)	1.3 (2.9)	—

Note: Numbers in parentheses are one degree-of-freedom  $\chi^2$  statistics.

Other entries in columns 2-5 are Mantel-Haenszel summary odds ratio with a category to the right of the "vs" in each column serving as the exposed.

Other entries in columns 6 and 7 are estimates of  $\beta_1$  in Eq. (6.1).

observations at the time of death. For the moment the reader should assume that all our prior results concerning *G*-null tests for death from all causes are also valid for such *G*-null tests of lung cancer mortality. Whether this is actually so is the subject of Sec. 12. Time since hire was the time scale on the tree graph and matching was to within 6 months on age at hire and within 3 years on calendar period of hire.  $\Delta t$  was chosen to be one year. The lag variable describes the weight to be given to each table as follows. Lag 0 gives weight 1 to all tables, Lag 9 (15) gives weight 1 to all tables more than 9 (15) years prior to death of the case, and 0 weight to all other tables (e.g.  $w(\underline{i}_k, t_s) = 0$  if  $|t_k - t_s| \leq 9$ ). Symmetrically, lag (-9) [(-5)] gives weight 0 to all tables more than 9 (5) years before the event and weight 1 to all other tables. The columns labelled "Type of Comparison" describe the test statistic computed for each table. In the column labelled L vs M vs H the Mantel-Haenszel test for trend was computed for each table for individuals at work with exposure scored as 0, 1, 2 for L, M, H. (Individuals off work do not contribute to the test.) The table specific trend statistic numerators were combined in an overall one degree-of-freedom Mantel-Haenszel test for trend. The  $\chi^2$  value for this test is given in parentheses with a + sign if there was a positive association of mortality with increasing exposure. Note, because individuals off work did not contribute any information, this test is exactly the *G*-null trend test for MPISTG 3.4. Therefore, the entry (lag 0, column 1) for all causes of death is the square of the Z-score, 3.01, in Table 3. The column labelled L vs (M, H) differs from the previous column only in that individuals who were M or H were scored identically as 1. The numbers in parentheses again represent the overall Mantel-Haenszel  $\chi^2$  statistic. The other number in each row of column 2 represents the summary Mantel-Haenszel odds ratio. The numbers in the succeeding three columns are also  $\chi^2$  statistics and M-H odds ratios. In the column labelled l vs (L, M, H) individuals off work were coded 0 and individuals at work were coded 1 (irrespective of exposure). In the column labelled l vs L, individuals off work were coded as zero, individuals at low exposed jobs were coded 1, and individuals at medium or high exposed jobs did not contribute to the test statistic. In the column labelled l vs (M, H) individuals off work were coded as 0, individuals at either high or medium exposed jobs were coded as 1, and individuals at low exposed jobs did not contribute. The entries in the Cumulative Exposure column are the partial likelihood estimates of  $\beta_1$  from Eq. (6.1). The entries in parentheses are the  $\chi^2$  statistic from the partial likelihood score test of Eq. (6.2). The exposure code gives the weights assigned to each exposure level when cumulating exposure. A lag of 9 years implies that exposures experienced in the 9 years before the death time of the case were ignored in computing the cumulative exposure of the cases and their matched controls. The same matched sets of cases and controls were used in the analysis represented in columns 6 and 7 as in the analysis in columns 1-5.

The most striking aspect of the lung cancer results in Table 4 is that there is no statistically significant association of cumulative exposure with lung cancer mortality, even if one lags exposure 15 years. As discussed in the Introduction, we know that among individuals in this cohort hired prior to 1935, arsenic exposure was a cause of lung cancer. Although clean up occurred in the 1920s, nevertheless, the SMR comparing the lung cancer mortality rate in the entire subcohort hired after 1935 to that of the general U.S. population is 1.4. This suggests a possible persistent effect of arsenic on lung cancer mortality that is not detected by lagging cumulative exposure 15 years when computing the score test for Eq. (6.2). In contrast, as shown in column 1, the *G*-null trend test based on MPISTG 3.4 shows a clear exposure effect ( $\chi^2 = 18$ ). Furthermore, the Mantel-Haenszel odds ratio associated with the *G*-null test for MPISTG 3.4 (when grouping medium and high exposure) is 1.9 (column 2). Can these results be explained by a healthy worker survivor effect whose duration exceeds 15 years for lung cancer? What evidence can we find for such an effect in Table 4? As we now explain, there is some evidence to be found in column 4.

Consider two individuals (technically, we should consider two groups, but speaking of individuals aids the exposition) hired at the same age and in the same calendar year with identical employment and exposure histories until, one day, one of the individuals finds himself out of work, and the other finds himself in a low exposure job. (For the present, we assume we gather no further information on their subsequent employment or exposure history.) Both individuals survive an additional 15 years from the day (which we will call  $D$ -day) on which their histories diverged. According to entry (lag 15, column 4) of Table 4, the subsequent lung cancer mortality rate in the individual at a low exposure job on  $D$ -day is 0.77 that of the individual who had left work (although the associated  $\chi^2$  is only 2.9). Assuming the relative risk of 0.77 is real, possible causal explanations for this finding include (1) the state of being off work is an independent causal risk factor for death from lung cancer; (2) the healthy worker survivor effect is operative, probably because the individual who left work is more likely to be a cigarette smoker; (3) during the 15 years that have elapsed since  $D$ -day, the arsenic exposure of the individual who was off work during  $D$ -day exceeded that of the individual at low exposure (this might occur if, for example, the individual off work on  $D$ -day later returned to a high exposure job); and (4) arsenic is protective against lung cancer, (4) we might rule out *a priori*. In addition, the  $\chi^2$  of 18 for the  $G$ -null test for MPISTG 3.4 would essentially exclude this option if MPISTG 3.4 were an FR MCISTG. If (3) were the explanation, we would have expected to find an association of cumulative exposure with lung cancer mortality in entry (lag 15, column 7). As an even better check, for each table  $(i_k, t_s)$  contributing to the test statistic in (lag 15, column 4) we determined whether the average cumulative exposure in the interval  $(t_k, t_s)$  was greater in subjects who were  $L$  at  $t_k$  than in those who were  $L$  when using the exposure scoring scheme in column 7. On average, it was not greater. In addition, the subjects who actually died do not contribute a greater average cumulative exposure. As always, explanations 1 and 2 cannot be empirically distinguished.

For the present, we shall assume that we are observing a pure selection effect (i.e. Assumption 9.3 holds) and that MPISTG 3.4 is an FR MCISTG. The observed relative risk of 1.5 in (lag 15, column 5) suggests that the risk associated with high or medium arsenic exposure was of greater magnitude than the selection effect associated with leaving work. Nonetheless, the relative risk of 1.8 in entry (lag 15, column 2) suggests the relative risk of 1.5 is almost certainly biased towards the null by the healthy worker survivor effect. Entry (lag -5, column 4) suggests that the magnitude of the healthy worker survivor effect is much greater in the first 5 years after an individual leaves work, presumably because individuals forced to leave work because of symptomatic lung cancer die within a 5 year period. Note that the relative risk of 1.8 in the entry in (lag -5, column 5) suggests the absence of a minimum latent period for lung cancer. But this estimate is so unstable as to be meaningless. (Although the reader might be surprised that the relative risk in columns 2 and 5 of row (lag -5) could be identical given the risk estimate in column 4, this may occur when there are but few controls per case in each table.)

Although we consider the above results to be some evidence for a healthy worker survivor effect lasting for more than fifteen years, we are skeptical whether the selection effect is of sufficient magnitude to explain the marked discrepancy between the  $\chi^2$  values associated with the  $G$ -null test of MPISTG 3.4 and those of the Cox cumulative exposure score tests of Eq. (6.2). Our present research is aimed at resolving this question.

In general, the overall character of the results for death from all causes is quite similar to those for deaths from lung cancer. It is interesting to note that the healthy worker survivor effect for deaths from all causes is apparently less than that for lung cancer. Furthermore, there is statistically convincing evidence in entry (lag -5, column 5) and entry (lag -5, column 2) that the minimum latent period for deaths from all causes is less than 5 years. One would not be surprised by this if accidental death was the main con-

tributor to the excess (since inexperienced workers tend to have accidents) and jobs with high arsenic exposure were also, incidentally, the dangerous jobs. But, interestingly, a similar analysis (not shown) for death from heart disease shows a similar lack of a minimum latent period. If a biological explanation for this finding cannot be found, one would have to carefully check that personnel policies did not systematically place individuals with cardiac risk factors into high-exposure areas.

Before one allows oneself to be persuaded of the causal effect of arsenic on lung cancer mortality in this cohort on the basis of the large  $\chi^2$  value produced by the  $G$ -null test for MPISTG 3.4, there are several other issues we must consider. To review, we argue in Section 11D that the association of lung cancer mortality with exposure lagged some number of years can underestimate the effect of arsenic exposure on mortality if cigarette smoking is an empirical risk factor for leaving employment. In Sec. 11E we apparently confirmed our prediction, provided that we believed MCISTG 3.4 was an FR MCISTG. But if we are willing to accept the hypothesis that cigarette smoking is a determinant of leaving employment, we must ask what is the implication of this hypothesis for our beliefs concerning whether MCISTG 3.4 is an FR MCISTG? And, even if we believe that 3.4 is an FR MCISTG, do our beliefs that cigarette smoking is a determinant of leaving employment, influence our thoughts about the likelihood that a test of the  $G$ -null hypothesis of FR MCISTG 3.4 is a test of the null hypothesis of no effect of exposure controlling for cigarette smoking. (That is, must we consider the possibility that the  $G$ -null test of MPISTG 3.4 rejected due to an effect of exposure on cigarette smoking behavior rather than due to a direct biological effect on the lungs?) And, finally, how do we account for competing causes of death?

To shed some light on this issue, in subsection 12C we shall give conditions under which (1) Gilbert's parameter when adjusting for  $x$ -lagged cigarette smoking history would be the basis of a valid test of the null hypothesis of no exposure effect on lung cancer controlling for cigarette smoking history, (2) Gilbert's parameter without controlling for  $x$ -lagged smoking would lead to a biased test, and (3) the  $G$ -null test for MPISTG 3.4 would be a valid test of the null hypothesis of no exposure effect controlling for cigarette smoking.

#### F. Can treatments at $t$ be determined by events that occurred prior to $t$ ?

We have seen that when  $x$  is less than the generalized minimum latent period of FR OCISTG 11.1, MPISTG 11.2 is the Stage 0 PL-sufficient reduction of the FR OCISTG 11.1. Is MPISTG 11.2 an FR MCISTG as well? Obviously, only if it is an MCISTG, which would require, for example, that the treatment given at  $t_3$  is an individual's employment and exposure status at  $t_1$ . Although one could take the position that it is philosophically inappropriate for treatment at  $t$  to be determined by events that occurred prior to  $t$ , we nonetheless examine whether MPISTG 11.2 satisfies the formal definition of an MCISTG given in Sec. 4. We consider two cases: (1)  $x$  is greater than the generalized minimum latent period of OCISTG 11.1 and (2)  $x$  is less than the generalized minimum latent period. If  $x$  is greater than the generalized minimum latent period, without loss of generality, we can assume there is some individual observed, say, to have received zero exposure at work at  $t_1$  and to have survived to  $t_3$  who, had he received high exposure at  $t_1$ , would have died before  $t_3$ . Using the formalism of Sec. 4C, we consider the value of  $D_{i,j_3}$  for that individual in the standard labelling of MPISTG 11.2, where  $j_3$  is the treatment "high exposure at work at  $t_1$ " given at  $t_3$ .  $D_{i,j_3}$  is undefined, since when in  $i'_3 j_3$  the individual (having died prior to  $t_3$ ) neither dies in the interval  $(t_3, t_4]$  nor survives past  $t_4$ . Thus, MPISTG 11.2 is not an MCISTG. Now suppose that  $x$  is less than the generalized minimum latent period for MCISTG 11.1. Then  $D_{i,j_3}$  will always be well defined and, in fact, for

each individual  $HT(\cdot i_s^{11.1})$  in MCISTG 11.1 naturally induces a unique  $HT(\cdot i_{s+x}^{11.2})$  where  $\cdot i_{s+x}^{11.2}$  is the image of  $\cdot i_s^{11.1}$ . Thus, formally MPISTG 11.2 is an MCISTG. In fact, within the context of our formal definitions, it is easy to show that

**THEOREM 11.7.** If  $x$  is less than the generalized minimum latent period of an FR MCISTG  $A$ , then MCISTG  $A(x)$  is an FR MCISTG.

*Proof.* See Ref. [7].

We can now prove in a simple manner that

**THEOREM.** If Assumptions  $\mathcal{G}.1$ ,  $\mathcal{G}.2$  and  $\mathcal{G}.3$  hold for OPISTG 11.1, then Eq. (11.3b) holds.

*Proof.* Given  $\mathcal{G}.1$  and  $\mathcal{G}.2$ , MPISTG 3.3( $x$ ) is an R MCISTG by the above theorem. But if  $\mathcal{G}.3$  holds, the supposition of Lemma F2 holds. The theorem is proved by now applying Lemma F2.

## 12. OTHER CONSIDERATIONS

### A. Competing risks and censoring

Consider a study represented by an MPISTG with followup from  $t_1$  to  $t_{S+1}$ . Often individuals will be lost to followup (censored) prior to  $t_{S+1}$ . In this section we generalize our theory of causal inference in observational mortality studies to allow for right censoring. (We shall assume that no censored individuals reenter into followup, although extension of our approach to include such interval censoring would not be difficult.)

A closely related (in fact, formally identical) problem is that of competing risks. We believe that an investigator who professes interest in the causal effect of exposure on cause of death  $D_1$  is often assuming (either implicitly or explicitly) that each individual has a potential death time from  $D_1$  that may be unobserved due to either censoring by loss to (or end of) followup or by death from other causes  $D_2$ . This assumption will form the basis of our formalization of the competing risks problem. Various investigators have professed displeasure with this assumption of potential but unobserved death times[24]. We discuss this issue further in Secs. 12B and 12C.

Our formal development can refer either to the pure competing risk problem, pure censoring problem, or the mixed problem. In the pure competing risk problem  $D_1$  refers to time of death from the cause of interest and  $D_2$  to time of death from all other causes. In the pure censoring problem,  $D_1$  refers to time of death from any cause and  $D_2$  to time of censoring. In the mixed problem  $D_1$  refers to time of death from the cause of interest and  $D_2$  to the minimum of the time of censoring and time of death from other causes. For consistency of exposition we have written this section from a single point of view, that of the pure competing risks problem.

We now develop our formal theory. First, we modify the definition of  $HT(\cdot i_s)$  in Sec. 4 by assuming that each individual in  $\cdot i_s$  has deterministic sets  $D_1(t_u)$  and  $D_2(t_u)$  in place of  $D(t_u)$ . Both are defined exactly like  $D(t_u)$ , except  $D_1$  refers to death from the cause of interest and  $D_2$  refers to death from all other causes. It follows from this definition that an investigator will consider an MPISTG to be MCISTG only if he believes that, at least conceptually, deaths from cause  $D_2$  could be eliminated in a manner that does not affect past or future covariate status or time of death from cause  $D_1$ . A modification of the competing risk problem that does not require this assumption is discussed in Sec. 12C.

