Research Article

Using the Negative Exponential Model to Describe Changes in Risk of Smoking-Related Diseases following Changes in Exposure to Tobacco

Peter N. Lee, John Hamling, John Fry, and Barbara Forey

P N Lee Statistics and Computing Ltd., 17 Cedar Road, Sutton, Surrey SM2 5DA, UK

Correspondence should be addressed to Peter N. Lee; peterlee@pnlee.co.uk

Received 13 April 2015; Accepted 13 July 2015

1. Introduction

A huge literature relates smoking to the risk of various major diseases. Although most of the evidence concerns comparison of risk in never and current smokers, many studies show that quitters have intermediate risks. Until now, few researchers have attempted formal modelling of the time pattern of the decline in relative risk (RR) following quitting [1, 2] and then only based on data from individual studies. Recently, however, publications on ischaemic heart disease (IHD) [3], lung cancer [4], chronic obstructive pulmonary disease (COPD) [5], and stroke [6] have fitted the negative exponential model (NEM) to multiple data sets to describe quantitatively the pattern of decline in RR with increasing time quit.

While these publications [3–6] provide important information on the benefits of quitting smoking, they are limited by not more generally considering how RR changes with variation in exposure to tobacco products. Thus, apart from quitting, smokers may change the amount they smoke, switch to a reduced exposure product, or may restart smoking having previously quit. Indeed, their level of exposure may change a number of times during their lifetime, and it would be useful to be able to model how RR varies over time for complex smoking histories. Not only would such a model allow a more comprehensive description of the risks associated with differing lifetime patterns of smoking, it may also be useful when attempting to predict the health impacts of novel products.

This paper describes extensions to the NEM to allow modelling of complex smoking histories. Section 2 gives a formal definition of the NEM; Section 3 then briefly summarizes material from the published papers on quitting [3–6] concerning the methodology used, the goodness-of-fit of the data for the different diseases studied, and evidence on between-study heterogeneity. Section 4 describes a first
extension to the NEM which applies to reducing cigarette consumption or switching to a reduced exposure product. A recent review [7] provides estimates of the reduction in lung cancer risk following a reduction in cigarette consumption, and it is shown that the NEM extension fits these data well. Section 5 describes a further extension of the NEM to allow for multiple exposure changes and illustrates it by some examples. No published epidemiological data are available to verify this, and we suggest that its validity might be tested using large data sets which have recorded changes in smoking habits at multiple time points. The methods described up to this point assume that exposure is linearly related to the excess relative risk (ER = RR − 1) and that the rate of its decline with time is invariant of age. Section 6 describes adaptations of the NEM which allow for nonlinearity of the dose-response and for the rate of decline in ER to vary with age.

The methods described and justified here give fuller detail than those summarized recently by Weitkunat et al. [8] when outlining a novel approach to assess the population health impact of introducing a modified risk tobacco product (MRTP). In that approach two sets of simulated individual smoking histories are compared, one scenario being derived on the assumption that the MRTP is introduced into the population and the other scenario derived assuming that it is not. The histories are derived from assumptions concerning the extent to which smokers will start smoking the two types of product (conventional cigarettes or MRTPs), will switch from one product to another, or will quit. The RR of the major smoking-related diseases (compared to never smokers) is then estimated for each individual’s smoking history using the formulae described here, and average risks over individuals for each scenario are then used to estimate the reduction in deaths attributable to the assumed level of adoption of the MRTP.

2. The Negative Exponential Model

The NEM predicts that the smoking-associated ER declines with time quit, so that if ER(a) is the ER in current smokers compared to never smokers, where a is age in years, then ER(a, t), the ER in quitters of the same age, t years after quitting, is given by

\[ ER(a, t) = ER(a) \exp \left(-\frac{tL}{H}\right), \]

where \( L = \log_2 2 \) and \( H \) is the “half-life,” that is, the time since quitting when the ER declines to half that associated with continuing smoking.

3. Quitting Smoking

Studies which relate risk to time quit typically present RRs and 95% confidence intervals (CIs), relative to never smokers, for current smokers and for quitters by grouped period of time quit. Some studies present RRs and CIs for never smokers and quitters, expressed relative to current smokers. Typically, the RRs are adjusted for potential confounding variables. As described in detail elsewhere [3], fitting the NEM requires a pseudotable to be estimated (using the method of Hamling et al. [9]). This pseudotable consists of numbers of subjects subdivided by smoking habit and disease status which correspond to the observed RRs and 95% CIs. For prospective studies the numbers are of cases and at risk, and for case-control studies they are of cases and controls. Using these data and estimates of \( t \), the midpoint time quit (taken to be zero for current smokers and infinite for never smokers) and maximum likelihood methods then allow fitting of the model. For prospective studies, the model is

\[ P_j = A + B \exp \left(-Ct_j\right), \]

where \( P_j \) is the absolute risk of disease at time quit \( t_j \) and \( A, B, \) and \( C \) are parameters to be estimated. Here \( j \) references smoking group (current smokers, quitters by time quit, or never smokers), \( A \) is the absolute risk in never smokers, and \( A + B \) is the absolute risk in current smokers. \( C \) is equal to \( L/H \).

