Critical examination of the Armitage-Doll model shows that neither the exact nor the approximate solution describes cancer incidence data well, and that models incorporating cell proliferation kinetics do better.[4] Be that as it may, for any stochastic model, stochastic heterogeneity is immediately introduced if the stochastic solution is used. Inter-individual variations in susceptibility arising from variations in the rates of critical biological processes can be modelled by assuming a distribution on the parameters of the model. Major gene defects, such as FAP, can be modelled along the lines suggested by Knudson[10] for retinoblastoma by assuming that one of the mutations along the pathway to carcinogenesis has been inherited by every cell in the tissue of interest. The critical point here is that all sources of inter-individual variation in susceptibility can be modelled using specific biological considerations. It is not necessary to use the artifice of multiplying the hazard function of the Weibull model by a frailty parameter, as Aalen *et al*. suggest.[1]

I agree with the authors that ignoring heterogeneity and frailty can yield misleading inferences. That said, another equally important factor in the misinterpretation of epidemiological data is the ubiquitous and often inappropriate application of the proportional hazards model for analysis and the virtually universal use of the relative risk as a measure of effect. It is becoming increasingly clear that summary measures of exposure, such as cumulative exposure, cannot capture the impact of complex temporal patterns of exposure on disease risk,[5,6] and that the relative hazard, which is the target of estimation with the proportional hazards model, has serious limitations.[7,8] For cohort data, the use of parametric hazard functions derived from multistage models of carcinogenesis that explicitly incorporate patterns of exposure can simultaneously address both issues and provide insights that are difficult or impossible to obtain using the proportional hazards model.[9]

Conflict of interest: None declared.

References

1. Aalen O, Valberg M, Grotmol T, Tretli S. Understanding variation in disease risk: The elusive concept of frailty. *Int J Epidemiol* 2015;**44**:1408–21.
2. Luebeck EG, Moolgavkar SH. Multistage carcinogenesis and the incidence of colorectal cancer. *Proc Natl Acad Sci U S A* 2002; **99:**15095–100.
3. Moolgavkar S. The multistage theory of carcinogenesis and the age distribution of cancer in man. *J Natl Cancer Inst* 1978;**61:**49–52.
4. Meza R, Jeon J, Moolgavkar SH, Luebeck EG. The age-specific incidence of cancer: phases, transitions and biological implications. *Proc Natl Acad Sci U S A* 2008;**105:**16284–89.
5. Peto J. That the effects of smoking should be measured in pack-years: misconceptions 4. *Br J Cancer* 2012;**107:**406–07.
6. Thomas DC. Invited commentary: is it time to retire the "pack-years" variable? Maybe not! *Am J Epidemiol* 2014;**179:**299–302.
7. Poole C. On the origin of risk relativism. *Epidemiology* 2010;**21:** 3–9.
8. Hernan MA. The hazards of hazard ratios. *Epidemiology* 2010; **21:**13–15.
9. Moolgavkar SH, Chang ET, Luebeck G *et al*. Diesel engine exhaust and lung cancer mortality: Time-related factors in exposure and risk. *Risk Anal* 2015, Feb 13. doi: 10.1111/risa.12315. [Epub ahead of print.]
10. Knudson AG. Two genetic hits (more or less) to cancer. *Nat Rev Cancer* 2001;**1:**157–162.

Authors' response: Understanding variation in disease risk

International Journal of Epidemiology, 2015, 1426–1428
doi: 10.1093/ije/dyv047
Advance Access Publication Date: 4 April 2015

From Odd O Aalen,[1,2,*] Morten Valberg,[1] Tom Grotmol[2] and Steinar Tretli[2]

[1]Oslo Centre for Biostatistics and Epidemiology, Department of Biostatistics, Institute of Basic Medical Sciences, University of Oslo, Oslo, Norway and [2]Cancer Registry of Norway, Institute of Population-Based Cancer Research, Oslo, Norway

*Corresponding author. Department of Biostatistics, Institute of Basic Medical Sciences, University of Oslo, POB. 1122, Blindern, N-0317 Oslo, Norway. E-mail: o.o.aalen@medisin.uio.no

Accepted 11 March 2015

First, we thank the authors of the three commentaries[1,2,3] for their interesting discussion of our paper.[4]

Our intention was to shed light on how the concept of frailty can contribute to the understanding of variations in disease risk that are due to unobserved or unknown factors. We discuss how such differences may arise, and what the consequences of such variations are when studying populations. Although Peto's statement that frailty retreats when biology advances is certainly true to some extent,[1] we argue that there will always be some unexplained variation left, due to stochastic elements involved in the development of many diseases. Even if all relevant risk factors, be they environmental or (epi)genetic, could be identified, there would still be considerable variation in risk that is unexplained. In a recent paper, Tomasetti and Vogelstein suggest that 'only a third of the variation in cancer risk among tissues is attributable to environmental factors or inherited predispositions'.[5] They further argue that the majority of cases are 'bad luck' due to random mutations. Although it is possible to question the claim that two-thirds of cancer cases occur randomly, which might be an exaggeration, it still gives recognition to the importance of randomness in the development of cancer. Frailty models offer one way of handling this in epidemiological studies.