Whenever we speak informally of the causal effect of exposure on cause of death  $D_1$ , we shall be referring formally to the effect of exposure on death in the world in which death from  $D_2$  has been removed. We denote such a world by  $RD_2$ .

We shall suppose that for any PISTG (even if not a CISTG) each individual has a deterministic vital status and covariate history (that lies on the PISTG) in  $RD_2$ .

*Definition.* Let  $HT_1(\cdot i_s)$  be  $HT(\cdot i_s)$  less the sets  $D_2(t_u)$ .

*Definition.* A PISTG  $A$  is an FR CISTG  $RD_2$  if Eq. (4.6) holds with  $HT_1(\cdot i_s)$  substituted for  $HT(\cdot i_s)$ . We represent the  $G$ -causal parameters of CISTG  $A$  in  $RD_2$  by  $S_{D_1}(t_s, G_1^A, G_2^A, RD_2)$ . Since, in  $RD_2$ ,  $D_1$  is our only cause of death, all of our previous results apply in such a world.

We shall need a number of definitions. In this section only, we shall assume that data on exact death times have been collected.  $\lambda_{D_1}(t \mid \cdot)$  is the instantaneous hazard of death from cause  $D_1$  at time  $t$  given the information in  $\cdot$ . Note that we use this symbol  $\lambda$  (rather than  $\gamma$ ) to represent instantaneous hazards.

*Definition 12A.1.* Given a PISTG, the full-independence assumption holds for  $D_1$  if and only if for all  $\cdot i_s$ ,  $\gamma(\cdot i_s \mid RD_2) = \gamma(\cdot i_s)$  and for all  $\cdot i_s j_s, t_{s+1} > t \geq t_s$ ,  $\lambda_{D_1}(t \mid \cdot i_s j_s, RD_2) = \lambda_{D_1}(t \mid \cdot i_s j_s)$ . The full independence assumption is similar to that usually made in the competing risk problem (i.e. independence of the competing risks conditional on all measured covariates).

*Definition 12A.2.*  $S_{D_1}(t \mid \cdot i_s j_s) = \exp[-\int_{t_s}^t \lambda_{D_1}(u \mid \cdot i_s j_s) du]$  and  $S(t \mid \cdot i_s j_s) = \exp[-\int_{t_s}^t \lambda_D(u \mid \cdot i_s j_s) du]$ .

*Definition 12A.3.*  $p(\cdot i_s, D \geq t \mid "G") = S[t \mid \cdot i_s j_s]p(\cdot i_s \mid "G")$  where  $\cdot i_s j_s \in "G"$ . Note  $p(\cdot i_s, D > t_{s+1} \mid "G") = p[\cdot i_s j_s(t_{s+1}) \mid "G"]$ .

*Definition 12A.4.*  $p^*(\cdot i_s, D > t \mid "G")$  is  $p(\cdot i_s, D \geq t \mid "G")$  with  $S_{D_1}(t_{k+1} \mid \cdot i_k j_k)$  substituted for  $S(\cdot i_k j_k)$  in the definition of  $p(\cdot i_s \mid "G")$  and  $S_{D_1}(t \mid \cdot i_s j_s)$  substituted for  $S(t \mid \cdot i_s j_s)$ . Note  $p(\cdot \mid \cdot, RD_2) = p^*(\cdot \mid \cdot, RD_2)$  since in  $RD_2$ , cause of death  $D_1$  is the only cause.

*Definition 12A.5.* For  $t_s \leq t \leq t_{s+1}$

$$\lambda_{D_1}(t \mid "G") \equiv \frac{\sum_{\cdot i_s j_s \in G} \lambda_{D_1}(t \mid \cdot i_s j_s) p(\cdot i_s, D \geq t \mid "G")}{\sum_{\cdot i_s j_s \in G} p(\cdot i_s, D \geq t \mid "G")}.$$

*Definition 12A.6.*  $\lambda_{D_1}^*(t \mid "G")$  is  $\lambda_{D_1}(t \mid "G")$  with  $p^*$  substituted for  $p$ . Note  $\lambda_{D_1}^*(\cdot \mid \cdot, RD_2) = \lambda_{D_1}(\cdot \mid \cdot, RD_2)$ .

**LEMMA 12A.1.** Full independence for  $D_1$ , implies  $\lambda_{D_1}^*(t \mid "G") = \lambda_{D_1}^*(t \mid "G" RD_2)$ .

*Definition 12A.7.* Given a PISTG, the  $G$ -independence assumption holds for  $D_1$  if and only if  $\lambda_{D_1}(t \mid "G") = \lambda_{D_1}(t \mid "G", RD_2)$ .

[Note if there is but one intranodal line per node (e.g. MPISTG 3.5) full and  $G$ -inde-

pendence are identical.] Note  $G$ -independence would hold if the observed MPISTG represented the PL-sufficient Stage 0 reduction of a double blind ordinary designed randomized trial in which  $\lambda_{D_1}(t | G) = \lambda_{D_1}(t | G, RD_2)$ .

It is of course nonidentifiable whether full independence or  $G$ -independence or neither assumption holds. In Sec. 12C we discuss what prior beliefs an investigator would need to have in order to wish to make the assumption of full or  $G$ -independence. Until Sec. 12B, we assume that full or  $G$ -independence are the only nonidentifiable assumptions that an investigator might entertain about the joint distribution of  $D_1$ ,  $D_2$ , and the covariates. We now examine the consequences of full and  $G$ -independence. We shall use \* to represent results related to full independence and no \* for results related to  $G$ -independence.

*Definition 12A.8.*  $S_{D_1}(t_s | "G_1^A") = \exp[-\int_{t_1}^{t_s} \lambda_{D_1}(u | "G_1^A") du]$ .

*Definition 12A.9.*  $S_{D_1}^*(t_s | "G_1^A")$  is defined as in Eq. (4.7) except any  $S(\cdot i_k j_k)$  is replaced by  $S_{D_1}(t_{k+1} | \cdot i_k j_k)$ . Note  $S_{D_1}^*(t_s | "G_1^A") = \exp[-\int_{t_1}^{t_s} \lambda_{D_1}^*(u | "G_1^A") du]$ .

**THEOREM 12A.1.** If (1) an MPISTG  $A$  is an FR MCISTG  $RD_2$  and (2) the full-independence assumption holds, then  $S_{D_1}^*(t_s | "G_1^A") = S_{D_1}(t_s | G_1^A, RD_2)$ , where the expression on the right refers to the survival curve in  $RD_2$  when the population is treated with  $G_1^A$ . Furthermore, under Large Sample Limiting Model 1,  $S_{D_1}(t | G_1^A, RD_2)$  can be estimated using Eq. (4.7) as previously described, except the NPMLE of any  $S(\cdot i_s j_s)$  is replaced by that of  $S_{D_1}(t_{s+1} | \cdot i_s j_s)$  [which is just the Kaplan-Meier estimator of the (conditional) survival curve for  $D_1$  in the interval  $(t_s, t_{s+1})$ ].

*Proof.* Obvious.

**THEOREM 12A.2.** If (1) an MPISTG  $A$  is an FR MCISTG  $RD_2$  and (2) the  $G$ -independence assumption holds, then  $S_{D_1}(t_s | "G_1^A") = S_{D_1}(t_s | G_1^A, RD_2)$ .

*Definition 12A.11.* The “ $G$ -independence” (“full-independence”) null hypothesis for  $D_1$  holds if  $S_{D_1}(t, "G_1", "G_2") \equiv 0$  ( $S_{D_1}^*(t, "G_1", "G_2") \equiv 0$ ) in notation similar to that we have used before.

*Definition.* Given a MCISTG, the sharp null hypothesis holds for  $D_1$  if the sharp null hypothesis holds in  $RD_2$ .

**LEMMA 12A.2.** If MPISTG  $A$  is an FR MCISTG  $RD_2$  and the sharp null hypothesis holds for cause of death  $D_1$ , then the  $G$ -independence (full-independence) assumption implies that the “ $G$ -independence” (“full-independence”) null hypothesis holds.

*Proof.* Straightforward.

#### B. The circumstance under which the $G$ -null tests for $D_1$ are valid and a worked example

If our interest is in whether an exposure has any effect on mortality from cause of death  $D_1$ , we would like to have available a valid nonparametric test of the “ $G$ - and full-independence” null hypotheses for  $D_1$  analogous to  $G$ -null test? The only natural candidate is the modification of the  $G$ -null test algorithm, in which we treat deaths from cause  $D_1$  as “deaths” and deaths from cause  $D_2$  as censored. We called this procedure the  $G$ -null test for  $D_1$  in Sec. 11E. Unfortunately the circumstances under which it is valid are rather

restrictive. In the following we assume that  $\Delta t \approx 0$  on our PISTG so that the probability of death from either  $D_1$  or  $D_2$  is small in any interval  $\Delta t$ . Then we may return to the use of  $\gamma$  rather than  $\lambda$ . Then, without loss of generality we can and do assume that all deaths from cause  $D_2$  in the interval  $(t_{s-1}, t_s]$  occur instantaneously right at  $t_s$ .

**THEOREM 12A.3.** If

$$\gamma_{D_2}(t_{s+1} | \cdot i_{kj_k}(t_k), \cdot i_{kj_k}i_{k+1}), \quad t_{k+1} \leq t_s \text{ does not depend on } \cdot i_{kj_k}i_{k+1} \quad (12.1)$$

then  $S_{D_1}^*(t, "G_1", G_2") \equiv 0$  implies

$$\gamma_{D_1}(t_{s+1} | \cdot i_k, \cdot i_{kj_k}) \text{ does not depend on } \cdot i_{kj_k}. \quad (12.1a)$$

The  $G$ -null test for  $D_1$  is, in general, not a valid test of the “full-independence” null hypothesis even when Eqs. (12.1) and (12.1a) hold since the contributions to the numerator can be correlated. On the other hand, if a PISTG has but one intranodal line per node the  $G$ -null test for  $D_1$  is a valid test of both the “full and  $G$ -independence null hypothesis.”

**THEOREM 12A.4.** The “ $G$ -independence” null hypothesis holds for both  $D_1$  and  $D_2$  if and only if Eq. (12.1a) holds, both as written and with  $D_2$  substituted for  $D_1$ . In this instance, together the  $G$ -null tests for both  $D_1$  and  $D_2$  constitute a valid test of the joint “ $G$ -independence” null hypothesis for  $D_1$  and  $D_2$  (see Corollary E1).

**THEOREM 12A.5.** If, for  $S \geq 0$ ,

$$\gamma_{D_2}(t_{s-2} | \cdot i_s j_s i_{s+1}) = \gamma_{D_2}(t_{s+2}) \text{ for all } \cdot i_s j_s i_{s+1} \quad (12.2)$$

then Eq. (12.1) holds, the “ $G$ -independence” null hypothesis holds for  $D_2$ , and  $\gamma_{D_1}^*(t_{s-1} | "G") = \gamma_{D_1}(t_{s+1} | "G")$ . Thus, when Eq. (12.2) holds the  $G$ -null test for  $D_1$  is a valid test of the “ $G$ -independence null hypothesis”.

*Proofs.* The trick is to recognize that  $D_2$  is just another covariate, in particular, a covariate that can itself be a treatment, since, if individuals experience  $D_2$ , we are assuming it can be removed. That is to say, the competing risk problem is just a question of inference about the effect of the treatment  $G$  on  $D_1$ , controlling for the covariate  $D_2$ . In Sec. 8D.3 and Appendix G, we solve this problem for cigarette smoking as a covariate. Using MPISTG 3.4 as a paradigmatic example, we shall without loss of generality establish the analogous results in the competing risk case. Assume  $\Delta t$  is small. Let  $D_2, \bar{D}_2$  individuals at each time  $t_s$  be labelled  $c$  and  $\bar{c}$  (where individuals who are  $c$  at  $t_s$  are individuals who died of cause  $D_2$  in the observed study at  $t_s$ ). Since  $D_2$  can be a treatment, if we add  $c$ -status to the set of internodal lines arising from each right circumference point of MPISTG 3.4 (technically of the  $t_1$ -modified version of MPISTG 3.4, see Sec. 8D.2), we have MPISTG 8.3. Now, since we have identified  $D_2$  with  $c$ ,  $D_1$  can play the role in MPISTG 8.3 that  $D$  did when  $c$  was cigarette smoking. Now, from the definitions, it is straightforward to check that with  $c$  as  $D_2$  and  $D$  as  $D_1$  Eq. (8.29) represents the “full-independence null hypothesis”. Equation (12.1) becomes Eq. (8.30). Equation (12.1a) becomes Eq. (8.23), Eq. (12.1a) with  $D_2$  substituted for  $D_1$  becomes Eq. (8.22). Equation (12.2) becomes Equation 8.31. And the “ $G$ -independence null hypothesis” for  $D_2$  and  $D_1$ , respectively, becomes Eqs. (8.19) and (8.20). Then Theorems 12A.3, 12A.4, and 12A.5 become Lemmas 8.15 (equivalently, Theorem G1), Lemma 8.11 (Theorem G4), and Lemmas 8.16 and 8.17 (equivalently, Theorems G2 and G3), respectively.

We now consider the circumstances under which we can empirically test the sharp null hypothesis for  $D_1$  when neither the  $G$ -independence nor the full-independence assumption holds and thus Lemma 12A.2 is not applicable.

**LEMMA 12A.3.** Given MCISTG  $A$  is an FR MCISTG, if the sharp null hypothesis holds for causes of death  $D_1$  and  $D_2$ , then the “ $G$ -independence” null hypothesis holds for  $D_1$  and  $D_2$ .

*Proof.* This result is just a restatement of Lemma 8.8 when MCISTG  $A$  is MCISTG 3.4 and, as above, we have labelled  $D_2$  as  $c$ . Note that the sharp null hypothesis for  $D_1$  becomes Eq. (8.17) and that for  $D_2$  becomes Eq. (8.18). [Recall that, in this section, the set  $HT(\cdot, t_s)$  in the definition of an FR MCISTG contains the two sets  $D_1(t_u)$  and  $D_2(t_u)$ .]

*Definition.* The *adverse effect assumption holds for  $D_1$*  if, under the above labelling of  $D_2$  as  $c$  and  $D_1$  as  $D$ , the adverse effect assumption holds as written in Sec. 8D.3 modified so that following the words “some  $t_s$ ”, we add “such that  $C^i(t_s)$  for individual  $i$  takes on the value  $\bar{c}$  for all times up to and including  $t_s$  (so that individual  $i$  is alive at  $t_s$ ).

*Definition.* The *adverse effect assumption holds for  $D_2$*  if the above modified version of the adverse effect assumption holds upon labelling  $D_1$  as  $c$  and  $D_2$  as  $D$ .

**LEMMA 12A.4.** If the adverse effect assumption holds for  $D_2$ , then if (1) MCISTG 3.4 is a FR MCISTG, (2) the “ $G$ -independence” null hypothesis holds for  $D_2$  and (3) the “ $G$ -independence” null hypothesis is false for  $D_1$ , then the sharp null hypothesis is false for  $D_1$ .

*Proof.* This is exactly Lemma 8.10 with  $D_1$  as  $c$  and  $D_2$  as  $D$ .

*Remark.* Lemma 12A.4 holds with the roles of  $D_2$  and  $D_1$  reversed.

