

ERRATA TO "A NEW APPROACH TO CAUSAL INFERENCE  
 IN MORTALITY STUDIES WITH A SUSTAINED  
 EXPOSURE PERIOD—APPLICATION TO CONTROL OF  
 THE HEALTHY WORKER SURVIVOR EFFECT"  
*Mathl Modelling* 7(9-12), 1393-1512 (1986)

J. M. ROBINS

Occupational Health Program, Harvard School of Public Health, 665 Huntington Avenue, Boston,  
 MA 02115, U.S.A.

Errata marked with an asterisk must be corrected by the reader to insure proper meaning. Errata not marked by an asterisk are of less importance. Any statement within the errata list that is found within brackets is a note to the reader and not necessarily a correction.

Page/Paragraph/Line	Uncorrected	Corrected
1398/6/6	thus	this
*1399/last/first	exposure histories	exposure histories or paths (i.e., OEH or OEP)
*1401/2nd line of condition 2C.3	$t_2$ (that until $t_1$	$t_1$ (that until $t_2$
*1401/Condition 2C.3/line 3	$t_{S+1}$	$S + 1$
1406/2/2	MCISTG $B$	MCISTG $A$
1413/2/3	survival would	survival, when controlling for employment history, would
*1415/3/2		
*1417/2/2	at $t_i$ consists	at $t_1$ consists
1417/3/1	factor apparent	factor, apparent
1421	[In Definition 4.4, if a subject is in both $[i'_s]$ and $[i'_{s+1}]$ in the observed study, the sets $HT(i'_s)$ and $HT(i'_{s+1})$ are assumed consistent with one another.]	
*1423/3rd to last/last	$t_1$ to $i_s$ .	$t_1$ to $i_s$ . $p(i_s   "G") \equiv 0$ if $i_s$ is not on "G".
*1423/2nd to last/first	$p[i_s j_s(t_{s+1})   "G"]$	If $i_s j_s(t_{s+1})$ is on "G", $p[i_s j_s(t_{s+1})   "G"]$
*1423/2nd to last/last	node $i_s j_s(t_{s+1})$	node $i_s j_s(t_{s+1})$ . $p[i_s j_s(t_{s+1})   "G"] \equiv 0$ if $i_s j_s(t_{s+1})$ is not on "G"
*1424/1/4	above propositions	above definitions
1425/Section F	[In the definition of a fundamental MPISTG any subset of protocol at $t_s$ is <i>not</i> eligible to be randomly assigned to a treatment at $t_s$ . Since a double blind ordinary designed randomized trial is a special case of an alternative designed randomized trial the discussion in this subsection is applic- able to ordinary designed randomized trials as well.]	

[continued overleaf

*continued]*

Page/Paragraph/Line	Uncorrected	Corrected
1428/2/1	Appendix A) for	Appendix A) for the curve
1434	[I've recently shown that, for MPISTG 3.5 (modified so that more than two exposure levels can arise from any right circumference point), there always exists a particular $G$ -null test defined by some table weight and exposure scoring function with asymptotic power equal to that of the Cox score test based on equation (6.2) for local alternatives given by equation (6.1). In the example given in the first paragraph on p. 1434, for a case failing at $t$ , the optimal $G$ -null test uses an exposure scoring function that assigns a score to a particular exposure level at $t_1$ proportional to the cumulative exposure up to $t$ , of the subject with that exposure level.]	
1435/1/3	its nominal level	their nominal level
1435/End of first full paragraph (just preceding Section 6.C)	[Another approach to increasing power for detecting alternatives to the $G$ -null hypothesis of FR MCISTG 3.4 can be based on the fact that if, among subjects at work at any $t_k$ , the probability of receiving any particular exposure level at $t_k$ depends only on employment and exposure history in the past $m$ time periods, then an (unconditionally) valid test of the $G$ -null hypothesis of FR MCISTG 3.4 (that will often have much greater power than the $G$ -null test) can be obtained based on the following. <i>m</i> -modified $G$ -null test algorithm. In step 3 of the $G$ -null test algorithm after constructing the tables $(\cdot i_k, t_s)$ , combine into a single table of exposure level at $t_k$ vs case control status at $t_s$ , those tables that have the same exposure history in the interval $[t_{k-m}, t_{k-1}]$ and employment history in the interval $[t_{k-m}, t_k]$ . Apply the remainder of the $G$ -null test algorithm to these "combined" tables [rather than to the original tables $(\cdot i_k, t_s)$ ]. A program that implements the $m$ -modified $G$ -null test algorithm and implements a test of the assumption that the probability of receiving a particular exposure level at $t_k$ is conditionally independent of exposure and employment history experience more than $m$ time periods previously is available from the author. The $m$ -modified $G$ -null test is not in general valid under the condition that among the subsets of the population with identical exposure and employment history for the last $m$ time periods, subsequent survival is independent of employment and exposure history experience more than $m$ time periods previously on account of non-null correlations between the table-specific numerators contributing to the $m$ -modified test.]	
*1437/2/1	Although the value of	Although the parameter estimated by
*1438/2nd to last/5	8.3 (see Fig. 8.3)	EM8.3 (as defined on p. 1453 following eq. (8.12))