For case-control studies, the model differs as follows:

\[ P_j^* = 1 + B^* \exp \left(-Ct_j\right). \]

Here \( P_j^* \) is the RR (versus never smokers, estimated by the odds ratio) not the absolute risk \( (P_j) \), and \( B^* \) is the ER, rather than the increase in absolute risk \( (B) \), in current smokers. The interpretation of \( C \) is the same as for prospective studies.

Lee et al. [3] describe methods for testing goodness-of-fit, carrying out meta-analysis of estimates of \( H \), and testing heterogeneity. They also describe sensitivity analyses which attempt to account for “reverse causation” (short-term quitters having an increased risk as quitting can be prompted by symptoms from undiagnosed disease), either by omitting short-term quitters or by reclassifying them to be current smokers. The methodology has recently been applied to all available data sets for four major smoking-related diseases. The analysis for IHD [3] was based on 41 independent blocks of RRs from 23 studies [10–32], that for COPD [5] on 14 blocks from 11 studies [23, 33–42], and that for stroke [6] on 22 blocks from 13 studies [15, 19, 20, 23, 25, 37, 38, 43–48]. The analysis for lung cancer [4] was based on the most data, involving 106 blocks of RRs from 85 studies, 30 in the USA or Canada [23, 45, 49–76], 25 in Europe [10, 13, 30, 34, 77–97], and 30 from other or multiple countries [19, 37, 38, 72, 98–122]. (Note that one publication [38] provided separate data for two different Australian studies, while another [72] provided separate data for a US and a Japanese study.) Based on the combined results, various conclusions can be made, which are summarized below.

3.1. Failure to Fit. The method may fail to converge to a valid solution \( (H > 0) \) or may converge to a solution with a huge variance, when a decline in risk following quitting is not clearly evident. This most often occurs when the current smoker RR is not or only minimally elevated. Failure (leading to exclusion from further analysis) occurred for none of 106 lung cancer data sets, one of 41 IHD sets, where the current smoker RR was 1.0, and one of 14 COPD sets, where no risk decline following quitting was evident. Failure often occurred...
for stroke, where current smoker RRs were below 1.4. Here, our meta-analysis used only those II (of 22) data sets which showed a stronger relationship with smoking.

3.2. Goodness-of-Fit. Where convergence did occur, goodness-of-fit, as assessed by comparison of observed and fitted numbers, was adequate for every COPD and stroke data set and for nearly all IHD sets, except where the decline in risk was clearly nonmonotonic. For lung cancer, model fit was poor when reverse causation was ignored but was much improved in sensitivity analyses omitting short-term quitters or counting them as current smokers. However, some misfit remained, seemingly resulting from odd data patterns in individual studies rather than from any systematic misfit.

3.3. Estimates of Half-Life. Meta-analysis H estimates (years) were larger for COPD, 13.32 (95% CI 11.86–14.96), and lung cancer, 9.93 (9.31–10.60), than for IHD, 4.40 (3.26–5.95), or stroke, 4.78 (2.17–10.50). It has been shown more recently [123] that the half-life is somewhat greater for adenocarcinoma than for squamous cell carcinoma of the lung, with the ratio of the Hs for the two cell types estimated as 1.32 (1.20–1.46).

These estimates provide more precise information on the pattern of the decline in ER following quitting than that found in various authoritative reviews, for example, [15, 124–126]. Thus, for example, the International Agency on Research and Cancer (IARC) [126] notes for IHD that, though the data were heterogeneous, “the body of evidence points toward the risk of CHD asymptotically approaching the risk of never smokers.” While imprecise, this conclusion seems correct. The same cannot be said of an earlier report, by the US Surgeon-General [15] which stated that “the excess risk of CHD caused by smoking is reduced by about half after 1 year of smoking abstinence and then declines gradually. After 15 years of abstinence, the risk of CHD is similar to that of persons who have never smoked.” Here the indicated decline in ER following quitting is much more rapid than indicated by the overall data currently available.