*Remark.* Given MCISTG 3.4 is a FR MCISTG  $RD_2$ , and that we (a priori) make the  $G$ -independence assumption for  $D_1$ , then, in the spirit of Sec. 8D.3, we should accept (reject) the sharp null hypothesis for  $D_1$  when the “ $G$ -independence” null hypothesis for  $D_1$  does (does not) hold. On the other hand, as in Sec. 8D.3, if we assume MCISTG 3.4 is an FR MCISTG and that the adverse effect assumption holds for  $D_2$ , then, when the “ $G$ -independence” null hypothesis holds for  $D_2$ , we would accept (reject) the sharp null hypothesis for  $D_1$  when the “ $G$ -independence” null hypothesis for  $D_1$  holds (does not hold). If the “ $G$ -independence” null hypothesis for  $D_2$  does not hold, then without further assumptions we will neither accept nor reject the sharp null hypothesis for  $D_1$  regardless of the truth of the “ $G$ -independence” null hypothesis for  $D_1$ .

*Remark 12.1.* When (1) MCISTG 3.4 is an FR MCISTG, (2) the sharp null hypothesis for  $D_1$  holds, (3) the adverse effect assumption holds for  $D_2$ , and (4) the sharp null hypothesis for  $D_2$  is false (so that the “ $G$ -independence” null hypothesis for  $D_2$  is false under the adverse effect assumption for  $D_2$  by Lemma 8.9 with  $D_1$  as  $c$  and  $D_2$  as  $D$ ) the minimal additional assumption necessary for the “ $G$ -independence” null hypothesis for  $D_1$  to hold is much less stringent than that of the  $G$ -independence assumption for  $D_1$ . In fact given (1)–(4) above, we only need assume that at each time  $t_s$ , for each  $G$ , the incidence of death from cause  $D_1$  at  $t_s + \Delta t$  for the subset of individuals who would be alive at  $t_s$  when treated with  $G \equiv 0$  (i.e. the treatment “always receive zero exposure”) but not

alive when treated with  $G$  is the same as the incidence of death from cause  $D_1$  for the set of all individuals who would be alive at  $t_s$  when treated with  $G \equiv 0$ . Given (1) and (3), if an investigator is willing to make the above assumption, he will reject the sharp null hypothesis for  $D_1$  whenever the “ $G$ -independence” null hypothesis for  $D_1$  is false.

*Remark 12.2.* An alternative definition of the sharp null hypothesis for  $D_1$  can be constructed without requiring that an individuals' covariate history or death time from  $D_1$  be defined after death from  $D_2$ .

*Definition.* The alternative sharp null hypothesis holds for  $D_1$  if for any  $G_1, G_2, t$  for which individual  $i$  would be at risk at  $t$  when treated with both  $G_1$  and  $G_2$ ,  $\gamma_{D_1}(t | G_1, i) = \gamma_{D_1}(t | G_2, i)$ . The statements of the adverse effect assumption for  $D_1$  and  $D_2$  and of Lemmas 12A.3 and 12A.4 remain unchanged under this alternative definition.

On the other hand, it is much less clear how to define the magnitude of the population effect of a generalized treatment on death from cause  $D_1$  without invoking the existence of  $RD_2$ . We cannot define the effect of treatment in terms of  $\gamma_{D_1}(t | G_1) - \gamma_{D_1}(t | G_2)$  for an MCISTG. For example,  $\gamma_{D_1}(t | G_1) - \gamma_{D_1}(t | G_2)$  may be nonzero even if the alternative sharp null hypothesis holds for  $D_1$ . This reflects the fact that if there exists a non-null effect of treatment on  $D_2$ , the set of individuals at risk at  $t$  under the hypothetical study defined by  $G_1$  will differ from those at risk in the study defined by  $G_2$ . We could, in theory, compute  $\gamma_{D_1}(t | G_1) - \gamma_{D_1}(t | G_2)$  only for individuals who are alive at  $t$  when treated with both  $G_1$  and  $G_2$ , but this leads to nontransitivity of treatment comparisons.

**EXAMPLE:** Suppose, for MPISTG 3.4,  $D_1$  is death from lung cancer, and  $D_2$  is death from all other causes plus censored individuals. Then Eq. (12.1) becomes

$$\gamma_{D_2}(t_{s+1} | E(t_k), L(t_k), l(t_{k+1})) = \gamma_{D_2}(t_{s+1} | E(t_k), L(t_k), \bar{l}(t_{k+1})) \text{ for } t_{k+1} \leq t_s. \quad (12.3)$$

Therefore, under the “full-independence assumption”, the chi-squared value of the  $G$ -null test for lung cancer given in entry (lag 0, column 1) of Table 4 is valid evidence against the null hypothesis of no overall exposure effect on lung cancer mortality provided (1) MCISTG 3.4 is an FR MCISTG  $RD_2$  (which is a nonidentifiable assumption), (2) Eq. (12.3) holds (which is an empirically testable assumption) and (3) the correlations mentioned in Theorem 12A.3 are unimportant. We assume (3) holds (because, as discussed below, Eq. (12.2) is “nearly true”). We have assumed that no censoring occurs in our cohort (see Sec. 5B). Therefore,  $D_2$  represents death from all causes other than lung cancer. Deaths from lung cancer represent only 116 out of 1782 deaths. Thus, as expected, the results for death from all causes (as shown in Table 4) are nearly identical to those for death from all causes but lung cancer (data not shown). Therefore, we can simply allow  $D_2$  to be deaths from all causes without introducing substantial bias. Entry (lag 0, column 3) in Table 4 for all causes of death is then a valid test of Eq. (12.3). Equation (12.3) is rejected ( $\chi^2 = 4.7$ ), although the magnitude of this effect (OR = .935) would, we believe, make it improbable that the  $\chi^2$  of 18 in entry (lag 0, column 1) for lung cancer could be due to this bias.

We now consider the entry (lag 9, column 3) for all causes of death. This entry represents a valid test of Eq. (12.3) when  $t_k$  is restricted so that  $t_{k+1} + 9 \leq t_s$ . Let us call this restricted version of Eq. (12.3), Eq. (12.3-lag 9). Note that the odds ratio (.964) in this entry is greater than its counterpart under lag 0 (OR = .935). Furthermore, the test of Eq. (12.3-lag 9) does not reject ( $\chi^2 = 1.8$ ). Why should this be the case? The obvious reason is that, by lagging nine years, the  $l$  individuals who left work because they were imminently going to die no longer contribute to the test statistic.

But if Eq. (12.3-lag 9) is true (at least, the modified  $G$ -null test does not reject), then the  $\chi^2$  value of 15 in entry (lag 9, column 1) for lung cancer constitutes valid evidence against the null hypothesis of no overall exposure effect on  $D_1$  (i.e., the  $G$ -null hypothesis in  $RD_2$  for MCISTG 3.4) under the full-independence assumption provided MPISTG 3.4 is an FR MCISTG  $RD_2$ . In fact, it is valid if MPISTG 3.4( $x$ ) with  $x = 9$  is an FR MCISTG  $RD_2$ , provided for individuals who are already  $x$ -doomed in  $RD_2$ , future exposure and employment history do not influence their death times. (This is a weaker assumption than the assumption that MPISTG 3.4 is an FR MCISTG  $RD_2$  since, for MPISTG 3.4( $x$ ) to be an FR MCISTG  $RD_2$  under the  $G$ -null hypothesis in  $RD_2$  for MCISTG 3.4, it is sufficient that for the subset of individuals who were not  $x$ -doomed at  $t_s$ , exposure at work, given past employment and exposure history, is at random in  $RD_2$ . This follows from Theorem 11.7, upon recognizing that, under the null hypothesis, for all  $x$ ,  $x$  is less than the generalized minimum latent period.

If we had believed *a priori* that MCISTG 3.4 were an FR MCISTG and that either the  $G$ -independence assumption held for  $D_1$  or (only) the adverse effect assumption held for  $D_2$  (and we considered low exposure to be essentially equivalent to zero exposure so that  $S(t \mid G^{3,4} = 0)$  is identifiable), then we would reject the sharp null hypothesis for lung cancer based on the extreme  $\chi^2$ -value of 18 for the  $G$ -null test for lung cancer when Eq. (12.2) holds. To see this note that by Theorem 12A.5, Eq. (12.2) implies the “ $G$ -independence” null hypothesis holds for  $D_2$ . Therefore, given the  $G$ -null test for  $D_1$  rejects, Theorem 12A.4 implies the “ $G$ -independence” null hypothesis is false for  $D_1$ . Hence under either the  $G$ -independence assumption for  $D_1$  or the adverse effect assumption for  $D_2$  (by using Lemma 12A.4), the sharp null hypothesis for  $D_1$  must be false. Equation (12.2) can be shown to imply that the “ $G$ -independence” null hypothesis for all causes of death (but lung cancer) holds for MPISTG 3.3. This in turn implies that to a good approximation the “ $G$ ”-null hypothesis for MPISTG 3.3 holds for all causes of death. Thus, all the tests represented in columns 1–5 of Table 4 would be null. But there is evidence in both column 1 and column 4 against the  $G$ -null hypothesis of MPISTG 3.3, MPISTG 3.3(9), MPISTG 3.3(15) and thus against Eqs. (12.2), (12.2-lag 9), and (12.2-lag 15). Why, in contrast, did we have the good fortune (when considering the full-independence assumption) to find the test in column 3 for lags 9 and 15 did not reject the null? The reason is that for the comparisons made in column 3, the increased mortality rate observed in individuals off work due to the healthy worker survivor effect (shown in column 4— $\chi^2 = 2.9$ ,  $\chi^2 = 4.1$ ) is partly balanced by the adverse effect of arsenic on the mortality of the healthier individuals who remained at work (shown in column 1— $\chi^2 = 6.2$ ,  $\chi^2 = 5.5$ ).

Even though Eq. (12.2) is likely false, we conjecture that Eq. (12.2) is close enough to being true (in terms of the magnitude of the non-null associations) that given the extreme  $\chi^2$ -value and large Mantel-Haenszel odds ratio associated with the  $G$ -null tests of  $D_1$ , the sharp null hypothesis for  $D_1$  can still be rejected. An active research issue is to turn this conjecture into a sharp theorem.

In Sec. 12C, we consider (1) whether it is reasonable to assume that MCISTG 3.4 is an FR MCISTG or even an FR MCISTG  $RD_2$ , (2) which assumptions among the full-independence,  $G$ -independence and adverse effect assumptions are *a priori* reasonable, and (3) whether we are successfully testing for the effect of exposure on lung cancer controlling for cigarette smoking and employment history.

### C. Testing for effects of exposure on lung cancer controlling for cigarette smoking in the presence and absence of data on cigarette smoking

It would not be unreasonable to assume that, among individuals free from clinical lung cancer, exposure, employment, and cigarette smoking status were received at random

with respect to unmeasured risk factors for lung cancer conditional on past exposure, employment and smoking history. It would also be a reasonable approximation to assume that all individuals with newly diagnosed clinical lung cancer will die within  $x$  years (e.g. five years). Then, OCISTG 11.1 modified so that  $x = 5\Delta t$  and the right circumference points labelled by  $\bar{l}s$  were the source of six internodal lines representing joint levels of exposure, employment and smoking status [i.e.  $(0, \bar{l}, \bar{c})$ ,  $(0, \bar{l}, c)$ ,  $(H, \bar{l}, \bar{c})$ ,  $(H, \bar{l}, c)$ ,  $(0, l, c)$ ,  $(0, l, \bar{c})$ ] would be an FR OCISTG  $RD_2$  (i.e. in a world in which the only cause of death was lung cancer). [Note that with  $x = 5$  years one would, in general, not consider the above modified version of OCISTG 11.1 to be an FR OCISTG if we considered causes of death other than lung cancer. For instance, individuals who leave work at  $t_s$  are more likely to be suffering from (unrecorded) angina pectoris than individuals with the same past history who stayed at work. Since, while alive, individuals with angina pectoris both continue to be at elevated risk of death from heart disease, and are not  $x$ -doomed for  $x = 5$  years (or even 15 years), employment status is not received at random (with respect to unmeasured risk factors for all causes of mortality) among non- $x$ -doomed individuals. Furthermore, even among those who left work at  $t_s$ , those who left due to angina pectoris may be more likely to give up cigarette smoking at that time than those who left work for purely socioeconomic reasons. As such, even cigarette smoking would not be received at random at  $t_s$  (conditional on  $C(t_{s-1}), E(t_{s-1}), L(t_s)$ ) among non- $x$ -doomed individuals.]

**LEMMA.** In a world in which lung cancer is the only cause of death, if the above modified version of OCISTG 11.1 is an FR OCISTG  $RD_2$  (i.e. the healthy worker survivor effect lasts for less than  $x$  years) and the minimum latent period for exposure and cigarette smoking (and employment history) is greater than  $x$  years, i.e. for each individual  $i$

$$\gamma_{D_i}(t + \Delta t | E(t), C(t), L(t), i) = \gamma_{D_i}(t + \Delta t | E(t - x), C(t - x), L(t - x), i)$$

then, F8.3( $x$ ) is an FR MCISTG  $RD_2$ .

*Proof.* Similar to Theorems 11.2 and 11.7.

If F8.3( $x$ ) is an FR MCISTG  $RD_2$  and if there is no exposure effect on lung cancer controlling for cigarette smoking and employment history (when defined in terms of the  $G$ -causal parameters of our modified FR OCISTG 11.1) then in  $RD_2$

$$\gamma_{D_i}(t + \Delta t | E(t - x), L(t - x), C(t - x)) = \gamma_{D_i}(t + \Delta t | L(t - x), C(t - x)). \quad (12.4)$$

Furthermore, it may be reasonable to assume that  $L$  is not a causal risk factor controlling for cigarette and exposure history, since it would be unlikely that, for example, loss of health insurance and increased poverty would have a causal effect on lung cancer mortality when controlling for cigarette smoking. Under this further assumption, the right-hand side of Eq. (12.4) reduces to

$$\gamma_{D_i}(t + \Delta t | C(t - x)). \quad (12.5)$$

Epidemiologic evidence suggests that the minimum latent period for cigarette smoking may be short, since the mortality from lung cancer among individuals who have given up smoking five years previously is observed to be less than the rate among continuing smokers (when controlling for smoking history up to five years before the time at risk). Thus, five years is probably the maximum value one would wish to use for the minimum latent period in the previous Lemma.

We might find it reasonable to assume that the full-independence (equivalently the  $G$ -

independence) assumption holds for MPISTG F8.3(x). This assumption would follow if one believed that, conditional on  $L$ -,  $C$ -, and  $E$ -history, unmeasured risk factors for lung cancer death were distributed independently of those for death from other causes. Under this assumption an empirical test of Eq. (12.4) (in this world) serves as a test of the null hypothesis of no exposure effect controlling for  $L$ - and  $C$ -history in  $RD_2$ .

*Remark.* Because of the assumed lack of correlation between unmeasured risk factors for lung cancer mortality and those for all other causes of death, the left side of Eq. (12.4) would equal

$$\gamma_{D_1}(t_s + \Delta t | G = [E(t_s), L(t_s), C(t_s)]) \text{ (where } G \text{ is a generalized treatment of the modified OCISTG 11.1)}$$

both in this world (as well as in  $RD_2$ ) even though the modified OCISTG 11.1 is not an FR MCISTG in this world.

