The National Marrow Donor Program Donor Register (www.ashi-hla.org/resourcesfiles/resources-frequencies.html) provides estimates of the ABDR haplotype frequencies for three North American populations. Using the data on the Caucasian American population in which 9642 haplotypes were estimated to have non-zero frequencies, only 184 had estimated frequencies $/H110220.1\%$. The three highest estimated frequencies were 0.0518, 0.0263, and 0.0215 (Mori et al. 2).

Table 2 shows that the probability of a matched sibling pair in this population is 0.2531. With the large number of haplotypes and the small frequencies associated with each, most parental pairs of identical siblings will be heterozygous with different haplotypes in the two parents. All such parental pairs would give a probability of 0.25 of a matched sibling pair and phenotypic frequencies the same as in a random population. The additional probability of 0.0033 arises from other types of parental pairs. The frequencies of the three commonest haplotypes and the six corresponding phenotypes shown in Table 2 demonstrate that this small set of additional parental pairs results in very small differences between the frequencies in the general population and in patients with a matching sibling. Clearly the genetic bias is very small and can safely be ignored particularly remembering that, to be important, the difference in phenotype has to affect differentially the consequences of the two treatments. Corresponding results for the African American and Asian American populations were even closer to 0.25, 0.2508 and 0.2516, respectively.

The differences will be doubled if the patient has a sibling who does not match, but the difference will still be negligible. The effect on the comparison when the patient has more than one sibling with varying numbers of them matching their phenotype is more difficult to quantify but again unlikely to be important.

Clearly, the number of possible HLA haplotypes at the $A$, $B$, and $DR$ loci and their low frequencies result in the majority of parents of siblings with identical HLA phenotypes being heterozygous and sharing no haplotypes in common. This means that differences in the distribution of phenotypes of siblings with and without an HLA identical sibling are very small. They should be ignored and concentration should be continued on the non-genetic biases that may occur between the types of patient being used to compare SCT with chemotherapy.

Acknowledgement
I am grateful to Dr Linda Shelper of UK Transplant for providing access to the data used in this paper.

References

DOI:10.1093/ije/dyi161
Advance Access publication 8 August 2005

Origins of the mutational origin of cancer
From LUTZ EDLER* and ANNETTE KOPP-SCHNEIDER

The authors would like to congratulate the editors of the *International Journal of Epidemiology* for their insight into the important role of theoretical concepts of carcinogenesis for the understanding, prevention, and treatment of cancer, and for their initiative to refresh the concept of multi-stage theory of cancer through reprinting of the pioneering paper of P Armitage and R Doll1 from 1954. The commentaries of Steven A Frank, 2 Suresh H Moolgavkar, 3 and Sir Richard Doll 4 provided further insight into the importance of this theory for examining and explaining the change in cancer mortality with age. Since 1954 the multi-stage theory has become a necessary prerequisite for understanding cancer data and developing advanced concepts.
As mentioned by Steven Frank,2 ‘an extensive mathematical literature has refined this theory and fit various models to more detailed sets of data’. One may note in this historical view that ~20 years after the first period of mathematical carcinogenesis modelling a second wave of modelling appeared in the late 1970s, which was initiated by the seminal articles by Whittome and Keller3 and by Mooigail and Venzon.4 The review article by Whittome and Keller5 stimulated research in multistage models and remained for many years a primary reference of carcinogenesis modelling. Mooigail and Venzon6 formulated a two-stage model with stochastic clonal expansion of both normal and intermediate cells and introduced a mathematical technique to analyse a two-stage model with deterministic growth of normal cells and stochastic growth of intermediate cells. A series of articles followed in which the model was applied to incidence data from animal experiments and epidemiological studies.7,8 Although originally the model was proposed by Kendall9 it is often termed the Mooigail–Venzon–Knudson (MVK) model. The mathematical and biological background for this model has been reviewed more recently by Tan10 and Kopp-Schneider.11 The two-stage model was generalized to incorporate more than one intermediate cell stage12,13 or to describe more than one pathway to malignancy. Since the 1990s multistage models have been used not only to describe the incidence of malignant neoplasia, but also to analyse the number and dimensions of premalignant lesions.14–16 It was obviously the combination of the concept of carcinogenesis as a multi-stage process with real life epidemiological data that led to the success of the ideas of Peter Armitage and Sir Richard Doll and initiated a wealth of methodological approaches for cancer research leading to the present state of cancer modelling.

As pointed out in the commentary of Sir Richard Doll,4 a basic component of carcinogenesis modelling is ‘the idea that a cancer might arise from a mutation in the hereditary material of a somatic cell’. Sir Richard Doll cites a science paper17 from 1930 as the reference for mutational causation of cancer, but remarks that ‘this also has been suggested some 15 years before’, but he had forgotten by whom. This remark intrigued us to explore our library at the German Cancer Research Center (DKFZ) in Heidelberg. We found that the surgeon and oncologist Karl Heinrich Bauer who founded the DKFZ in 1964 had published in 1928 a book titled Mutation Theory of the Origin of Tumors—Transitions of Somatic Cells into Tumor Cells by Gene Changes (Mutationstheorie der Geschwulst-Entstehung—Übergang von Körperzellen in Geschwulstzellen durch Gen-Änderungen)18 in which he develops and justifies his mutation theory. Since the book is written in German we take the opportunity to shortly summarize Bauer’s review of the mutation theory of cancer.