For lung cancer, the US Surgeon-General [15] stated that “after 10 years of abstinence, the risk of lung cancer is about 30–50% of the risk for continuing smokers; with further abstinence, the risk continues to decline.” While not incorrect, the statement does not fully quantify the decline in ER with time.

As far as we are aware, there have been no attempts, for any of the four diseases considered, to formally quantify the shape of the decline in risk following quitting based on all the available epidemiological evidence.

3.4. Sensitivity Analyses. Although improving the fit, the sensitivity analyses had little effect on the meta-analysis estimates of H. This was unsurprising as, for each data set, the estimated H was usually close to where the ratio of ERs in quitters and continuing smokers reduced to below 50%. This would occur longer after quitting than when the short-term increases were seen.

3.5. Heterogeneity. Heterogeneity between individual H estimates for a disease was assessed by likelihood-ratio tests. Significant ($p < 0.05$) heterogeneity was not seen for COPD but was seen for the other diseases, markedly so for IHD (where $H$ was estimated as <2 years for 10 of 40 data sets and as >10 years for 12) and for stroke. For lung cancer, $H$ was somewhat higher in males and in older populations. For IHD and stroke, the relationship of $H$ to age was also evident.

For all four diseases, the NEM was found to provide a simple method for summarizing quitting data, generally fitting it well. Misfits seen were mainly due to unusual results, which no plausible model could be expected to fit, rather than to any systematic misfits. However, the evidence of a greater $H$ in older populations, seen for three of the four diseases studied, suggests that some adaptation of the simple NEM may be required to allow for this. This is described later.

4. Reducing Cigarette Consumption

Quitting can be regarded as reducing the effective exposure from 1 to 0 units. The NEM can be extended, assuming that the ER is linearly related to exposure to estimate the ER, $\text{ER}_R(a, t)$, $t$ years after a reduction in exposure from 1 to $F$ units, by

$$\text{ER}_R(a, t) = \text{ER}_S(a) \left( F + (1 - F) \exp \left( -\frac{tL}{H} \right) \right). \quad (4)$$

Here $t$ is the time since the reduction occurred. When $F = 0$, formulae (1) and (4) are identical. Formula (4) can apply, not only to reducing cigarettes/day from, say, 25 to 15 ($F = 0.6$), but also to switching to a product with a reduced uptake of relevant smoke constituents. Note that the formula implies a linear dose-response relationship. Modification of the formula to allow for nonlinearity is discussed later (see Section 6.1).

A recent review on the effect on lung cancer risk of reducing amount smoked [7] summarized evidence from three prospective and three case-control studies which involved varying reductions in amount smoked over a varying time. Each study demonstrated a risk reduction, with a meta-analysis estimating the overall RR for reducers versus nonreducers as 0.81 (95% CI 0.74–0.88).

To compare these results with the predictions of formula (4), data are required on the extent of reduction of consumption in the reducers, the length of follow-up since the reduction, and the average age of the subjects at the time of reduction. Such data were only available for the prospective studies, each of which involved two separate examinations when smoking habits were recorded and then a mortality follow-up period of 10 or more years. The study details and the observed and NEM predicted RRs are shown in Table 1.

