THE ROLE OF FRAILTY MODELS AND ACCELERATED FAILURE TIME MODELS IN DESCRIBING HETEROGENEITY DUE TO OMITTED COVARIATES

NIELS KEIDING and PER KR GH NDERSEN

Department of Bio tati tic , Univer ity of openhagen, Blegdam vej 3,
DK-2200 openhagen N, Denmark

and

JOHN P. KLEIN

Department of Bio tati tic , The Medical ollege of i con in,
8701 aterto n lank oad, Mil aukee, i con in 5322 , U.S.A.

SUMM R

In urvival analy i , deviation from proportional hazard may ometime by e plained by unaccounted random heterogeneity, or frailty. Thi note recall the literature on omitted covariate in urvival analy i and how in a ca e tudy how un tably frailty model might behave when a ed to account for unob erved heterogeneity in tandard urvival analy i with no replication per heterogeneity unit. ccelerated failure time modelling eem to avoid the e difficultie and al o to yield ea ily interpretable re ult .

We propo e that it would be advantageou to upgrade the accelerated failure time approach along ide the hazard modelling approach to urvival analy i .

1. INT ODUCTION

Stati tical modelling of heterogeneity may be ba ed on tratification according to factor , regre ion on covariate , or by a uming a probability di tribution of time framewor for interpretation of covariate effect in urvival analy i with random heterogeneity.

The purpo e of thi note i to briefly recapitulate the above framewor and to pre ent another ca e tudy which, li e that of Hougaard et al.[16], indicate that accelerated failure model may be preferable in accounting for (re idual) heterogeneity in univariate (" ingle- pell") urvival time due to "mi ing" (omitted, unrecorded) covariate .

Section 2 pre ent a brief partial urvey on approache to the tudy of omitted covariate in the 1980 , and Section 3 briefly recall the proportional hazard frailty model with a pect of current technique for it   tati tical analy i . Section 4 pre ent and lightly e tend the Struther -Kalbflei ch heuri tic on omitted covariate in urvival analy i ba ed on a normal-theory linear model equivalent to the accelerated failure time model. Section 5 pre ent equi k3fl1T01Df1()the)Tjfl1CHES0TDflTDfl(to)TOjfl1721TDfl(to)THEjfl17291failureac200(ec

parameter , ma imum li

Let W have a tandard e treme value di tribution of a minimum, that i , the den ity of W i e p$(w - e^w), -\infty < w < \infty$. Then T follow the above Weibull di tribution, where

$$Y = \log T = -\frac{\log \kappa}{\nu} - \frac{\beta_1}{\nu} x_1 - \frac{\beta_2}{\nu} x_2 + \frac{W}{\nu}.$$

Thi i an *accelerated failure time model*: an ordinary regre ion problem of log(urvival time) on x_1 and x_2 with e treme value di tributed re idual with cale parameter ν^{-1}, regre ion coefficient $-\beta_1/\nu$ and $-\beta_2/\nu$ and intercept $-\nu^{-1} \log \kappa$. Borrowing e perience from normal-theory linear regre ion (i.e. a uming W tandard normal (0,1)), it i   een that the regre ion coefficient and intercept are e timated by the u ual regre ion e timate , in particular $E(\widehat{\beta_1/\nu}) = \beta_1/\nu$, ν^{-1} i e timated by the u ual re idual empirical variance s^2, and for large amw27TD0TDfl(24 2336TD^9c8180TDfl9-3040560TDfl(empide)TTg9t above, $\beta\tau$ i e timated by the u ual regre ion e timate, o $E(\widehat{\beta\tau}) = \beta\tau = \beta_1/\nu$ (= the theoretical regre ion of Y on x_1). Therefore $\hat{\beta} \xrightarrow{P} \beta = \beta_1 \nu^{-1/2}/\tau$, which i clo er to 0 than β_1: there i the well- nown *attenuation due to an omitted covariate.* Furthermore

$$\text{a . var.}(\hat{\beta}) = \frac{1}{n}\left(\frac{1}{\sigma^2_{x_1}} + \frac{\beta_1^2}{2\nu\tau^2}\right) < \text{a . var.}(\hat{\beta}_1) \quad ;$$