Page/Paragraph/Line	Uncorrected	Corrected
*1438/2nd to last/7 *1440/2/2 1446/4/at end	PISTG 8.3 $D(t G^A)$	PISTG EM8.3 $S(t G^A)$
	[A consistent test is a test that, given a sufficiently large sample size, would reject the null hypothesis when false with probability 1. A valid test is a test that would reject at its nominal level for all states of nature consistent with the null hypothesis.]	
*1448/2/2 1448/4/1 1449/1/3 1449/3/4	FR MSCISTG and or MPISTG valid valid	R MSCISTG and/or MPISTG valid and/or consistent consistent
*1449/3/Lemma 8.3, Conclusion (2)	$E(t)$ if and only if	$E(t)$ implies
*1449/3/Lemma 8.3, Conclusion (2)	valid and consistent tests	consistent tests
*1449/4	[In light of the corrected version of Lemma 8.3 given just above, the example that follows Lemma 8.3 is slightly misleading. Specifically, even were exposure not a predictor of future $L$ -status, valid tests of the $G$ -null hypothesis of FR MCISTG 3.4 would have to be based on the $G$ -null test for MPISTG 3.4 rather than on a test of whether $L$ was an independent population risk factor for death controlling for exposure.]	
*1451/Heading C	estimated	tested
*1456/4/Remark 8.3	[Replace all of Remark 8.3 by the following: (a) and (b) together imply (c) holds by Lemma 8.3 as corrected in these errata.]	
*1457/4/Remark 8.9	Eq. (8.19) implies that Eq. (20) holds $\Leftrightarrow$	Together Eqs. (8.19) and (8.20) imply
*1457/6/9 *1460/Lemma 8.12/last line	i.e. do not Eq. (8.18) holds	then do not Eq. (8.18) holds. More precisely in Lemma 8.12, we are assuming MPISTGs CF8.3 and 8.3 (see legend to Fig. 8.3) are MCISTGs. Eqs. (8.24)–(8.26) are propositions about these MCISTGs.
	[Specifically, Eq. (8.24) is equivalent to $p[C(t + \Delta t) C(t), G_1^{CF8.3} = [E_1(t_s), L(t_s), i] = p[C(t + \Delta t) C(t), G_2^{CF8.3} = [E_2(t_s), L(t_s), i]$ (8.26a) for all $E_1(t_s), E_2(t_s), L(t_s), i$ . Eq. (8.26) is equivalent to Eq. (8.26a) with the roles of $E$ and $L$ interchanged. Eq. (8.25) is $p[L(t + \Delta t) L(t), G_1^{8.3} = [G_1^{3,4}, C(t_s), i] = p[L(t + \Delta t) L(t), G_2^{8.3} = [G_2^{3,4}, C(t_s), i]$ for all such $G_1^{8.3}, G_2^{8.3}$ and $i$ .]	
*1460/Lemma 8.14	then we have Eq. 8.27 holds if and only if	and Eq. 8.27 holds, then
*1461/2	[The proof of the above corrected version of Lemma 8.14 is given in Ref. [7].]	

[continued overleaf]