Study 1 provided one RR estimate for sexes combined, study 2 provided one estimate per sex, and study 3 provided an estimate for males for different degrees of reduction. Each RR was adjusted for age and confounders. $F$, the ratio of consumption of post- to prereduction, varied from 0.17 to 0.50. $F$ values came from the source for studies 1 and 2.
Table 1: Summary of lung cancer results for reducing amount smoked from three prospective studies.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Study 1 [148]</th>
<th>Study 2 [149]</th>
<th>Study 3 [150]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location</td>
<td>Denmark</td>
<td>Norway</td>
<td>Korea</td>
</tr>
<tr>
<td>Time between the two examinations</td>
<td>5 to 10</td>
<td>3 to 13</td>
<td>2</td>
</tr>
<tr>
<td>Follow-up period from examination</td>
<td>18</td>
<td>21.2</td>
<td>10.7</td>
</tr>
<tr>
<td>Definition of a continuing smoker</td>
<td>Heavy smoker (at least 15 cigs/day)</td>
<td>Heavy smoker (at least 15 cigs/day)</td>
<td>At each exam subjects were described as heavy (20+, N 57730), moderate (10–19, N 101347), or light (&lt;10 cigs/day, N 76754)</td>
</tr>
<tr>
<td>Definition of a reducing smoker</td>
<td>(1) Heavy smoker at first exam reducing by at least 50% at the second exam (N 832)</td>
<td>(2) As for study 1, males (N 348)</td>
<td>(4) Heavy smoker at first exam reducing to moderate smoker at second exam (N 15936)</td>
</tr>
<tr>
<td></td>
<td>(3) As for study 1, females (N 127)</td>
<td>(5) Heavy to light (N 4293)</td>
<td>(6) Moderate to light (N 19168)</td>
</tr>
<tr>
<td>Ratio of average consumption at the</td>
<td>(1) 8.5/20.2 = 0.42</td>
<td>(2) 9.7/22.9 = 0.42</td>
<td>(4) 0.50</td>
</tr>
<tr>
<td>two exams in reducers (F)</td>
<td>(3) 8.6/19.5 = 0.44</td>
<td>(5) 0.17</td>
<td>(6) 0.33</td>
</tr>
<tr>
<td>Time from reduction to end of follow-up (years)</td>
<td>21.4</td>
<td>24.0</td>
<td>11.7</td>
</tr>
<tr>
<td>Cases in reducers</td>
<td>(1) 52</td>
<td>(2) 10</td>
<td>(4) 63</td>
</tr>
<tr>
<td></td>
<td>(3) 2</td>
<td>(5) 95</td>
<td>(6) 53</td>
</tr>
<tr>
<td>Average age at the time of reduction</td>
<td>(1) 49.0</td>
<td>(2) 39.0</td>
<td>(4) 42.6</td>
</tr>
<tr>
<td>((\bar{\alpha}))</td>
<td>(3) 38.2</td>
<td>(5) 42.9</td>
<td>(6) 42.4</td>
</tr>
<tr>
<td>Observed RRs (95% CIs)(^a)</td>
<td>(1) 0.73 (0.54–0.98)</td>
<td>(2) 0.71 (0.36–1.39)</td>
<td>(4) 0.72 (0.59–0.89)</td>
</tr>
<tr>
<td></td>
<td>(3) 0.51 (0.12–2.08)</td>
<td>(5) 0.63 (0.46–0.86)</td>
<td>(6) 0.79 (0.64–0.98)</td>
</tr>
<tr>
<td>NEM predicted RRs</td>
<td>(1) 0.64</td>
<td>(2) 0.60</td>
<td>(4) 0.80</td>
</tr>
<tr>
<td></td>
<td>(3) 0.61</td>
<td>(5) 0.67</td>
<td>(6) 0.73</td>
</tr>
</tbody>
</table>

\(^a\)Estimated as the time of follow-up plus half the time between the two examinations.

\(^b\)Reducers relative to continuing smokers.

but for study 3 were based on assumed midpoints of 30, 15, and 5 cigs/day for the three categories used. The time from reduction to end of follow-up was estimated as length of follow-up plus half the time between examinations.

RRs were predicted using formula (4), as \(F + (1 - F)V\), where \(V\) is a weighted mean value of \(\exp(-tL/H)\) over the time from reduction to end of follow-up, the values being computed at 0.1 year intervals. Based on the Doll-Peto formula [127] the weighting factor was taken as \((\bar{\alpha} - 22.5 + t)^{1.5}\), \(\bar{\alpha}\) being the average age at time of reduction and \(t\) being the time since the reduction occurred. This factor is necessary because the observed RR derives from cancers occurring at varying times after the reduction started, whereas formula (4) relates to the reduction in ER at a specific time. It also accounts for the absolute risk in continuing smokers rising with increasing age (or duration). \(H\) was taken as 9.93, based on earlier work [4]. No account was taken of any survival differences between continuing smokers and reducers, likely to have relatively minor effect given the ages of the populations.

As seen in Table 1, the predicted RR is always near its observed value and lies well within its 95% CI. For the six observed RRs, meta-analysis yields an estimate of 0.73 (95% CI 0.64–0.82) with no heterogeneity. Using weights as for the observed data, the estimate based on the NEM predicted
RRs, 0.71 (0.63–0.80), was similar. Although various assumptions and simplifications are involved, unavoidable without access to the full study data, and study 3 estimates are not independent (estimates 4 and 5 both involving comparisons with continuing heavy smokers), the predicted RR from formula (4) is clearly a good approximation to the observed value.

5. Multiple Changes in Exposure

Denoting the negative exponential function \( \exp(-tL/H) \) by \( N(t) \), formula (4) can be rewritten as

\[
ER_k(a, t) = N(t) ER_0(a) + (1 - N(t)) ER(a, t) \quad (5)
\]

where \( t \) is the time since switch. Setting \( F_1 = 1 \) and \( F_2 = F \), the two formulae are clearly the same.