Now suppose data on cigarette smoking is missing. We now consider whether MPISTG 3.4 is an MCISTG. If so,  $\gamma_{D_1}(t_s | G_1^{3,4}, RD_2)$  would be well defined for  $G_1^{3,4}$  "if at work, receive high exposure" even when there is an independent exposure effect [i.e. Eq. (12.4), is false in  $RD_2$ ]. Does this make sense? Consider an individual who, when treated with  $G_1^{3,4}$  in this world, leaves work at age 48 with disabling angina and dies of coronary artery disease at age 55. In  $RD_2$ , he survives past 55. But does he go back to work? His time of death from lung cancer will, in general, depend on the answer. Possibly he returns to work at age 56 if his angina remits? But does his angina remit? It is clear why Kalbfleish and Prentice may have argued against believing in  $RD_2$ .

[This discussion is further complicated by the fact that one might want to eliminate the assumption (contained in the definition of a CISTG) that covariate history prior to time of death from  $D_2$  would not be influenced when cause of death  $D_2$  is removed. For example, one might want to assume that if death from heart disease were removed, disabling angina pectoris should also be removed. Formally we could accomplish this by unlinking the definition of a CISTG A  $RD_2$  from that of CISTG A. If so, our individual who, in this world left work at age 48 with disabling angina, would not have left work in  $RD_2$ , thus accumulating even larger exposures.]

Nonetheless, from a practical point of view, the above example is not that problematic. We are not able to identifiably estimate survival as a function of a generalized treatment in  $RD_2$  unless the  $G$ - or full-independence assumption holds. So it makes little difference whether we consider outcomes in  $RD_2$  as ill-defined or as (arbitrarily) well-defined but not identifiable (for example, by arbitrarily assuming that all individuals who would have died of heart disease at age  $t$  become symptom free at  $t + \Delta t$  in  $RD_2$ ). We made essentially the same point at the end of Sec. 3E when we decided to arbitrarily define MPISTG 3.3 to be an MCISTG.

In contrast with the problems of estimation, we shall now see that testing of the null hypothesis of no exposure effect on  $D_1$  controlling for cigarette smoking and employment history (in the absence of data on smoking) can be, relatively speaking, less problematic, provided that the null hypothesis of no independent effect of exposure on cause of death  $D_2$  holds. To begin, it seems quite reasonable to believe both that exposure is given at random conditional on past cigarette smoking, exposure, and work history, and that Eq. (8.15) holds (in this world). Now, if Eqs. (8.35) and (8.34) hold both for cause of death  $D_1$  and for cause  $D_2$  [i.e. when  $D_1$  ( $D_2$ ) is substituted for  $D$  in Eqs. (8.34) and (8.35)], then, at times, it may be reasonable to assume that Eqs. (8.24) and (8.26) hold as well. (See discussion at the end of Sec. 8D.3.) If so, then MPISTG 3.4 is an FR MCISTG, and for, MCISTG 3.4, the sharp (and alternative sharp) null hypotheses for  $D_1$  and  $D_2$  will

hold. This essentially follows from Lemma 8.24 and Corollary 8.22. But these sharp null hypotheses imply the “ $G$ -independence” null hypothesis for both  $D_1$  and  $D_2$  (by Lemma 12A.3). Thus, if the  $G$ -null test for  $D_1$  and  $D_2$  both accept (see Theorem 12A.4), we might well accept that exposure has no effect on death from either  $D_1$  or  $D_2$  when controlling for cigarette and employment history.

Nonetheless, even in this setting, the full- and  $G$ -independence assumptions will not in general hold, especially if, as discussed at the end of Sec. 8D.3, an individual’s health status influences cigarette smoking behavior (since, then  $\gamma_{D_1}(t_s | G_1^{3\cdot4}, RD_2)$  will again depend on whether our previous subject’s angina remits after the would be date of cardiac death.) Even if illness status does not determine cigarette smoking behavior the full- and  $G$ -independence assumptions cannot hold for MPISTG 3.4 [or 3.4(x)], even under the null, if they hold for MPISTG F8.3(x) and cigarette smoking is a cause of both lung cancer and death from other causes.

Suppose that if, for MPISTG 3.4, the “ $G$ -independence” null hypothesis held for  $D_1$ , then, for MCISTG 3.4, we would accept the sharp null hypothesis holds for  $D_1$ . (If one were willing to make the assumptions concerning MCISTG 3.4 described in Remark 12.1, one might be willing to use this decision rule). Then if, for MPISTG 3.4, the “ $G$ -independence” null hypothesis for  $D_2$  failed to hold and that for  $D_1$  held, one would accept the sharp null hypothesis of MCISTG 3.4 for  $D_1$ , but this would not count as evidence for the null hypothesis of no effect of exposure on  $D_1$  controlling for cigarette smoking and employment history. This follows because the sharp null hypothesis of MCISTG 3.4 for  $D_2$  must be false. A likely explanation for which is that either Eq. (8.34) or (8.35) is false for  $D_2$ . But, as discussed at the end of Sec. 8D.3, if Eq. (8.34) or (8.35) is false (even if only for  $D_2$ ) then, in general, either Eq. (8.24) or (8.26) will be false. If Eq. (8.24) or (8.26) is false, then, even when there is no exposure effect on  $D_1$  controlling for cigarette smoking and employment history, the sharp null hypothesis of MCISTG 3.4 for  $D_1$  will, in general, not hold. Thus, even when we accept that discover the sharp null hypothesis for  $D_1$  does hold, we will not be inclined to conclude that exposure has no effect on  $D_1$  controlling for  $L$ - and  $C$ -history.

#### D. Extension of our results to other studies

It is common for risk factors for death to be determinants of future exposure in a variety of mortality studies outside of occupational epidemiology. For example, in studying the effect of exogenous estrogens on mortality, an investigator needs to be aware that physicians frequently withdraw women from exogenous estrogens at the time they develop hypertension or angina.

Similarly, in observational studies of the efficacy of cervical cancer screening on mortality, women who have had operative removal of their cervix due to malignancy are no longer at risk for further screening (i.e. exposure). This example was pointed out to the author by Alan Morrison. For example, suppose MPISTG 3.4 represented the results of an observational mortality study of screening for cervical cancer where women screened at  $t_s$  are  $H$ , women not screened at  $t_s$  are  $0$ ,  $|t_s - t_{s-1}| = 1$  day, individuals with (without) a cervix at  $t_s$  are  $\bar{l}$  ( $l$ ), respectively, and we suppose among women with a cervix that, conditional on past screening history, screening status at  $t_s$  is received at random. Then, MPISTG 3.4 is an FR MCISTG. (Note women who are  $l$  at  $t_s$  have probability 0 of becoming  $\bar{l}$  at later times.)

In this setting, Sasco *et al.* (25) recently suggested one could analyze the results of such a study from case-control data by using design and analysis strategy 3 as defined in Sec. 9C. But, as pointed out in Sec. 9C, if exposure (screening) history is a determinant of  $l$ -status (which it almost certainly is, even under the null hypothesis), then design and

analysis strategy 3 does not produce valid tests. Rather, if, as is possible,  $L$ -status is a determinant of mortality (since individuals with invasive cancer are among those who have their cervix removed), then only the  $G$ -null test for MCISTG 3.4 is a valid test of the null hypothesis of no overall effect of screening history on mortality. Furthermore, based on the results given in Sec. 9A, we are unable to test, using case-control data, the null hypothesis of no direct effect of screening on mortality, controlling for cigarette smoking, if women, sobered by the experience of needing to have their cervixes removed, change their cigarette smoking habits. (In this setting, as in Fig. 9.1a and 9.1b, MCISTG 8.3 will be an FR MCISTG, smoking is an independent risk factor for death,  $L$  is an independent predictor of future smoking history and of death, and exposure is an independent predictor of future  $L$ -status.)

The extension of our results to outcomes other than mortality is, at least in a theoretical sense, straightforward. For example, if an investigator believed that MCISTG 3.4 was an FR MCISTG, the population distribution of employment history  $L(t_{s+1})$  can be estimated for the hypothetical studies defined by the generalized treatments of FR MCISTG 3.4. The interpretation is complicated by the fact that death censors  $L(t_{s+1})$ . One might be interested in the effect on the covariate's history in a world in which death has been removed. In that setting the “ $G$ -independence” and “full independence” assumptions and null hypotheses for death from  $D_1$  can be generalized to arbitrary time-dependent covariates.

To be precise: suppose that whenever we speak informally of the causal effect of exposure on a time dependent covariate  $I_s^B$ , as defined in Appendix G, [e.g. on cigarette smoking history,  $C(t_s)$ , where paths of possible  $C(t_s)$  are the various  $\cdot i_s^B$ ] we are referring to the effect of exposure on  $I_s^B$  in a world without death. We denote such a world by  $RD$ . We have the following definitions and results. (These results are numbered to correspond to their isomorphic counterparts in Secs. 12A and B. Since we have changed the ordering of the definitions slightly the numbers are slightly out of order.)

*Definition 12D.1.* Given a PISTG  $A$ , the full-independence assumption holds for  $I_s^B$  if and only if for all  $\cdot i_s^A$ ,  $\gamma(\cdot i_s^A | RD) = \gamma(\cdot i_s^A)$ .

*Definition 12D.4.*  $p^\Delta(\cdot i_s^A | "G_1^A") \equiv p^\Delta(\cdot i_s j_s(t_{s+1})^A | "G_1^A")$  is the product of intranodal probabilities  $\gamma(\cdot i_k)$  [ending with  $(\cdot i_s)$ ] on the sequence of inter- and intranodal lines connecting the left circumference of the  $t_1$  node to node  $\cdot i_s j_s(t_{s+1})$ .

*Definition 12D.5.*  $\gamma(\cdot i_{s+1}^B | "G_1^A") \equiv p[\cdot i_{s+1}^B | "G_1^A", \cdot i_s^B]$ . The term on the right is defined in Definition G.1.

*Definition 12D.6.*  $\gamma^\Delta(\cdot i_{s+1}^B | "G_1^A")$  is defined exactly like  $\gamma(\cdot i_{s+1}^B | "G_1^A")$  except  $p^\Delta(\cdot i_s j_s(t_{s+1})^A | "G_1^A")$  is substituted for  $p(\cdot i_s j_s(t_{s+1})^A | "G_1^A")$ . Note  $\gamma(\cdot i_{s+1}^B | "G_1^A", RD) = \gamma^\Delta(\cdot i_{s+1}^B | "G_1^A", RD)$ .

*Definition 12D.8.*  $p_{I_s^B}[\cdot i_s^B | "G_1^A"] = p(i_1^B) \prod_{k=2}^s \gamma(\cdot i_k^B | "G_1^A")$ .

*Definition 12D.9.*  $p_{I_s^B}^\Delta[\cdot i_s^B | "G_1^A"]$  is similarly defined except  $\gamma^\Delta$  replaces  $\gamma$ . Note:  $p_{I_s^B}^\Delta[\cdot i_s^B | "G_1^A"] = p^\Delta[\cdot i_s^B | "G_1^A"]$ , where  $p^\Delta[\cdot i_s^B | "G_1^A"]$  is defined like  $p[\cdot i_s^B | "G_1^A"]$  (which is defined under Definition G1) except  $p^\Delta[\cdot i_s^A | "G_1^A"]$  replaces  $p[\cdot i_s^A | "G_1^A"]$ .

**LEMMA 12D.1.** Full independence for  $I_s^B$  implies  $\gamma^\Delta(\cdot i_{s+1}^B | "G_1^A") = \gamma(\cdot i_{s+1}^B | "G_1^A", RD)$ .

*Definition 12D.7.* The  $G$ -independence assumption holds for  $I_s^B$  if and only if  $\gamma(\cdot i_{s+1}^B | "G_1^A") = \gamma(\cdot i_{s+1}^B | "G_1^A", RD)$ . Note if within each node all intranodal lines have a unique value of  $\cdot i_s^B$ , then  $\gamma(\cdot i_{s+1}^B | "G_1^A") = \gamma^A(\cdot i_{s+1}^B | "G_1^A")$  (e.g. this would be the case for MPISTG 3.4 with  $L(t_s)$  determining levels of  $\cdot i_s^B$ ).

**THEOREM 12D.1.** If MPISTG A is an FR MCISTG RD and the full-independence assumption holds, then  $\gamma^A(\cdot i_{s+1}^B | "G_1^A") = \gamma(\cdot i_{s+1}^B | G_1^A, RD)$  where the expression on the right refers to probability statements about outcomes in RD when the population is treated with  $G_1^A$ .

**THEOREM 12D.2.** If the  $G$ -independence assumption holds for FR MCISTG RD A,  $\gamma(\cdot i_{s+1}^B | "G_1^A") = \gamma(\cdot i_{s+1}^B | G_1^A, RD)$ .

*Definition 12D.11.* The “ $G$ -independence” (“full-independence”) null hypothesis holds for  $I_s^B$  if the  $\gamma[\cdot i_{s+1}^B | "G_1^A"]$  ( $\gamma^A[\cdot i_{s+1}^B | "G_1^A"]$ ) are the same for all  $G_1^A \in G^A$ .

**THEOREM 12D.3.** If Eq. (12.1) holds with  $D$  substituted for  $D_2$  and the “full-independence” null hypothesis for  $I_s^B$  holds then for  $t_k \leq t_s$

$$p[\cdot i_{s+1}^B | \cdot i_s^B, \cdot i_k, \cdot i_k j_k, D > t_{s+1}] \text{ does not depend on } \cdot i_k j_k. \quad (12.6)$$

If each intranodal line corresponds to a different level of  $\cdot i_s^B$ , both the “full- and  $G$ -independence” null hypothesis for  $I_s^B$  imply Eq. (12.6) holds.

**THEOREM 12D.4.** If Eq. (G5) holds, then the “ $G$ -independence” null hypothesis for  $I_s^B$  implies Eqs. (12.6) and (G3A) hold.

**THEOREM 12D.5.** If Eq. (12.2) holds with  $D$  substituted for  $D_2$ , then  $\gamma(\cdot i_{s+1}^B | "G_1^A") = \gamma^A(\cdot i_{s+1}^B | "G_1^A")$ . A test for Eq. (12.6) can be constructed in a manner exactly like that described in Sec. 9 for testing Eq. (8.22). It can be shown simply by changing the rotation in Corollary E1 that the contributions to the test statistic numerator from tables  $[\cdot i_k, t_{s+1}]$  are uncorrelated when Eqs. (12.6) and (G3A) both hold.

*Acknowledgements*—Research for this paper was funded in part by grants from the following organizations: National Institute of Environmental Health Sciences #5 R23 ES03405, American Lung Association #LA 3/22/85, National Institute for Occupational Safety and Health #5 RO3 OH01867, National Institute of Environmental Health Sciences #5 P30 ES00002, The Alfred P. Sloan Foundation and Mobil Corporation.

The author wishes to thank Mary Firestone, Janna Frelich, Virgi Hepner, Mike Wold, and Donald Blevins for technical assistance and Lee Robins for reviewing the manuscript.