*continued*

Page/Paragraph/Line	Uncorrected	Corrected
*1463/2	[Lemma 8.18 should read: If Eqs. 8.28 and 8.29 hold then Eqs. 8.29 and 8.21 hold. The proof of the corrected version is given in Ref. [7].]	
1466/1/2nd to last	with a given	within a given
1468/5/11	100,000/160,000	10,000/16,000
*1468/5/2nd to last	adverse effect	beneficial effect
*1468/5/last line	.08	.16
1468/6/2nd to the last	data	data, provided exposure is received at random within levels of smoking.
1468/6/last	controlling for smoking behavior	within levels of smoking
1471/2/3	ample	sample
*1471/4/3	[The expression in brackets in paragraph four is true only under large sample limiting model 1. Under sparse data limiting model, the unconditional power of the $G$ -null test for MPISTG 3.4 may be greater than that of MPISTG C8.3. Furthermore, under a sparse data model, approximate ancillaries are not well defined.]	
*1472/3/last line	If not	If $C$ was not a causal risk factor,
*1472/4/10	. Corollary E1	; and use a M-H test to combine information over tables. Corollary E1
*1482/1/3rd to last	presumably at	presumably less than
*1482/1/2nd to last	for less than $x$	for $x$
1482/end of page	[One possible approach, which we do not use in Section 11.E, is to note that if MCISTG 3.4 is an FR MCISTG, we can estimate the minimum latent period $x$ using steps 1 and 2 of the lagged exposure test algorithm as given on the bottom of p. 1481 and top of p. 1482. Then, if we can assume assumption G.3 holds <i>a priori</i> , a test of whether Eq. (11.7) holds becomes a consistent test of whether the healthy worker survivor effect lasts for more than $x$ years. Even if there is no $x$ for which the healthy worker survivor effect lasts for less than $x$ years, nonetheless if, for some $x$ , $p[L(t_s - x + \Delta t)   L(t_s - x), E(t_s - x), D > t_s + \Delta t]$ does not depend on $E(t_s - x)$ , then a test of whether $\gamma_D[t_s + \Delta t   E(t_s - x), L(t_s - x)]$ depends on $E(t_s - x)$ is a consistent (but not a valid) test of the null hypothesis of no direct exposure effect. This follows because under the null of no exposure effect, for any $x$ , MPISTG 3.4( $x$ ) is an FR MCISTG for which the " $G$ "-null hypothesis holds. Our result follows by applying the corrected version of Lemma 8.3 to MPISTG 3.4( $x$ ).]	
1484/Table 4	Outline	Outcome
1491/Theorem 12A.5	$S \geq 0$	$s \geq 0$
1491/last paragraph	[Note that STG 8.3 modified such that $D_2$ is $c$ and $D_1$ is $D$ will be an MPISTG only if, in addition to assuming that deaths from $D_2$ occur precisely at $t_s$ , covariate status (e.g.,	

*[continued overleaf*

Page/Paragraph/Line	Uncorrected	Corrected
	employment status) at $t$ , must have been measured at $t_s - \sigma$ , $\sigma$ infinitesimal, rather than at $t_s$ . This follows because if the modified STG 8.3 is an MPISTG, we must know the 1-status at $t_s$ of individuals who are $D_2$ at $t_s$ . Since subjects have an 1-status only if alive, their 1 status at $t_s$ must have been ascertained prior to their death at $t_s$ from cause $D_2$ . (Even so, this modified STG 8.3 is not strictly an MPISTG because subjects who are $D_2$ at $t_s$ do not have $D_1$ status or 1-status recorded at $t_{s+1}$ . Since we are only concerned with the generalized treatments of this modified MPISTG 8.3 that assign $\bar{D}_2$ at each node, no difficulty arises. Also see Section AD.5 of the accompanying Addendum.]	
1495 and 1496 (all occurrences)	FR OCISTG FR MCISTG	R OCISTG R MCISTG
1496/1/2	[This change reflects the fact that smoking may not be received at random with respect to future $L$ -history.]	
1497/3/3rd to last	on $L$ -, $C$ -	on $x$ -lagged $L$ -, $C$ -
1499/Theorem 12 D.5/4th line	that discovery the	that the
*1500/3/5	rotation	notation
	$\theta_1^x \theta_1^{N-x} \theta_3^z \theta_3^{N-x-y-z}$	$\theta_1^x (1 - \theta_1)^{N-x} \theta_3^z (1 - \theta_3)^{N-x-y-z}$
*1501/3/3rd to last	projected exposure path	the PEH
1504/2/5	used to estimate	used the Monte Carlo
	the Monte Carlo	
1504/2/5	algorithm, for	algorithm to estimate,
		for
*1507/8/2	of PISTG 8.3	of PISTG EM8.3 (as defined on page 1453 following Eq. (8.12))
		PISTG EM8.3
*1507/3rd from last/4	PISTG 8.3	PISTG EM8.3
*1508/Definition/2-3	Stage 0 counterpart	Stage 0 reduction
*1509/1/4	melded reduction A	melded reduction of A
*1509/Lemma F2	[The proof given for Lemma F2 is actually the proof of Lemma F1. The proof of Lemma F2 is given in Ref. [7].]	
*1509/Lemma F2/2nd line	of the supposition hold	of Lemma F1 hold
*1511/Theorem G4	If Eq. (G1)	Eq. (G1)
1511/Theorem G4/last line	[Delete the last sentence and replace by: Together, Eqs. (G1) and (G5) imply $S(t_1, "G_1^A", "G_2^A") \equiv 0$ .]	