Using frailty models for analysing single time-to-event data may be quite speculative, and they should be based on sound biological knowledge to be used for any kind of mechanistic inferences. Whether the estimate of only 12% of the US population being susceptible to colon cancer, as suggested by Soto-Ortiz and Brody,[6] is reasonable or not, is of course a relevant question. Nevertheless, simple frailty models may still be quite useful in a hypothesis-generating way. Mathews and Hopper point out that frailty explanations and biological explanations are not 'competing', as we state in the paper, and they further claim that frailty must have a biological explanation.[2] The point we wanted to make is that there is a distinction between a mechanistic explanation of, say, a peak in the incidence rate and a selection explanation which is the frailty one. Of course, selection also has a biological basis but might be seen as competing with a mechanistic biological view. For instance, a mechanistic explanation of the peak in testicular cancer incidence at around 30 years of age could be declining testosterone level; however, there is no basis for this view and the competing selection, or frailty, explanation is likely to hold.

When times-to-events are related, the amount of speculation is reduced in a frailty model. A major topic in our paper is that of familial clustering of disease. One particular interest lies in what a so-called familial relative risk larger than one implies. Since such estimates compare the risk in individuals who have a certain familial history of the disease in question, with the average risk level in the population (or with the risk in individuals with a different familial history), it does not immediately say anything about how the risk is distributed across the population. An important implication, also appreciated in the commentaries, is that even moderate familial relative risks have to mean that there are potentially very large differences in risk between individuals in the population. Frailty models seem to be very suitable for analysing this kind of data, and are not only able to provide estimates of very detailed familial relative risks (given any kind of familial history), but can also provide information on how the underlying risk is distributed in the population.[7,8]

Moolgavkar argues that frailty and heterogeneity are two terms that should be kept apart; frailty should be reserved for situations where a subgroup of the population is exclusively at risk, or at vastly increased risk.[3] We fail to see the reasoning behind this statement. A population that contains two risk groups is heterogeneous, and a population with a continuous spectrum of underlying risk is also heterogeneous, even if the variation is modest. In our terminology, varying frailty between individuals expresses the heterogeneity in risk in a population, regardless of how it is distributed. That being said, situations where the frailty effects are most striking are perhaps those where the frailty distribution is much skewed.

We also provide a brief discussion of 'frailty and models of carcinogenesis', which Moolgavkar finds to be 'unclear.'[3] It is true that our wording was not the best when stating that the Armitage-Doll model is a 'sensible approximation to the carcinogenic process within an individual', when it actually gives an approximation to the hazard function in an initially homogeneous population. Moolgavkar states that: 'Inter-individual variations in susceptibility arising from variations in the rates of critical biological processes can be modelled by assuming a distribution on the parameters of the model'. This is exactly what we are suggesting. We randomize a parameter in the Weibull (i.e. Armitage-Doll) model, and let it be distributed over the entire population. The remaining criticism seems to be the use of the 'poor' Weibull approximation. However, as long as the probability of the event (cancer) occurring is small, the Weibull approximation will be good. Our point is that we are interested in studying selection effects in a population that is inherently heterogeneous, which can indeed be done by combining a Weibull model with the notion of a varying frailty. The resulting model is intuitively easy to understand, although there are more sophisticated models that provide a more accurate description of carcinogenesis itself.

Funding

This work was partially supported by a grant from the Norwegian Research Council (191460/V50), and by the Norwegian Cancer Society (project/grant number 4493570).

References

1. Peto J. Commentary: Frailty and Cancer. *Int J Epidemiol* 2015; **44**:1423–24.
2. Mathews JD, Hopper JL. Commentary: Age and frailty—not quite the same thing. *Int J Epidemiol* 2015;**44**:1421–23.
3. Moolgavkar SH. Commentary: Frailty and heterogeneity in epidemiologic studies. *Int J Epidemiol* 2015;**44**:1425–26.
4. Aalen OO, Valberg M, Grotmol T, Tretli S. Understanding variation in disease risk: the elusive concept of frailty. *Int J Epidemiol* 2015;**44**:1408–21.
5. Tomasetti C, Vogelstein B. Variation in cancer risk among tissues can be explained by the number of stem cell divisions. *Science* 2015;**347**:78–81.
6. Soto-Ortiz L, Brody JP. Similarities in the age-specific incidence of colon and testicular cancers. *PLoS One* 2013;**8**: e66694.
7. Moger TA, Aalen OO, Heimdal K, Gjessing HK. Analysis of testicular cancer data using a frailty model with familial dependence. *Stat Med* 2004;**23**:617–32.
8. Valberg M, Grotmol T, Tretli S, Veierød MB, Moger TA, Aalen OO. A hierarchical frailty model for familial testicular germ-cell tumors. *Am J Epidemiol* 2014;**179**:499–506.