## APPENDIX A

In this Appendix we review some concepts that will be used in the paper. Consider a statistical model with minimal sufficient statistics  $(T, Q)$  and parameters  $P = (P_1, P_2)$ , where  $T, Q, P_1, P_2$  are possibly vector valued with the following property.

$$f(T, Q; P_1, P_2) = f(T | Q; P_1)f(Q; P_2), \quad (A1)$$

where  $f$  either is a density function or a probability function, and any value of  $P_2$  can arise with any value of  $P_1$ . The following are trivial to show:

- (1) In a Bayesian framework, if  $P_1, P_2$  have independent priors, then the posterior distribution of  $P_2$  depends on the data only through  $Q$ .
- (2) The maximum likelihood estimator of  $P_2$  and the observed information matrix or expected information evaluated at the maximum likelihood estimate depend on the data only through  $Q$ .

(3) Pure likelihood inference based on maximized relative likelihoods for  $P_2$  depend on the data only through  $Q$ .

Furthermore, by definition,  $Q$  is called a cut, and  $Q$  is said to be  $S$ -sufficient for  $P_2$ , and  $S$ -ancillary for  $P_1$  [26]. By extended principles of conditionality and sufficiency, frequentist inference on  $P_2$  is from the marginal distribution of  $Q$ , and on  $P_1$  is from the conditional distribution of  $T$  given  $Q$ .

More generally suppose,

$$f(T, Q; P_1, P_2) \propto L_1(P_1; Q, T)L_2(P_2; Q). \quad (\text{A2})$$

If any value of  $P_1$  and  $P_2$  can appear together we continue to call  $Q$  a cut, contrary to standard usage. If Eq. (A2) holds, then (1)–(3) above still hold. If the functions  $L_1$  and  $L_2$  in Eq. (A2) cannot be interpreted as the conditional distribution of  $T$  given  $Q$  and the marginal distribution of  $Q$ , respectively,  $Q$  is not  $S$ -sufficient. (Note that if  $P_1$  is infinite dimensional and  $P_2$  is finite dimensional, the standard large sample properties of the maximum likelihood estimator of  $P_2$  will be valid under regularity conditions if  $L_2(P_2; Q)$  is a partial likelihood).

**EXAMPLE.** Suppose  $X, Y, Z, V$  are random variables with  $X \sim \text{Bin}(N, \theta_1)$ ,  $N$  a known constant;  $Y | X \sim \text{Bin}(N - X, \theta_2)$ ;  $Z | Y, X \sim \text{Bin}(N - X - Y, \theta_3)$ ;  $V | X, Y, Z \sim \text{Bin}(Y, \theta_4)$ . Let  $Q = (X, Y, Z)$ ,  $T = V$ ,  $P_2 = (\theta_1, \theta_3)$ , and  $P_1 = (\theta_2, \theta_4)$ . Then Eq. (A2) holds with

$$\begin{aligned} L_1(P_1; Q, T) &= \theta_2^Y(1 - \theta_2)^{N-X-Y}\theta_4^V(1 - \theta_4)^{Y-V} \\ L_2(P_2; Q) &= \theta_1^X\theta_3^{N-X}\theta_3^Z\theta_3^{(N-X-Y-Z)}. \end{aligned}$$

Equation (A1) is false and  $L_2$  is a partial likelihood.

This example could represent a study of survival time in which  $\theta_1$  was the probability of survival from start of follow-up to some time  $t_2$ . At  $t_2$  survivors were independently and instantaneously censored with probability  $\theta_2$ .  $\theta_3$  represents the probability of survival from time  $t_2$  to end of follow-up in uncensored individuals.  $\theta_4$  might represent the probability that a censored individual is a male. Thus,  $Y$  would represent the censored individuals and  $V$ , the male censored individuals, and some way or another the value of  $V$  is made available. It follows that inference on the probability of survival up to end of follow-up, i.e.  $\theta_1\theta_3$ , would, for a likelihood-based frequentist or a Bayesian with independent priors, depend only on the data through  $Q$ .

When Eq. (A2) holds, we will say that  $Q$  is  $L$ -sufficient for  $P_2$ . When  $L_2$  is a partial likelihood, we will say that  $Q$  is  $PL$ -sufficient for  $P_2$ .

Consider that Eq. (A2) held but  $P_1 - \frac{1}{2}P_2 = 0$  was known *a priori*. Then the priors for  $P_1$  and  $P_2$  are no longer independent, and thus the posterior distribution will depend on the data through  $T$ ; the restricted maximum likelihood estimator of  $P_2$  given  $P_1 - \frac{1}{2}P_2 = 0$  and maximized relative likelihoods for  $P_2$  will depend on the data through  $T$ ; every value of  $P_2$  cannot arise with every value of  $P_1$ , so by definition  $Q$  cannot be a cut for  $P_2$ . A prior restriction that functionally depends on both  $P_1$  and  $P_2$  will be called a cross over since it crosses over the “cut” between  $P_1$  and  $P_2$ . On the other hand, the restrictions such as  $P_1 = 5$  or  $P_1 = 6$  do not cross over, and properties 1–3 above continue to hold.

## APPENDIX B

**LEMMA B1.** If  $p(G | i) = p(G)$  and Eqs. (2.1) and (2.3) hold then  $\gamma_L(t | E(t), G) = \gamma_L(t | E(t))$ .

*Proof.*  $\gamma_L(t | E(t), G) = \sum_i \gamma_L(t | E(t), i, G)p(i | E(t), L > t, D > t, G)$ . By Eq. (2.1) the first factor in the sum does not depend on  $G$ . The lemma follows if we can show

$$p(i | E(t), L > t, D > t, G_1) = p(i | E(t), L > t, D > t). \quad (\text{B1})$$

Let  $I[G_1(t)]$  be the subset of the population such that  $L > t, D > t$  when assigned any  $G$  such that  $G(t) = G_1(t)$ . By (2.1) and (2.3),  $I[G_1(t)]$  is well defined since by (2.1) and (2.3)  $I[G_1(t)]$  cannot depend on  $g(u)$  for  $u > t$ . Now Eq. (B1) may be restated as  $p(i | G_1) = p(i)$ ,  $i \in I[G_1(t)] \cap \{i; G_i(t)$

$= G_1(t)\}$  (where  $\{i; G_i(t) = G_1(t)\}$  is the set of individuals assigned a  $G$  such that  $G(t) = G_1(t)$ ), which follows immediately from the fact that  $p(i | G) = p(i)$ . The proofs that the nonidentifiable temporal assumptions plus randomization imply the other two identifiable temporal assumptions are similar.

**THEOREM B1.** If the identifiable temporal assumptions hold *a priori*, then  $\{(L_i, E(L_i), D_i)\}$  are sufficient in the ordinary sense (i.e.  $B$ -sufficient) for  $\{(p(D > t | G); \text{for all } G \text{ on test})\}$ .

*Proof.* For notational convenience in this proof, we without loss of generality define  $L_i = D_i$  if an individual  $i$  dies on treatment protocol.

$$p(L, D, E(L), G) = p(G | L, E(L), D)p(L, E(L), D).$$

If the identifiable temporal assumption holds, then the first factor in the above likelihood factorization equals  $p(G | E(L))$ , which depends only on the investigator's known choice for the distribution of projected exposure paths under study, i.e. the randomization scheme. Thus  $\{(L_i, E(L_i), D_i)\}$ ,  $i$  indexing the individuals studied is sufficient in the ordinary sense (i.e.  $B$ -sufficient) for  $p(D > t | G)$ .

## APPENDIX C

*Proof of Theorem 4.1.* The theorem follows by induction from the following three lemmas.

(For notational convenience, we use  $p(\cdot | \cdot i_s, \cdot)$  to mean  $p(\cdot | \cdot i_s, D > t_s, B, \cdot)$  where  $B$  is the CISTG of the theorem.)

**LEMMA 1.** If Eq. (4.5) holds and if, for all  $G$ ,

$$p(H(\cdot i_s) | \cdot i_s, G) = p(H(\cdot i_s) | \cdot i_s) \quad (\text{C1})$$

then  $p(H(\cdot i_s) | \cdot i_s j_s, \cdot i_s) = p[H(\cdot i_s) | \cdot i_s, \cdot i_s j_s, G]$ ,  $S(\cdot i_s j_s | G) = S(\cdot i_s j_s)$ , and  $\gamma(\cdot i_{s+1} | G) = \gamma(\cdot i_{s+1})$ .

*Proof.*

$$S(\cdot i_s j_s) = \sum_{H(\cdot i_s)} p(D > t_{s+1} | \cdot i_s j_s, H(\cdot i_s))p(H(\cdot i_s) | \cdot i_s j_s, \cdot i_s).$$

But

$$p(H(\cdot i_s) | \cdot i_s j_s, \cdot i_s) = p[H(\cdot i_s) | \cdot i_s] = p[H(\cdot i_s) | \cdot i_s, G] = p[H(\cdot i_s) | \cdot i_s, \cdot i_s j_s, G]. \quad (\text{C2})$$

The first equality follows from the definition of an R ISTG [i.e. Eq. (4.5)], the second from superposition (C1) and the third from the fact that, in the unobserved study defined by  $G$ ,  $\cdot i_s$  determines  $\cdot i_s j_s$ . Also

$$p[D > t_{s+1} | \cdot i_s j_s, H(\cdot i_s)] = p[D > t_{s+1} | \cdot i_s j_s, H(\cdot i_s), G] \quad (\text{C3})$$

since  $H(\cdot i_s)$  includes information on survival to  $t_{s+1}$  given  $\cdot i_s j_s$ . Therefore,  $S(\cdot i_s j_s) = S(\cdot i_s j_s | G)$ . Now

$$\gamma(\cdot i_{s+1}) = \sum_{H(\cdot i_s)} p(\cdot i_{s+1} | \cdot i_s j_s, D > t_{s+1}, H(\cdot i_s))p[H(\cdot i_s) | \cdot i_s j_s, D > t_{s+1}, \cdot i_s].$$

By definition of  $H(\cdot i_s)$ , the first term is 1 or 0 and is unchanged in the study defined by  $G$ . By Bayes theorem

$$p[H(\cdot i_s) | \cdot i_s j_s, D > t_{s+1}, \cdot i_s] = \frac{p(D > t_{s+1} | H(\cdot i_s), \cdot i_s j_s, \cdot i_s)p(H(\cdot i_s)) | \cdot i_s j_s, \cdot i_s}{\sum_{H(\cdot i_s)} p(D > t_{s+1} | H(\cdot i_s), \cdot i_s j_s, \cdot i_s)p[H(\cdot i_s) | \cdot i_s j_s, \cdot i_s]}. \quad (\text{C4})$$

But by (C2) and (C3), no term on the right side of Eq. (C4) is changed in a study defined by  $G$ . Thus,  $\gamma(\cdot i_{s+1}) = \gamma(\cdot i_{s-1} | G)$ .

**LEMMA 2.** If  $p(H(\cdot i_s) | \cdot i_s) = p(H(\cdot i_s) | \cdot i_s, G)$  and Eq. (4.5) holds, then  $p(H(\cdot i_{s-1}) | \cdot i_{s+1}) = p[H(\cdot i_{s-1}) | \cdot i_{s+1}, G]$ .

*Proof.*  $p[H(\cdot i_{s+1}) | \cdot i_{s+1}] = \sum_{H(i_s)} p[H(\cdot i_{s+1}) | H(i_s), \cdot i_{s+1}] p[H(\cdot i_s) | \cdot i_{s+1}]$ . But, since  $H(\cdot i_s)$  determines  $H(\cdot i_{s+1}) \equiv H(\cdot i_s j_s i_{s+1})$ , the first term in the sum is 1 or 0 and is identical in a study defined by  $G$ .

$$\begin{aligned} p[H(\cdot i_s) | \cdot i_{s+1}] &\equiv p(H(\cdot i_s) | \cdot i_s, \cdot i_s j_s, D > t_{s+1}, \cdot i_s j_s i_{s+1}) \\ &= \frac{p[D > t_{s+1}, \cdot i_s j_s i_{s+1} | \cdot i_s j_s, H(i_s)] p[H(\cdot i_s) | \cdot i_s j_s, \cdot i_s]}{p[\cdot i_s j_s i_{s+1} | D > t_{s+1}, \cdot i_s j_s] p[D > t_{s+1} | \cdot i_s j_s]} = \frac{AB}{CD} \end{aligned}$$

where  $A$ ,  $B$ ,  $C$ , and  $D$  correspond to the four terms in the previous expression and we have used Bayes' theorem. By definition of  $H(\cdot i_s)$ ,  $A$  is unchanged in a study defined by  $G$ . By Lemma 1,  $B$ ,  $C$ , and  $D$  are also unchanged in such a study. This proves the lemma.

**LEMMA 3.**  $p(H(\cdot i_1) | i_1) = p[H(\cdot i_1) | \cdot i_1, G]$ .

*Proof.* Obvious, since the subset  $\cdot i_1$  is the same subset in the observed study as in a study in which all individuals are given  $G$ .

## APPENDIX D

In this appendix we discuss the results of analyzing the copper smelter cohort data by the methods described in Sec. 5B under the model specifications

$$\begin{aligned} \beta_D \cdot X_D &= \beta_{1,D}[ce(t) \cdot [1 - l(t)]] + \beta_{2,D}l(t) \\ &\quad + \beta_{3,D}[Cl(t) \cdot l(t)] + \beta_{4,D}[ce(t) \cdot l(t)], \end{aligned} \quad (\text{D1})$$

$$\beta_L \cdot X_L = \beta_{1,L}ce(t) + \beta_{2,L}[I_R \cdot CR(t)] + \beta_{3,L}[ce(t) \cdot (t - t_h)], \quad (\text{D2})$$

$$\beta_R \cdot X_R = \beta_{1,R}ce(t) + \beta_{2,R}[ce(t) \cdot (t - t_h)] + \beta_{3,R}[Cl(t)], \quad (\text{D3})$$

where

- (1) letting  $e(u)$  be a quantitative estimate of exposure concentration received at time  $u$ ,  $ce(t)$  is the sum of the biannual measurements  $e(u)$ , taken on an individual up to age  $t$ . It is a discretized version of lifetime cumulative exposure.
- (2)  $t_h$  is age at hire.
- (3)  $l(t) = 1$  if an individual is out of work at  $t$  and  $l(t)$  is zero otherwise.
- (4)  $Cl(t)$  is the number of years since last at work.
- (5)  $I_R$  is an indicator variable taking the value 1 if an individual has ever been off work since time of hire and 0 otherwise.
- (6)  $CR(t)$  is the number of years elapsed since an individual was last off work.

The above models are simplistic and are used for illustrative purposes. They are not the basis of a definitive analysis. They do have a number of desirable features:

- (1) For the outcome of death, the variable  $t$  is used to represent age as mortality rates are most strongly related to age [13].
- (2) The incidence of death should be allowed to depend on time since hire, since at hire workers were selected into the workforce on the basis of health. Stratification on five year intervals of age at hire performs this function.
- (3) The relative risk associated with cumulative exposure (for mortality) may differ for workers at work at  $t$  and workers out of work at  $t$  (i.e.  $\beta_{1,D} \neq \beta_{4,D}$ ). In addition, the effect on mortality of being out of work at  $t$  (i.e.  $\beta_{2,D}$ ) may be modified by the length of time one has been out of work (i.e.  $\beta_{3,D}$ ).
- (4) The conditional probability of leaving work at  $t$  may depend on the time since an individual

Table 5. Maximum partial likelihood estimates of coefficients of Eqs. (5.1)–(5.3) with covariate specification given in Eqs. (D1)–(D3)

	Death ( $D$ )	Failure type Leaving ( $L$ )	Return ( $R$ )
$\beta_1$	-.003 (.0015)†	-.049 (.011)	.023 (.025)
$\beta_2$	.16 (.06)	-.008 (.013)	-.0001 (.0070)
$\beta_3$	-.008 (.003)	.0016 (.0004)	-.367 (.031)
$\beta_4$	.002 (.0013)	—	—

† Standard errors in parenthesis.

was last out of work (i.e.  $\beta_{2,L}$ ). In addition, the effect of cumulative exposure on leaving work may be influenced by the number of years since hire (i.e.  $\beta_{3,L}$ ).