His earliest reference is to D van Hansemann19 in 1897 followed by Th Boveri20 in 1914. Hansemann19 defined anaplasia of tumour cells in analogy to the change of animal races under domestication, but he did not include hereditary elements in his characterization of tumour cells. In contrast, Boveri20 supposed abnormal chromosomal aggregation during cell division as a causal origin of cancer. Later, motivated by an idea of Aichel,21 who postulated the aggregation of cells with qualitative abnormal distribution of chromosomes as the origin of cancer, Boveri22 explains the origin of a malignant tumour as a consequence of abnormal chromatin content in the cell. This was the first time that carcinogenesis was attributed to chromosomal change in cells.

In chronological order, the next references are two papers from 1921.23,24 Levy23 postulates the origin of a tumour cell from a mutation (i.e. an alteration of the hereditary content) of the cell, e.g. a cell of the epithelium or connective tissue, and refers to Boveri’s hypothesis as well as to Th H Morgan’s breeding investigations of Drosophila.25 The nature of mutation was, however, at this time rather nebulous and not associated with specific changes in the chromosomes but more generally envisaged as an erroneous combination of cells and their chromatin (‘Zell(kern)verschmelzung’). In order to explain the causes of cancer, Gade24 started his considerations with an analysis of the age distribution of cancer mortality in Norway between 1901 and 1915 (Table 1), which shows a decline in the oldest age group, seen many times afterwards. Gade points out in his paper read before the Medical Society of Kristiania on September 21, 1921 that there is ‘a considerable parallelism between mutation and the phenomenon accompanying the transformations of a normal body-cell into a tumour cell’ and that ‘the cancer cell will transfer the newly acquired qualities to their future generations’.24 Although we cannot be sure about other earlier events marking the mutation theory of cancer, this date is at least a remarkable one ~10 years earlier than the landmark paper17 cited by Sir Richard Doll.

The parallelism between mutations in biological inheritance and carcinogenesis was also mentioned by Lenz26 in 1921 and then later by Bauer27 himself as well as by Schwarz28 in 1923. Obviously, the beginning of the mutational theory of cancer was in the early 1920s. Bauer concludes his review on the state of the art of the theory of mutational origin of cancer by referring to four scientists: Boveri, Morgan, Muller and Stomps,29 but he attributes most of the beginnings of the mutation theory to Boveri and his work in 1914. Therefore, when Richard Doll assumes that the mutation hypothesis of McCombs and McCombs from 1930 had a predecessor 15 years before, it seems that he had Boveri in mind. However, a full developed mutation theory of cancer had to wait until 1928 when Bauer’s book appeared. Bauer’s work is comprehensive and addresses intracellular phenomena, characteristics of the tumour cell and its division and proliferation, metastasis and tumour growth. He draws far-reaching conclusions on the treatment of cancer still worth reading and reflecting on today. Notably, Bauer warns of an overestimation of the importance of exogenous factors and he is rather critical about the success of prevention, an interesting thought for present times in an era of cancer research where both prevention studies and gene

<table>
<thead>
<tr>
<th>Age</th>
<th>Yearly cancer deaths per 100 000 living</th>
</tr>
</thead>
<tbody>
<tr>
<td>1–20</td>
<td>0.024</td>
</tr>
<tr>
<td>20–30</td>
<td>0.34</td>
</tr>
<tr>
<td>30–40</td>
<td>2.27</td>
</tr>
<tr>
<td>40–50</td>
<td>9.55</td>
</tr>
<tr>
<td>50–60</td>
<td>23.73</td>
</tr>
<tr>
<td>60–70</td>
<td>45.80</td>
</tr>
<tr>
<td>70–80</td>
<td>67.30</td>
</tr>
<tr>
<td>80–90</td>
<td>55.60</td>
</tr>
<tr>
<td>90 and above</td>
<td>30.67</td>
</tr>
</tbody>
</table>
expression play a prominent role. Even more interesting is that K H Bauer came to his conclusions via a probabilistic argument.

Acknowledgement
This work was supported by the Collaborative Linkage Grant PST.CLG 979045 of the NATO Science Program ‘Stochastic Modeling of Cancer Detection and Survival’.

References

Epidemiology vs epidemiology: the case of oil exploitation in the Amazon basin of Ecuador

AK HURTIG¹,²* AND M SAN SEBASTIÁN¹,²

In a recent letter, Terracini raises several important issues concerning the development of environmental epidemiology in Latin America and adds to the debate around our research on the health effects of oil exploitation in the Amazon basin of Ecuador.¹ Terracini argues that the tension between the two often used perspectives on the mission of epidemiology, ‘scientific’ or ‘action-oriented’, ignores the importance of the context in which epidemiology carries out its mission.

¹ Umeå International School of Public Health, Public Health and Clinical Medicine, Umeå University, Umeå, Sweden.
² Instituto de Epidemiología y Salud Comunitaria ‘Manuel Amunáriz’, Coca, Umeå International School of Public Health, Public Health and Clinical Medicine, Umeå University, SE-901 85 Umeå, Sweden.
* Corresponding author. Umeå International School of Public Health, Public Health and Clinical Medicine, Umeå University, SE-901 85 Umeå, Sweden.
E-mail: anna-karin.hurtig@epiph.umu.se