the tandard error i al o attenuated, indeed if $\sigma^2_{x_1}$ i large, the WaldTt-dte 5703111141389998792061113111 bution ha changed, now being that of $(W+U)/\nu$. gain borrowing e perience from normal-theory linear regre ion, $-\beta_1/\nu$ would be e timated by the u ual regre ion e timate, $E(\widehat{\beta_1/\nu}) = \beta_1/\nu$, but if we had erroneou ly a umed no frailty ($U=0$), $\nu-1$ would have been overe timated by the factor $\eta = (\mathrm{Var}W + \mathrm{Var}U)/\mathrm{Var}W$ and the hazard model regre ion parameter $\beta_1 = (\beta_1/\nu)/\nu^{-1}$ imilarly undere timated by the factor $\eta 1$, leading to *attenuation by di regarding frailty.*

onclu ion. For the Weibull model the accelerated failure time parametrization conveniently eparate regre ion coefficient from di per ion parameter , allowing unchanged e timation of regre ion coefficient under the frailty-amended model, which only contribute to the di per ion. Thi wa previou ly pointed out by Hougaard et al.[16].

5. EXAMPLE

nder en et al.[2] con idered in their E ample VII.3.1, VII.3.4 and IX.4.3 urvival after operation for malignant melanoma for 205 patient frailtIX.4.3fl/T141.Tffl10Dfl[(t e141 and uai 3/yT1Tffl1who0

andT14h.n

imilar way of incorporating the e covariate . If the covariate are included in a tandard Co model the e timated regre ion coefficient and tandard error were

log(tumour thic ne)	0.610 (0.176)
ulceration	0.971 (0.321)

but graphical chec (nder en et al.[2], Fig . VII.3.3 and VII.3.6) rai ed ome u picion that hazard for patient without and with ulceration, were not proportional but rather converging. Therefore a time-dependent covariate to account for po ible time $\times$ covariate interaction wa added:

log(tumour thic ne)	0.607 (0.177)
ulceration	1.082 (0.357)
ulceration $\cdot(\log(t) - 7)$	-1.198 (0.589);

here t i mea ured in day and $7 \sim \log(3 \times 365)$. li elihood ratio te t of no effect of the latter variable yielded $= .02$, giving ome evidence to upport the u pected deviation from proportionality.

Semiparametric frailty model.

Becau e thi deviation might be interpreted a a election effect in a heterogeneou population ari ing from important unmea ured confounder not being included in the analy i , a frailty model wa po tulated. To the

Co regre ion model pecification of the death inten ity with the two covariate wa multiplied a frailty factor Z, a umed gamma di tributed with $E(Z) = 1, \mathrm{Var}(Z) = \delta$. The fitted parameter were (with the no-frailty model e timate attached for compari on)

	Frailty	No frailty
log(tumour thic ne)	1.370 (0.472)	0.610 (0.176)
ulceration	1.696 (0.686)	0.971 (0.321)
frailty variance	4.215 (2.266)	0 (-)

with li elihood ratio te t tati tic of no frailty variance yielding = .007. For detail on e timating the tandard error under the frailty model, cf. nder en et al.[27].

It i thu een that incorporation of unmea ured population heterogeneity in thi ca e *deattenuate* the effect of the mea ured covariate (a well a of their tandard error) by a factor of about 2.

eibull frailty model.

nder en et al.[2] noted that the underlying inten itie of the fitted Co regre ion model varied o regularly that a hypothe i of Weibull underlying inten ity hould be acceptable. In order to tudy theE etatio re vey by Klein et al.[4], a well a the power variance family $(\alpha, \psi, \ )$ due to Hougaard[32], of which all of the e are pecial ca e . Hougaard' model i mo t ea ily characterized by the Laplace tran form

$$\text{e p}\left\{-\frac{\psi}{\alpha}\left[(\ + s)^{\alpha} - \ ^{\alpha}\right]\right\} \quad .$$

Our gamma di tribution i $(0, \delta^{-1}, \delta^{-1})$, while $(\alpha, \psi, 0)(0 < \alpha < 1)$ are the po itive table di tribution and $(\frac{1}{2}, \psi, \ )$ the inver e Gau ian di - tribution . i well nown, the po itive table frailty di tribution lead to unidentifiability in the pre ent ca e of ob erving only one event per individual. For the other frailty model , with the no frailty model included for compari on, the e timate are given in Table 1.