- (5) The conditional probability of returning to work might depend on the number of years since an individual was last at work (i.e.  $\beta_{3,R}$ ). In addition, the effect of exposure on the conditional probability of returning to work may be modified by time since hire.

*Results.* In the data available to us, exposure at work was measured as high, medium, or low (as opposed to high or zero as in Fig. 3.4) and coded as 3, 2, 1, respectively, when computing  $ce(t)$ . Thus, MCISTG 3.4 needs to be modified to reflect the actual study data by having three internodal lines (representing the three exposure levels) originating from each right circumference point that represents individuals at work. Table 5 gives parameter estimates and standard errors (computed from the inverse of the estimated expected information matrix) based on fitting models (5.1)–(5.3) by the method of conditional logistic regression.

For individuals who in the observed trial had  $Z(t_1)$  defined by being born in 1902, beginning work in 1935, and staying at work continually at high-exposure jobs until start of follow-up in 1938, Table 6 gives Monte Carlo estimates (and, in parentheses, Monte Carlo standard errors) of the survival curves up to age 35, 55, and 75 that would have been observed in two different hypothetical studies. The first of these studies, labelled  $G_H$ , is that defined by the generalized treatment of the modified MCISTG 3.4 “if at work (after 1938), receive high exposure”. The second study, labelled  $G_0$ , is a hypothetical study in which each individual in the subset received “zero exposure when at work (after 1938)”. Such a study is technically not represented by a generalized treatment of the modified MCISTG 3.4, because no unexposed jobs existed in the factory. Nonetheless, we assumed that models (5.1)–(5.3) would give valid estimates of  $S[\cdot i_s | G_0, [Z(t_1)]]$  and  $\gamma(\cdot i_s | G_0, [Z(t_1)])$  if they do so in the hypothetical studies defined by the generalized treatments of the modified FR MCISTG 3.4 (i.e. we assumed, we could use the models to extrapolate beyond the range of the data). The Monte Carlo standard errors are based on a  $N$  of 200 Monte Carlo trials. We did not carry out any bootstrap replications. Therefore, the true standard error of the estimates of, e.g.  $S(t | "G_H^{3,4}"', Z(t_1))$  is not known. Thus, although there is a suggestion of an adverse effect of “high exposure if at work” on the probability of survival to age 75, we cannot tell if it is statistically significant. We return to the question of its statistical significance in Sec. 6, at which time we

Table 6. Comparison of the estimated survival probabilities in hypothetical studies of a cohort of arsenic-exposed copper smelter workers

Study	Age		
	35	55	75
$G_H^{\dagger}$	.992§ (.0005)¶	.873 (.013)	.325 (.005)
$G_0^{\ddagger}$	.992 (.0005)	.874 (.004)	.370 (.002)

† Generalized treatment of “if at work past start of follow-up, receive high exposure” for the subset of the observed study population hired at age 32 in 1935 remaining on work at high exposure until start of follow-up in 1938.

‡ Generalized treatment of “if at work past start of follow-up, receive zero exposure” for the same subset of the population.

§ Estimated survival probability.

¶ Empirical Monte Carlo standard error.

compare our results with those obtained using the standard approach of assessing the relationship of mortality with cumulative exposure.

Since  $\beta_{2,D}$  and  $\beta_{3,D}$  were significantly different from zero, it appears that employment history is a population risk factor controlling for exposure history (as defined in Sec. 3G). As discussed in Sec. 3G, that raises the question of whether the empirical healthy worker survivor effect is operative. We have performed a partial but not exhaustive check of the empirical healthy worker survivor effect as follows. We used to estimate the Monte Carlo version of  $G$ -computation algorithm, for a group of workers who remained at work at high exposure through age 56 [with the  $Z(t_1)$  used above], the probability of survival to 75 for the subgroup who left work at age 57 compared to the subgroup who remained at work at 57 when both subgroups were treated, beginning at age 57, with the generalized treatment "if at work receive zero exposure." Survival probability for the "off work at 57 subgroup" was .26 and for the "at work at 57 subgroup" .31. Similar survival differences were obtained comparing groups at work and off work at age 47 and at age 52 (in each case when treated with the generalized treatment "if at work receive zero exposure" beginning at the corresponding age). Thus, it appears that "the empirical healthy worker survivor effect is operative."

We give an informal discussion of why the results in Table 5 and Table 6 might be as observed in order to give the reader a sense of the issues involved. Our discussion is overly schematic and simplistic. It treats the models as correct and their estimates as precise. It is a form of Monday morning quarterbacking, since we are making up explanations to fit our results. Other quite different explanations may be equally plausible. Nonetheless, we do think it is valuable to give a sense of how one might think about the plausibility of the results and the causal mechanisms that might have produced them. First we show how the estimates in Table 5 might lead to the survival curves in Table 6. Then we discuss possible sociobiologic explanations for why the results in Table 5 were observed. First we summarize the main results in Table 5. In Table 5, the group of unexposed individuals off work at  $t$  are, in general, at greater risk for death than the group of individuals at work at  $t$  ( $\beta_{2,D} = 0.16$ ). This excess risk wears off with increasing time off work. Unexposed individuals off work for more than 20 consecutive years have a lower risk than unexposed individuals at work (i.e.  $\beta_{2,D} + 20\beta_{3,D} = 0$ ). In addition, increasing exposure increases the risk of death in the subgroup of individuals off work but not the subgroup at work ( $\beta_{1,D} < 0$ ,  $\beta_{4,D} > 0$ ). Exposure protects individuals from leaving work within the first 30 years from time of hire. After 30 years, increasing exposure increases the rate of leaving work (i.e.  $\beta_{1,L} + 30\beta_{3,L} \approx 0$ ). Exposure has little effect on the incidence of returning to work.

The estimates in Table 5 might predict the adverse effect of high exposure, as seen in Table 6, due to the conjunction of the following facts. (Caveat: Here we are predicting the results of the Monte Carlo estimates of the  $G$ -causal parameters based only on estimates of the coefficients without considering the nuisance hazards. This is fraught with danger but we give it a try.)

It is known that most deaths occur at age 60 and above (i.e. greater than 28 years from hire for individuals treated with  $G_H$  or  $G_0$ ). Highly exposed individuals treated with  $G_H$  have a tendency to remain at work, possibly until age 60. They then start leaving work at a rate faster than unexposed individuals treated with  $G_0$ . Therefore, at age 60 there are a number of highly exposed individuals recently off work with high cumulative exposures. They will experience high mortality rates due to the fact that  $\beta_{2,D}$  and  $\beta_{4,D}$  are positive. In contrast, workers treated with  $G_0$  have a tendency to leave work soon after start of follow-up in 1938 and, if they do not return to work, the excess relative risk associated with leaving will be dissipated by the time they reach the age (i.e. 60) at which the baseline risk of death is finally high. These remarks could explain the difference in survival curves that is observed.

The estimates observed in Table 5 could reflect some or all of the following sociobiologic factors.

(1) Most individuals leaving work soon after time of hire leave not for health related reasons (but rather for economic and social reasons).

(2) In this factory, it might have been the case that for social and economic reasons, the turnover rate in the first few years after hire is greater in low than high-exposure jobs.

(3) Individuals who leave employment at later ages (e.g. age 50–60) are often individuals who have developed chronic diseases.

(4) Exposure is a cause of chronic disabling disease. This would explain why the leaving rate finally increases with increasing exposure after 25–30 years from hire.

(5) 55–60 year old individuals who left work due to disability have a high mortality rate compared

to workers who are still remaining at work (i.e.  $\hat{\beta}_{2,D} = 0.16$ ). (Note that since most deaths occur after age 55 it will be these deaths that determine the estimate of  $\beta_{2,D}$ . In particular, if the model is misspecified and  $\beta_{2,D}$  varies with age, our estimate of the relative risk associated with recently leaving work will be biased for the young individuals.)

(6) Among individuals who left work near age 60, by (4) above, exposure is a marker of chronic disability. Thus, increasing exposure is associated with a higher relative risk (i.e.  $\hat{\beta}_{4,D} = 0.002$ ). Exposure may also be an irritant that affects unhealthy individuals. Therefore, unhealthy individuals, whether unhealthy from the effects of exposure or not, tend to leave highly exposed jobs in their later years. Thus, individuals who accumulate high exposure and remain at work must be healthy individuals and thus  $\hat{\beta}_{1,D} < 0$ .

(7) Individuals who left work soon after start of follow-up tend to be healthy years later (i.e.  $\hat{\beta}_{2,D} + 20\hat{\beta}_{3,D} = 0$ ) for two separate reasons. (1) The group of individuals leaving work at any age is a mixture of healthy and disabled workers. Disabled workers die off over time, leaving healthy workers. (2) Most individuals who left long ago in their twenties and thirties did so for socioeconomic reasons and not due to chronic disabling illness. Thus, their risk will not be high years later. (From other occupational health studies there is some evidence that individuals who leave work soon after hire may be individuals who have difficulty holding a job due to problems with drinking, etc. As such, they are often observed to have higher mortality rates in their later years. We see no evidence of this in our analysis, although we do find such evidence in Sec. 11E when we reanalyze this data using another analytic approach.)

## APPENDIX E

*Proof of Theorem 6.1.*  $\Leftarrow$ . Since  $A$  is an R CISTG it follows from Corollary 4.1 that we need to show that for any PISTG, if for all  $\cdot i_s j_s, \cdot i_s j'_s$ ,  $S(t | \cdot i_s j_s) = S(t | \cdot i_s j'_s)$ , then  $S(t | "G_1^A") = S(t | "G_2^A")$  for all " $G_1^A$ ", " $G_2^A$ ". The proof follows by induction. We assume the theorem is true if for a PISTG of  $S$  generations and show that it is true for  $S + 1$  generations. Suppose  $A$  has  $S + 1$  generations. Now by supposition  $p(D > t_k | \cdot i_1 j_1)$  and  $S(\cdot i_1 j_1)$  can be written as  $p(D > t_k | \cdot i_1 -)$  and  $S(\cdot i_1 -)$ . Thus

$$p[D > t_k | \cdot i_1 j_1(t_2)] = \frac{p(D > t_k | \cdot i_1 -)}{S(\cdot i_1 j_1)} = p(D > t_k | (\cdot i_1 - (t_2))).$$

Furthermore,

$$p[D > t_k | "G^{A[\cdot i_1 j_1(t_2)]}"] = p[D > t_k | A[\cdot i_1 j_1(t_2)]] \quad (E1)$$

since we have assumed the theorem holds for PISTGs of  $S$  generations and we have used Lemma 4.2. But

$$p[D > t_k | A[\cdot i_1 j_1(t_2)]] = p[D > t_k | \cdot i_1 j_1(t_2)] = p[D > t_k | (\cdot i_1 - (t_2))], \quad (E2)$$

where the first equality is definitional and the second follows from above. Now, from its definition, for  $k \geq 3$ ,

$$p(D > t_k | "G^A") = \sum_{i_1=1}^N p[D > t_k | "G^{A[\cdot i_1 j_1(t_2)]}"] S(i_1 j_1) p(i_1), \quad (E3)$$

where the  $j_1(i_1)$  are determined by " $G^A$ " and " $G^{A[\cdot i_1 j_1(t_2)]}$ " is the highlighted subgraph of graph  $A[\cdot i_1 j_1(t_2)]$  induced by " $G^A$ ". But by (E1), (E2) and our supposition, no term on the right of the equality sign in Eq. (E3) depends on  $j_1(i_1)$  and thus on " $G^A$ ". The theorem now follows by checking that it is true for  $t_2$ , and seeing that it is true for a PISTG of two generations.

$\Rightarrow$ . The proof in this direction follows by noting that  $p(D > t | \cdot i_s j_s) - p(D > t | \cdot i_s j'_s)$  can be written as  $[S(t | "G_1^B") - S(t | "G_2^B")]/p(\cdot i_s | "G^A")$  for a PISTG  $B$  coarser than  $A$ . The proof follows by applying Lemma 4.3.

*Preliminaries to the statement to Theorem E1.* Consider a cohort mortality study represented by an MPISTG  $A$  with  $t_1, \dots, t_m$  being the ordered death times. Let  $D(\cdot i_s, t_m)$  be the number of individuals who died at  $t_m$  who are in  $\cdot i_s$  and let  $JD(\cdot i_s, t_m)$  be the  $D(\cdot i_s, t_m)$ -dimensional vector of the treatments  $j_s(\cdot i_s)$  they received at  $t_s$ . Similarly, let  $K(\cdot i_s, t_m)$  be the set of individuals at risk to die at  $t_m$  who are in  $\cdot i_s$ , and  $JK(\cdot i_s, t_m)$  be the vector of treatments they received at  $t_s$ . Let  $F(\cdot i_s, t)$  be the event that, for each individual in  $\cdot i_s$ , records the treatment  $j_s(\cdot i_s)$  received at  $t_s$  and subsequent failure and censoring history through  $t$ . Let  $F(\cdot i_s, t^-)$  be  $F(\cdot i_s, t)$  except that it records failure and censoring history only to a time  $t^-$  just preceding  $t$ . [In our discrete set-up this would be to  $t - \Delta t$ .] Let  $F_{\cdot i_s}(\cdot i'_s, t)$  contains the information in  $F(\cdot i'_s, t)$  only for those individuals who were also in  $\cdot i_s$ . Let  $O(\cdot i_s, t_m)$  and  $E(\cdot i_s, t_m)$  be respectively functions of  $JD(\cdot i_s, t_m)$  and  $JK(\cdot i_s, t_m)$  if  $D(\cdot i_s, t_m) > 0$ . Otherwise, both  $O(\cdot i_s, t_m)$  and  $E(\cdot i_s, t_m)$  are defined to be zero. The functions  $O(\cdot i_s, t_m)$  and  $E(\cdot i_s, t_m)$  may be many valued but they must have the same dimension. Define  $\Delta(\cdot i_s, t_m) \equiv O(\cdot i_s, t_m) - E(\cdot i_s, t_m)$ .

**THEOREM E1.** Suppose for all  $\cdot i_s, \cdot i_s j_s, \cdot i_s j'_s$ ,

$$\gamma_D(t \mid \cdot i_s j_s) = \gamma_D(t \mid \cdot i_s j'_s), \quad t > t_s \quad (\text{E4})$$

and  $E[\Delta(\cdot i_s, t_m) \mid F(\cdot i_s, t_m^-), D(\cdot i_s, t_m)] = 0$  and there is no censoring before end of follow-up. Then

$$\text{Cov}[\Delta(\cdot i_s, t_m), \Delta(\cdot i'_s, t_{m'})] = 0 \quad \text{if } \cdot i_s \neq \cdot i'_s \text{ or } t_m \neq t_{m'}.$$

*Proof.* The theorem would follow upon taking unconditional expectations if we can show that for all  $t'_s \leq t_s$ ,

$$E[\Delta(\cdot i'_s, t_{m'}) \Delta(\cdot i_s, t_m) \mid F(\cdot i_s, t_m^-), D(\cdot i_s, t_m), F(\cdot i'_s, t_{m'})] = 0. \quad (\text{E5})$$

Since  $\Delta(\cdot i'_s, t_{m'})$  is fixed given the conditioning event, and since the event  $[F(\cdot i'_s, t_{m'}) - F_{\cdot i_s}(\cdot i'_s, t_{m'})]$  does not contain information on individuals in  $\cdot i_s$ , it follows from the assumption of independent random sampling that we need only to show that  $E[\Delta(\cdot i_s, t_m) \mid F(\cdot i_s, t_m^-), D(\cdot i_s, t_m), F_{\cdot i_s}(\cdot i'_s, t_{m'})] = 0$ . This will follow if  $JD(\cdot i_s, t_m) \perp\!\!\!\perp F_{\cdot i_s}(\cdot i'_s, t_{m'}) \mid F(\cdot i_s, t_m^-), D(\cdot i_s, t_m)$  where the notation  $A \perp\!\!\!\perp B \mid C$  means  $A$  is conditionally independent of  $B$  given  $C$  as in Dawid [18]. We need to consider three cases.

*Case 1.* If  $t_{m'} < t_m$  the  $F_{\cdot i_s}(\cdot i'_s, t_{m'})$  is fixed given  $F(\cdot i_s, t_m^-)$  and (E5) follows immediately.

*Case 2.* If  $t_{m'} = t_m$  and  $t_s = t_{s'}$ , then since  $\cdot i'_s \neq \cdot i_s$ ,  $F_{\cdot i_s}(\cdot i'_s, t_{m'})$  is the empty set, and (E5) follows.

*Case 3.* The only interesting case is if  $t_{m'} \geq t_m$  when  $t'_s < t_s$ . Equation (E5) then follows by the superposition of equal survival curves without censoring.

By essentially the same argument, we have

**Corollary E1.** Theorem E1 holds in the presence of right censoring if Eq. (E4) holds for the “censoring” hazard as well as for the death hazard.

## APPENDIX F

**THEOREM F1.** If MPISTG  $B$  is a Stage 0 reduction of MPISTG  $A$  and Assumptions R defined below holds, then  $B$  is the Stage 0 PL-sufficient reduction of  $A$  under Assumptions R.

**Assumptions R.** For each  $\cdot i_s j_s^B$  and for all  $\cdot i_s j_s^A$  such that  $[\cdot i_s j_s^A] \subset [\cdot i_s j_s^B]$ ,

$$S(\cdot i_s j_s^A) = S(\cdot i_s j_s^B) \quad (\text{F1})$$

$$p(\cdot i_{s+1}^B \mid \cdot i_s j_s^B(t_{s+1})) = p(\cdot i_{s+1}^B \mid \cdot i_s j_s^B(t_{s+1}), \cdot i_s j_s^A). \quad (\text{F2})$$

*Proof.* See Ref. [7].

*Corollary F1.* Under the suppositions of the above theorem if  $G_1^A$  has a Stage 0 counterpart  $G_1^B$  then  $S(t_s | "G_1^A") = S(t_s | "G_1^B")$  and  $p(\cdot i_s^B | "G_1^A") = p(\cdot i_s^B | "G_1^B")$ .

*Proof.* See Ref. [7].

Our next major result, Theorem F2, requires some introductory definitions.

*Definitions.* If a PISTG  $B$  is a Stage 0 reduction of a PISTG  $A$  and for any  $\cdot i_s j_s^B$ , given  $[\cdot i_s^A]$  contained in  $[\cdot i_s^B]$  there exists an  $[\cdot i_s j_s^A]$  contained in  $[\cdot i_s j_s^B]$ ; we say that  $B$  is a semimelded reduction of PISTG  $A$ . If the  $\cdot i_s j_s^A$  is unique, we say that  $B$  is a melded reduction of  $A$ .

**LEMMA.** If  $B$  is a semimelded reduction of PISTG  $A$ , then  $B$  is a  $B$ -complete Stage 0 reduction of  $A$ .

*Proof.* Given a  $G^B$  for each  $\cdot i_s j_s^B$  on the highlighted subgraph  $G^B$  find the complete set of  $\cdot i_s^A$  that is contained in  $\cdot i_s^B$ . Choose from each such  $\cdot i_s^A$  a single  $\cdot i_s j_s^A$  contained in  $\cdot i_s j_s^B$ . This is possible by the definition of semimelded. The set of such  $\cdot i_s j_s^A$  are seen to determine a  $G^A$  with counterpart  $G^B$ .

**LEMMA.** If  $B$  is a melded Stage 0 reduction of PISTG  $A$ , then  $B$  is a unique  $B$ -complete Stage 0 reduction of  $A$ .

*Proof.* Obvious.

**EXAMPLES.** PISTG 3.5 is a melded reduction of PISTG 8.1. PISTG 3.5 is a semimelded reduction of PISTG 8.3.

*Remark.* The converses of the last two lemmas are false.

**EXAMPLE.** PISTG 7.1b is a unique  $B$ -complete reduction of PISTG 7.1a, but it is neither a semimelded nor a melded reduction.

*Remark.* If PISTG  $B$  is the melded Stage 0 reduction of PISTG  $A$ , then PISTG  $B'$  coarser than  $B$  is the melded Stage 0 reduction of a unique PISTG  $A'$  coarser than  $A$ .  $A'$  is constructed as follows: if, in forming  $B'$  from  $B$ , the set  $\cdot i_s j_s^B$  arising from any  $\cdot i_s^B$  is divided into  $K$  mutually exclusive subsets (each with its own right circumference point); the  $\cdot i_s j_s^A$  arising from each  $\cdot i_s^A$  contained in  $\cdot i_s^B$  are divided into  $K$  mutually exclusive sets in such a way that the  $\cdot i_s j_s^A$  remain contained in the original  $\cdot i_s j_s^B$ . We call  $A'$  the antimeld of  $B'$ .

**THEOREM F2.** Suppose PISTG  $B$  is a melded reduction of PISTG  $A$ , and  $G_1^B$  is a Stage 0 counterpart of  $G_1^A$ . If for all  $\cdot i_s^A$  and  $\cdot i_s'^A$  on highlighted subgraph  $G_1^A$  such that  $[\cdot i_s^A]$  and  $[\cdot i_s'^A]$  are contained in the same  $[\cdot i_s^B]$ ,  $\gamma(\cdot i_s j_s^A) = \gamma(\cdot i_s' j_s'^A)$  (where  $\cdot i_s j_s^A$  and  $\cdot i_s' j_s'^A$  are on subgraph  $G_1^A$ ), then  $S(t_s | "G_1^A") = S(t_s | "G_1^B")$  and  $p(\cdot i_s^B | "G_1^A") = p(\cdot i_s^B | "G_1^B")$ .

*Proof.* See Ref. [7].

*Remark.* The above theorem is false, when PISTG  $B$  is a unique  $B$ -complete Stage 0 reduction of PISTG  $A$ , but is not a melded reduction. PISTG 7.1A and its unique  $B$ -complete Stage 0 reduction PISTG 7.1B, serve as an example of this. The theorem is also false if  $B$  is only a semimelded reduction of PISTG  $A$ . The reader can verify that PISTG 8.3 and its semimelded reduction PISTG 3.5 provide an example.

If PISTG  $A$  is an FR CISTG in Theorems F1 and F2, under what circumstances will the Stage 0 reduction  $B$  also be a FR CISTG? We develop sufficient conditions. We begin with an example to motivate the necessary definitions.

Given that PISTG 8.1 is a CISTG, is its melded Stage 0 counterpart PISTG 3.5 also a CISTG? By definition, a necessary condition would be that if an investigator had intervened and gave to

an individual who received zero exposure at  $t_2$  in the observed study, high exposure at  $t_2$ , the individual's survival history through  $t_3$  must be well defined. Now, since the investigator, in giving a treatment of MCISTG 3.5, does not choose employment at  $t_2$ , nature must deterministically do so when  $l$  is a causal risk factor. Otherwise, survival through  $t_3$  would not be well defined. Although there is not sufficient formal structure in our definition of a CISTG to require nature to "choose," we shall assume that nature does indeed choose so that MPISTG 3.5 is a CISTG. Furthermore, it is natural to suppose (there is, of course, no observational verification possible) that when the investigator intervenes to change exposure, nature leaves employment status unchanged. In that case, the generalized treatment of  $G^{3.5}$ , "always receive high exposure," is exactly equivalent to the generalized treatment of  $G^{8.1}$  defined by "always receive high exposure" in terms of both exposure and employment status. Thus, we should have

$$S(t | G^{3.5}, i) = S(t | G^{8.1}, i). \quad (\text{F3})$$

In fact, Eq. F3 would hold for any  $G^{3.5}$  and its Stage 0 counterpart  $G^{8.1}$ . In addition, an equivalent relationship would hold between any generalized treatment of a coarser MCISTG derived from MCISTG 3.5 and that generalized treatment of its antimeld which is its counterpart. Nonetheless, it is important to remember that if nature did not leave employment status unchanged, then, even though two generalized treatments of MCISTG 3.5 and 8.1 had the same name, "always receive high exposure", they would be different treatments if  $l$  were a causal risk factor. Furthermore, Eq. F3 would be false. We now provide the following generalizations.

*Definition.* If  $B$  is a melded Stage 0 reduction of CISTG  $A$ , we say  $B$  is a *causal* melded Stage 0 counterpart of CISTG  $A$  if (1)  $B$  is a CISTG; (2)

$$p(D > t_{s+1} | \cdot i_s j_s^A, G^B, i) = p(D > t_{s+1} | \cdot i_s j_s^A, G^A, i) \quad (\text{F4})$$

$$p(\cdot i_{s+1}^A | \cdot i_s j_s^A(t_{s+1}), G^B, i) = p(\cdot i_{s+1}^A | \cdot i_s j_s^A(t_{s+1}), G^A, i) \quad (\text{F5})$$

for  $G^B$  and  $G^A$  that are Stage 0 counterparts; and (3) (F4), (F5) hold with  $G^{B'}$  and  $G^{A'}$  (also  $B'$  and  $A'$ ) replacing  $G^B$  and  $G^A$  (also  $B$  and  $A$ ), respectively, where  $B'$  is coarser than  $B$  and  $A'$  is the antimeld of  $B'$ . Note (F4) and (F5) generalize (F3).

*Remark.* To give further clarification of the necessity for the definition of a causal melded reduction, suppose MPISTG F8.1 is an MCISTG. Since our formal definition of a CISTG assigns no meaning to names of treatments, we relabel the four treatments (representing joint levels of exposure and employment status) at any right circumference point as  $a, b, c, d$ . Consider two coarser MCISTGs  $A_1$  and  $A_2$  such that  $A_1$  ( $A_2$ ) has treatments  $a$  and  $c$  ( $a$  and  $d$ ) arising from one right circumference point and  $b$  and  $d$  ( $b$  and  $c$ ) from another. There exists a melded reduction, say,  $A_R$ , of both  $A_1$  and  $A_2$  which has two treatments—one characterized by " $a$  or  $b$ " and the other by " $d$  or  $c$ "—arising from a single right circumference point.  $A_R$  is not a CISTG without further assumptions, since it could be either the causal melded reduction of  $A_1$  or of  $A_2$ . If it is the causal melded reduction of  $A_1$  ( $A_2$ ) individuals who receive  $c$  in the observed trial receive  $a$  ( $b$ ) when treated with the treatment " $a$  or  $b$ ".

**THEOREM F3.** If the suppositions of Theorem F2 hold for each " $G^A$ " that is a Stage 0 counterpart of some " $G^B$ ", and  $B$  is a causal melded Stage 0 reduction of FR CISTG  $A$ , then  $B$  is an FR CISTG.

**THEOREM F4.** If  $B$  is a causal melded Stage 0 reduction of FR CISTG  $A$  and Assumptions R of Theorem F1 hold, then  $B$  is an FR CISTG.

*Proofs.* For both theorems, we must show  $p[H^B(\cdot i_s^B) | \cdot i_s j_s^B] = p[H^B(\cdot i_s^B) | (\cdot i_s^B)]$ . But

$$p[H^B(\cdot i_s^B) | \cdot i_s j_s^B] = \sum_{[\cdot i_s j_s^A] \subset [\cdot i_s j_s^B]} p[H^B(\cdot i_s^B) | \cdot i_s j_s^A, \cdot i_s^B] p[\cdot i_s j_s^A | \cdot i_s j_s^B]. \quad (\text{F6})$$

Now under the suppositions of Theorem F3, Eq. (F6) equals

$$\sum_{\cdot i_s^A \in \cdot i_s^B} p[H^B(\cdot i_s^B) | \cdot i_s^B, \cdot i_s^A] p[\cdot i_s^A | \cdot i_s^B] \quad (\text{F7})$$

where the sum in Eq. (F6) over  $\cdot i_s j_s^A$  has been replaced by a sum over  $\cdot i_s^A$  in Eq. (F7) because  $B$  is a melded reduction  $A$ ;  $\cdot i_s j_s^A$  is replaced by  $\cdot i_s^A$  in the first term in the summand of Eq. (F6) because  $A$  is an R CISTG and, since  $B$  is a causal melded Stage 0 reduction of  $A$ ,  $H^B(\cdot i_s^B)$  is a function of  $H^A(\cdot i_s^A)$ . Finally,  $\cdot i_s^A$  and  $\cdot i_s^B$  can replace  $\cdot i_s j_s^A$  and  $\cdot i_s j_s^B$  in the second term in the summand by the supposition of Theorem F2. Summing over the  $\cdot i_s^A$  proves Theorem F3. To prove Theorem F4, it is sufficient to show that the first term in the summand of Eq. (F7) does not depend on  $\cdot i_s^A$  under Assumptions R (see Ref. [7]).

**LEMMA F1.** Let PISTG  $B$  be a stage 0 reduction of SCISTG  $A$  such that no two intranodal lines in the same node on  $A$  have the same intranodal line on  $B$  as counterpart. Suppose that whenever  $\cdot i_s j_s^A$  and  $\cdot i_s' j_s'^A$  are in the same  $\cdot i_s^B$  (and whenever the following suppositions are defined for some  $G_1^A$  and  $G_2^A$ ), for any individual  $i$ , (1)  $S(\cdot i_s j_s^A | G_1^A, i) = S(\cdot i_s' j_s'^A | G_2^A, i)$  and (2)  $p(\cdot i_{s+1}^B | \cdot i_s^B(t_{s+1}), \cdot i_s j_s^A, i, G_1^A) = p(\cdot i_{s+1}^B | \cdot i_s j_s^B(t_{s+1}), \cdot i_s' j_s'^A, i, G_2^A)$ ; then  $H(\cdot i_s^B)$  induces a natural unique  $H(\cdot i_s^B)$  so that PISTG  $B$  is an SCISTG  $B$ .

*Proof.* See Ref. [7].

**LEMMA F2.** If the suppositions of Lemma F1 hold and SCISTG  $A$  is an R SCISTG, then SCISTG  $B$  is an R SCISTG and suppositions (1) and (2) of the supposition hold when not conditioned on  $i$ .