It i een that the re ult from the all-inclu ive power variance frailty model are virtually indi tingui hable from that of the gamma frailty model, which in turn fit ignificantly better than the inver e Gau ian frailty and the no frailty/po itive table frailty (the latter two having the ame li elihood). l o, the e timate for no frailty and gamma frailty are well compatible with the emiparametric e timate quoted above, and al o there i a deattenuation factor of 2 to 3 on the regre ion parameter when con idering the gamma frailty model. The a umption of inver e Gau ian frailty yield intermediate re ult , and judging from the li elihood al o a le effective accounting for the heterogeneity.

Table 2 record the e timated correlation between the e timated frailty parameter (indicating the pread of the frailty di tribution) and the e timate of the regre ion coefficient and the Weibull hape parameter. The po itive correlation reflect the inherent negative correlation between two alternative way of de cribing the ob erved heterogeneity in urvival time : either by a large frailty parameter (wide frailty di tribution), or by a “flat” underlying inten ity (mall Weibull hape parameter). Indeed, while the underlying Weibull di tribution in the no-frailty model i in ignificantly different from an e ponential di tribution (hape parameter=1), a much more concentrated underlying di tribution i e timated for the gamma and inver e Gau ian frailty model .

The po itive correlation between e timated frailty parameter and e -timated regre ion parameter reflect the deattenuation effect de cribed in Section 3. Intuitively: The interindividual variation i *either* de cribed by covariate (high regre ion coefficient) *or* frailty (large frailty parameter).

Accelerated failure time interpretation.

lternatively, we may tart from the accelerated failure time (FT) interpretation outlined toward the end of Section 3. We then obtain the re ult of Table 3, accounting for the multiplicative indeterminacy in the po itive table frailty di tribution and till a uming underlying Weibull di tribution.

It i een that in the FT interpretation, the variou model agree. Let

$$
\begin{aligned}
\log(\text{ urvival time}) = \text{con t.} \quad & - \quad 0.60 \times \text{log tumour thic ne} \\
& - \quad 0.75 \times \text{ulceration} \\
& + \quad \text{noi e} \quad .
\end{aligned}
$$

That i , for fi ed value of ulceration, if tumour thic ne increa e by a factor α, urvival time will decrea e by a factor $\alpha^{0.60}$. Similarly, for fi ed value of tumour thic ne , ulceration of the tumour will decrea e life by a factor of $e^{-0.75} \approx 0.47$ compared to what it would have been if the tumour wa not ulcerated.

6. DI CU ION

railty interpretation: individual or population ri k. The original impetu for the frailty concept uch a defined by Vaupel et al.[1] wa to clarify the behaviour of the *mean ha ard among the urvivor* in a heterogeneou population. In our e ample we ob erved a (light) deviation from80TDfioTDfl(ed)Tjfl1o0020TDfl(t) the only lightly wor e fitting inver e Gau ian frailty di tribution deattenuation wa halved, and for the po itive table frailty model it (the parameter above) i inherently unidentifiable. (Motivated in part by thi feature of the po itive table frailty di tribution, Robin and Greenland[14,15] di cu ed con equence of uch unidentifiability problem for compen ation cheme). It i well nown that ratio of regre ion coefficient are much le en itive to model mi pecification than the regre ion coefficient them elve , ee Solomon[19] for e ample from the pre ent conte t and Li and Duan[31] for a careful general di cu ion with review of earlier wor . Thi i al o very apparent in our e ample.

conceptual e planation may be obtained from the ob ervation above about trong po itive correlation between the e timate of the Weibull hape parameter ν and the pread of the frailty di tribution. The ingle- pell data contain only limited power a to di tingui hing the random variation a within-individual (large ν) or between-individual (large frailty pread), and therefore interpretation ba ed only on the within-individual hazard are un table.

Accelerated failure time interpretation: een above the FT interpretation (which wa here fea ible tarting from log-Weibull error di tribution) avoid the unidentifiability problem by hifting attention of the dependence on covariate from the elu ive concept of 'individual hazard' to the acceleration factor of the life time it elf, thereby combining the within- and between-individual component of variation into much more tably determined functional . The heterogeneity i conveniently relegated to an overdi per ion elemen

REFERENCES

1. Vaupel, J.W., Manton, K.G. and Stallard, E. 'The impact of heterogeneity in individual frailty on the dynamic of mortality', *Demography*, **16**, 439-454 (1979).

2. nder en, P.K., Borgan, Ø., Gill, R.D. and Keiding, N. *Stati tical Model Ba ed on ounting roce e* , Springer, New or , 1993.