*Proof.* Use the following construction. Given individual  $i$  is in  $\cdot i_s^B$  in the observed study, select a  $G_1^A$  such that there exists some  $\cdot i_s^A$  on  $G_1^A$  and contained in  $\cdot i_s^B$ . Let  $(\cdot i_s^B, D)$  be that individual's covariate history (on MPISTG  $B$ ) and death time when treated with  $G_1^A$ . Then, for each  $G^B$  such that  $\cdot i_s^B \in G^B$  define  $(\cdot i_s^B, D)$  to be that individual's covariate history and death time when treated with  $G^B$ . As the  $G_1^A$ 's vary over  $G^A$ , this procedure will exhaust the  $G^B \in G^B$  with  $\cdot i_s^B \in G^B$ . This process defines a unique (because of suppositions (1) and (2))  $H(\cdot i_s^B)$ .

## APPENDIX G

### The effect of a generalized treatment controlling for a time-dependent covariate

Given a PISTG  $A$ , suppose we have covariate histories  $i_1 i_2 \dots i_s^B \equiv \cdot i_s^B$  defined as follows. Each  $[\cdot i_s^A]$  is contained in some  $[\cdot i_s^B]$  (abbreviated  $\cdot i_s^A \in \cdot i_s^B$ ),  $[\cdot i_{s-1} \cdot i_s^B] \subset [\cdot i_{s-1}^B]$ , and the union of the  $[\cdot i_s^B]$  for fixed  $s$  is the entire population alive at  $t_s$ .

**EXAMPLE.** For PISTG C8.3 (as defined in Sec. 8), we could let a particular  $\cdot i_s^B$  be the subset of the population with a particular cigarette smoking history  $C(t_s)$ .

*Definitions G1.*

$$\begin{aligned} \gamma_D(t_{s+1} | "G_1^A", \cdot i_s^B) &\equiv \frac{\sum_{\cdot i_s^A \in \cdot i_s^B \cap G_1^A} \gamma_D(t_{s+1} | \cdot i_s^A j_s^A) p(\cdot i_s^A | "G_1^A")}{\sum_{\cdot i_s^A \in \cdot i_s^B \cap G_1^A} p(\cdot i_s^A | "G_1^A")} \\ p[\cdot i_s^B | "G_1^A"] &\equiv \sum_{\cdot i_s^A \in \cdot i_s^B \cap G_1^A} p(\cdot i_s^A | "G_1^A") \\ p[\cdot i_{s+1}^B | "G_1^A", \cdot i_s^B] &\equiv \frac{\sum_{\cdot i_s^A \in \cdot i_s^B \cap G_1^A} p(\cdot i_{s+1}^B | \cdot i_s j_s^A(t_{s+1})) p(\cdot i_s j_s^A(t_{s+1}) | "G_1^A")}{\sum_{\cdot i_s^A \in \cdot i_s^B \cap G_1^A} p(\cdot i_s j_s^A(t_{s+1}) | "G_1^A")} \end{aligned}$$

where  $j_s^A(\cdot i_s^A)$  is determined by “ $G_1^A$ ,” and  $\cdot i_s^B \cap G_1^A$  are the  $\cdot i_s^A$  on graph “ $G_1^A$ ” with history  $\cdot i_s^B$ .

If  $A$  is an R CISTG then, when the quotation marks are erased from around  $G_1^A$ , the above definitions refer to probability statements about the controlled trial defined by  $G_1^A$ .

**Definition.** The “ $G$ ”-independence null hypothesis holds for  $I_S^B = \{\cdot i_s^B; s \in (1, \dots, S + 1)\}$  if and only if

$$p[\cdot i_{s+1}^B | "G_1^A", \cdot i_s^B] = p[\cdot i_{s+1}^B | "G_2^A", \cdot i_s^B] \quad (\text{G1})$$

for all  $G_1^A, G_2^A \in G^A$  and all  $\cdot i_s^B$ .

Given a PISTG  $Q$ , suppose that for each  $\cdot i_s^Q$  the  $\cdot i_s j_s^Q$  can be grouped into one of  $K$  categories.

**EXAMPLE 1.** Let PISTG 8.3 be PISTG  $Q$  and group individuals in each  $\cdot i_s j_s^{8.3}$  into one of two categories, depending on whether or not they are a current smoker at  $t_s$ .

Consider a coarser PISTG  $A$  formed from PISTG  $Q$  by, for each  $\cdot i_s^Q$ , giving the  $\cdot i_s j_s^Q$  in each category  $k, k \in (1, \dots, K)$ , a separate right circumference point. Then each  $[\cdot i_s^A]$  is the union of the  $[\cdot i_s j_s^Q]$  in a single category  $k$ . Define  $[i_1^B]$  to be the union (both within and across nodes) of the  $[i_1^A]$  associated with category  $k$ . There are  $K$  such  $[i_1^B]$ . Let each of the  $K^2 [i_1 i_2^B]$  be the union of the  $[i_1 i_2^A]$  with  $[i_1^A]$  in some level  $k_1$ , and  $[i_1 i_2^B]$  in some level  $k_2$ , and so on.

**EXAMPLE 1 (continued).** PISTG C8.3 would be the coarser PISTG  $A$  if PISTG 8.3 were PISTG  $Q$ . Again, the  $[\cdot i_s^B]$  would be defined by cigarette smoking history through  $t_s$ .

**Definition.** “ $G^Q$ ” = (“ $G^A$ ”,  $\cdot i_s^B$ ) is the generalized treatment of PISTG  $Q$  defined by the property  $\cdot i_s j_s^Q \in "G^Q" \Leftrightarrow [\cdot i_s j_s^Q] \subset [\cdot i_s^B]$  and  $[\cdot i_s j_s^Q] \subset [\cdot i_s j_s^A]$  for some  $\cdot i_s j_s^A \in G^A$ , where  $\cdot i_s^B$  is the initial part of  $\cdot i_s^Q$ .

**Definition.** If  $Q$  is an FR MCISTG we say that there is no population effect of  $G^A$  controlling for  $\cdot i_s^B$  if and only if for all  $G_1^A, G_2^A, \cdot i_s^B$ ,

$$S[t, "G^Q"] = ("G_1^A", \cdot i_s^B), "G_2^Q" = ("G_2^A", \cdot i_s^B)] \equiv 0. \quad (\text{G2})$$

**EXAMPLE 1 (continued).** Note that the above definition suggests that when FR MCISTG  $Q$  is FR MCISTG 8.3,  $G^{C8.3}$  rather than  $G^{3.4}$  should be in Eq. (8.10). Of course, this is not necessary because there is many-to-one map from the  $G^{C8.3}$  to the  $G^{3.4}$  when, as in Eq. (8.10),  $C(t_s)$  is fixed.

**THEOREM G1.** Given  $Q, A$ , and  $\cdot i_s^B$  defined as above if (1)

$$p[\cdot i_{s+1}^B | \cdot i_s^B, D > t_{s+1}, \cdot i_k j_k(t_{k+1})^Q, \cdot i_k j_k i_{k+1}^Q] \text{ does not depend on } \cdot i_k j_k i_{k+1}^Q \quad (\text{G3})$$

for  $1 \leq k \leq s$  and (2) Eq. (G2) holds, then, for  $1 \leq k \leq s$ ,

$$\gamma_D[t_s + \Delta t | \cdot i_s^B, \cdot i_k j_k^A, \cdot i_k^A] \text{ does not depend on } \cdot i_k j_k^A. \quad (\text{G3a})$$

**Proof.** See Ref. [7].

**THEOREM G2.** Equation (G1) and (G3) hold  $\Leftrightarrow$  for  $s \geq 1$

$$p[\cdot i_{s+1}^B | \cdot i_s^B, D > t_{s+1}, \cdot i_s j_s i_{s+1}^Q] = p[\cdot i_{s+1}^B | \cdot i_s^B, D > t_{s+1}] \quad (\text{G4})$$

**Corollary G2.** Given Eq. (G1) and (G3) hold

(1) Equation (G2) holds  $\Leftrightarrow S(t, "G_1^A", "G_2^A") \equiv 0 \Leftrightarrow$

$$\gamma_D(t_{s+1} | "G_1^A", \cdot i_s^B) = \gamma_D(t_{s+1} | "G_2^A", \cdot i_s^B) \quad (\text{G5})$$

for all  $G_1^A, G_2^A \in G^A$  and all  $\cdot i_s^B$ .

Theorem G2 continues to hold when Eqs. (G3) and (G4) are modified such that  $1 \leq k \leq s$  is replaced by  $0 \leq k \leq s$  in Eq. (G3), and  $s \geq 1$  is replaced by  $s \geq 0$  in Eq. (G4).

**THEOREM G3.** Equation (G4), when modified as above,  $\Rightarrow$

$$\gamma_D[t_{s+1} | "G^Q"] = ("G_1^A", \cdot i_s^B) = \gamma_D(t_{s+1} | "G_1^A", \cdot i_s^B) \quad (G6)$$

where  $\cdot i_s^B$  is the initial part of  $\cdot i_s^B$ .

**THEOREM G4.** If Eq. (G1) and Eq. (G5) hold  $\Leftrightarrow$  Eq. (G3a) holds and

$$p[\cdot i_{s+1}^B | \cdot i_s^B, \cdot i_k j_k^A, \cdot i_k^A] \text{ does not depend on } \cdot i_k j_k^A. \quad (G7)$$

for  $1 \leq k \leq s$ . In addition, Eq. (G1) implies that  $S(t, "G_1^A", "G_2^A") \equiv 0 \Leftrightarrow$  Eq. (G5) holds.

**Corollary G4.** Given a subset  $\tau^A$  of  $G^A$ , if Eq. (G1) holds (only) for the  $G_1^A, G_2^A \in \tau^A$ , then  $S(t, "G_1^A", "G_2^A") \equiv 0$  for  $G_1^A, G_2^A \in \tau^A \Leftrightarrow$  Eq. (G5) holds for  $G_1^A, G_2^A \in \tau^A$ .

**THEOREM G5.** If PISTG A is an FR CISTG and for each individual  $j$ , (1)  $p[\cdot i_{s+1}^B | G_1^A, \cdot i_s^B, j]$  does not depend on  $G_1^A$  for all  $G_1^A \in \tau^A \subset G^A$  and (2)  $S(t, G_1^A, G_2^A, j) \equiv 0$  for all  $G_1^A, G_2^A \in \tau^A$ , then Eqs. (G1) and (G5) hold for any  $G_1^A, G_2^A \in \tau^A$ . Since (1) implies each individual  $j$  has a unique history  $\cdot i_s^B$  when treated with a  $G_1^A$  (which we write as  $\cdot i_s^{j,B}$ ), the outcome of an individual  $j$  when treated with  $G^Q = (G_1^A, \cdot i_s^{j,B})$ ,  $G_1^A \in \tau$  is well defined even when PISTG Q is not a CISTG. Thus for  $G_1^A, G_2^A \in \tau^A$  supposition (1) implies that

$$\begin{aligned} S(t, G^Q) &= ("G_1^A", \cdot i_s^{j,B}), \quad G_2^Q = ("G_2^A", \cdot i_s^{j,B}), j) \equiv 0 \\ &\Leftrightarrow S(t, G_1^A, G_2^A, j) \equiv 0. \end{aligned}$$

*Proofs.* See Ref. [7].

## REFERENCES

1. E. S. Gilbert, Some confounding factors in the study of mortality and occupational exposures. *Am. J. Epidemiology* **116**, 177–189 (1982).
2. J. M. Robins, A statistical method to control for the healthy worker effect in intracohort comparisons. *Am. J. Epidemiology* **120**, 465 (1984).
3. J. M. Robins, A new theory of causality in observational survival studies—Application to the healthy worker effect. *Biometrics* **41**, 331 (1985).
4. D. B. Rubin, Bayesian inference for causal effects: The role of randomization. *Ann. Statist.* **6**, 34–58 (1978).
5. P. R. Rosenbaum and D. B. Rubin, Comment: Estimating the effects caused by treatment. *J. Am. Statist. Assn.* **79**, 26–28 (1984).
6. J. W. Pratt and R. Schlaifler, On the nature and discovery of structure. *J. Am. Statist. Assn.* **79**, 9–21 (1984).
7. J. M. Robins, Some theorems relating to causal tree graphs. Unpublished manuscript (1986).
8. A. M. Lee and J. F. Fraumeni, Arsenic and respiratory cancer in man: An occupational study. *J. Natl. Cancer Inst.* **42**, 1045–52 (1969).
9. Richard Monson, private communication (citation p. 5-3).
10. D. R. Cox, Regression models and life tables (with discussion). *J. R. Statist. Soc. B* **34**, 187–220 (1972).
11. D. C. Thomas, Appendix to F. D. K. Liddell, J. C. McDonald, D. C. Thomas. Methods of cohort analysis. Appraisal by application to asbestos mining (with discussion). *J. R. Statist. Soc. A* **140**, 469–491 (1977).
12. D. Oakes, Survival times: Aspects of partial likelihood. *Int. Statist. Rev.* **49**, 235–264 (1981).
13. N. E. Breslow, J. H. Lubin, P. Marek and B. Langholz, Multiplicative models and cohort analysis. *J. Am. Statist. Assn.* **78**, 1–12 (1983).
14. N. E. Breslow, Contribution to discussion of paper by D. R. Cox. *J. R. Statist. Soc. B* **34**, 216–217 (1972).
15. J. M. Robins and H. Morgenstern, Confounding and Prior Knowledge, Technical Report No. 1. Occupational Health Program, Harvard School of Public Health, Boston (1983).
16. O. Miettinen and F. Cook, Confounding: Essence and detection. *Am. J. Epidemiology* **134**, 593–603 (1981).
17. J. M. Robins, The Statistical Foundations of Confounding in Epidemiology, Technical Report No. 2. Occupational Health Program, Harvard School of Public Health, Boston (1983).
18. A. P. Dawid, Conditional independence in statistical theory (with Discussion). *J. Roy. Stat. Soc. B* **41**, 1–31 (1979).

19. P. R. Rosenbaum, The consequences of adjustment for a concomitant variable that has been adversely affected by treatment. *J. Roy. Stat. Soc. A* **147**, 656–666 (1984).
20. D. A. Hsieh, C. F. Manski and D. McFadden, Estimation of response probabilities from augmented retrospective observations. *J. Am. Statist. Assn.* **80**, 651–662 (1985).
21. Charles Poole, Exposure opportunity in case-control studies. *Am. J. Epidemiology* **123**, 352–358 (1986).
22. D. C. Thomas, General relative risk models for survival time and matched case-control analysis. *Biometrics* **37**, 673–686 (1981).
23. A. S. Whittemore and A. McMillan, Lung cancer mortality among U.S. uranium miners: A reappraisal. *J. Natl. Cancer Inst.* **71**(3), 489–499 (1983).
24. J. D. Kalbfleisch and R. L. Prentice, *The Statistical Analysis of Failure Time Data*. Wiley, New York (1980).
25. A. J. Sasco, N. E. Day, and S. D. Walter, Case-control studies for the evaluation of screening. *J. Chronic Dis.* **39**, 399–405 (1986).
26. O. Barndorff-Nielsen, *Information and Exponential Families in Statistical Theory*. Wiley, New York (1978).