3. Niel en, G.G., Gill, R.D., nder en, P.K. and Sren en, T.I. . ' counting proce approach to ma imum li elihood e timation in frailty model ', *Scandinavian Journal of Stati tic* , **19**,

8. Oa e , D. ‘Frailty model for multiple event time ’, in Klein, J.P. and Goel, P.K. (ed), *Survival Analy i : State of the Art*, Kluwer, Dordrecht, 1992, pp. 371-379.

9. Pon O., Kaddour, . and de Turc heim, E. ‘ nonparametric approach to dependence for bivariate cen ored data’, in Klein, J.P. and Goel, P.K. (ed), *Survival Analy i : State of the Art*, Kluwer, Dordrecht, 1992, pp. 381-392.

10. Prentice, R.L. and Cai, J. ‘Marginal and conditional model for the analy i of multivariate failure time data’, in Klein, J.P. and Goel, P.K. (ed), *Survival Analy i : State of the Art*, Kluwer, Dordrecht, 1992, pp. 393-406.

11. Turnbull, B. ‘Multivariate failure time analy i : Di cu ion of paper by Oa e ; Pon , Kaddour and de Turc heim; and Prentice and Cai’, in Klein, J.P. and Goel, P.K. (ed), *Survival Analy i : State of the Art*, Kluwer, Dordrecht, 1992, pp. 407-414.

12. Hougaard, P. ‘Modelling heterogeneity in urvival data’, *Journal of Applied robability*, **2** , 695-701 (1991).

13. a hin, ., Vaupel, J. and Iachine, I. ‘Correlated individual frailty: n advantageou approach to urvival analy i of bivariate data’, *Mechani m of ageing and development*, **5**, 1-10 (1995).

14. Robin , J. and Greenland, S. ‘The probability of cau ation under a tocha tic model for individual ri ’, *Biometric* , **45**, 1125-1138 (1989).

15. Robin , J. and Greenland, S. 'E timability and e timation of e pected year of life lo t due to a hazardou e po ure', *Stati tic in Medicine*, **10**, 79-93 (1991).

16. Hougaard, P., Myglegaard, P. and Borch-John en, K. 'Heterogeneity model of di ea e u ceptibility, with application to diabetic nephropathy', *Biometric* , **50**, 1178-1188 (1994).

17. Gail, M.H., Wieand, S. and Piantado i, S. 'Bia ed e timate of

29. Murphy, S. . 'Con i tency in a proportional hazard model incorporating a random effect', *The Annal of Stati tic* , **22**, 712-731 (1994).

30. Murphy, S. . ' ymptotic theory for the frailty model', *The Annal of Stati tic* , **23**, 182-198 (1995).

31. Struther , C. . *A ymptotic propertie of linear rank te t ith cen ored data*, Ph.D. The i , Department of Stati tic , Univer ity of Waterloo, Ontario, 1984.

32. Hougaard, P. 'Survival model for heterogeneou population derived from table di tribution ', *Biometrika*, **73**, 387-396 (1986).

33. Li, K.-C. and Duan, N. 'Regre ion analy i under lin violation', *The Annal of Stati tic* , **17**, 1009-1052 (1989).

34. alen, O.O. ' linear regre ion model for the analy i of life time ', *Stati tic in Medicine*, , 907-925 (1989).

Table 1. E timate for Weibull frailt

Table 2. Weibull frailty model . Correlation between e timated frailty parameter and parameter e timate a pecified.

	Gamma frailty emiparametric	Gamma frailty Weibull	Inver e Gau ian frailty Weibull
Weibull hape parameter	—	.882	.793
log(tumour thic ne)	.632	.598	.323
ulceration	.532	.511	.430

Table 3. Weibull frailty model . Hazard rate regre ion coefficient contra ted to accelerated failure time regre ion coefficient .

	Gamma frailty	Inver e Gau ian frailty	No frailty (a umed=1) or Po itive table frailty (indeterminate)
Weibull hape parameter	2.917 (0.718)	1.747 (0.299)	1.150 · (0.131 ·)
log(tumour thic ne)	1.754 (0.592)	0.932 (0.281)	0.577 · (0.175 ·)
ulceration	2.180 (0.875)	1.512 (0.518)	1.020 · (0.322 ·)
$\frac{\text{log(tumour thic ne)}}{\text{Weibull hape parameter}}$	0.60 (0.15)	0.53 (0.18)	0.50 (0.16)
$\frac{\text{ulceration}}{\text{Weibull hape parameter}}$	0.75 (0.25)	0.87 (0.28)	0.89 (0.29)