Figure 1 illustrates various patterns of observed \( ER_5 \). The five lines in red relate to smokers of conventional cigarettes \( (F = 1) \) who, at the age of 40, either continue to smoke conventional cigarettes, switch to differing MRTPs \((F = 0.5, 0.75, 0.25)\), or quit \((F = 0)\). The blue lines are for smokers of modified products \((F = 0.5, 5)\) who, at the age of 40, either continue smoking these, switch to conventional cigarettes or to other modified risk products \((F = 0.75, 0.25)\), or quit \((F = 0)\). \( H \) is set at 10 throughout.

This formulation can be extended to allow for multiple periods of exposure. We first define the following:

- \( G + 1 \): number of periods \((g = 0, 1, \ldots, G)\),
- \( Z_g \): the time of the \( g \)th exposure change,
- \( F_g \): the exposure in period \( g \), with \( F_0 \) taken as zero,
- \( X(a) \): the ER for a continuing smoker with \( F = 1 \), at age \( a \), the value being derived from epidemiological data,
- \( ER(k, a, F_1, F_2, \ldots, F_k) \): the ER in period \( k \) at age \( a \) for the given pattern of exposure. Note that ER is only defined for \( a > Z_k \).

The method of estimation of \( ER(k, a, F_1, F_2, \ldots, F_k) \) is illustrated below for four switches \((G = 4)\) working through the five periods in turn.

In period 0, \( ER(0, a) \) is zero.

In period 1, the ERs are estimated by multiplying the exposure by \( X(a) \). To allow calculation for later periods, this must be estimated for each later exposure. Thus, we have

\[
ER(1, a, F_1) = F_1 X(a), \\
ER(1, a, F_2) = F_2 X(a), \\
ER(1, a, F_3) = F_3 X(a), \\
ER(1, a, F_4) = F_4 X(a). \\ (7)
\]

In period 2, the ERs are calculated from period 1 ERs as follows:

\[
ER(2, a, F_1, F_2) = N(a - Z_2) ER(1, a, F_1) + (1 - N(a - Z_2)) ER(1, a, F_2),
\]

\[
ER(2, a, F_1, F_3) = N(a - Z_2) ER(1, a, F_1) + (1 - N(a - Z_2)) ER(1, a, F_3), \\
\] (8)

\[
ER(2, a, F_1, F_4) = N(a - Z_2) ER(1, a, F_1) + (1 - N(a - Z_2)) ER(1, a, F_4). \\
\]

In period 3, the ERs are calculated from period 2 ERs as follows:

\[
ER(3, a, F_1, F_2, F_3) = N(a - Z_3) ER(2, a, F_1, F_2) + (1 - N(a - Z_3)) ER(2, a, F_1, F_3),
\]

\[
ER(3, a, F_1, F_2, F_4) = N(a - Z_3) ER(2, a, F_1, F_2) + (1 - N(a - Z_3)) ER(2, a, F_1, F_4). \\
\] (9)

Similarly in period 4, we have

\[
ER(4, a, F_1, F_2, F_3, F_4) = N(a - Z_4) ER(3, a, F_1, F_2, F_3) + (1 - N(a - Z_4)) ER(3, a, F_1, F_2, F_4). \\
\] (10)
Thus the person of interest, with exposures $0, F_1, F_2, F_3,$ and $F_4$, has the ERs in each period given by the first equation shown for each period. The other ERs are used only to calculate these.

Figures 2, 3, and 4 show illustrative results for three smoking patterns, each starting with smoking of conventional cigarettes ($F = 1$) up to the age of 40 when the ER is taken as 10 but involving differing patterns later. $H$ is set as 10, as for Figure 1. Figure 2 concerns smokers who quit at the age of 40, resume at the age of 50 ($F = 1$), quit again at the age of 60, and then resume at a lower consumption at the age of 70 ($F = 0.5$). The ER is taken to be 10 for $F = 1$ at the age of 40. $H$ is set at 10 years.

Figure 3 concerns smokers who reduce consumption slightly at the age of 40 ($F = 0.75$), further reduce consumption at the age of 50 ($F = 0.5$), and then return to their original smoking habits at the age of 60 ($F = 1$), and then quit at the age of 70.

Note that, as far as we are aware, there are no published epidemiological data available giving changes in risk following multiple periods of exposure which would allow formal comparisons to be made of observed risks and those predicted by the formulae given above.

### 6. Modifications of the NEM

#### 6.1. Allowing for Variation in the Dose-Response Relationship

The methods described in Sections 4 and 5 assume that exposure is linearly related to the ER. Although recent dose-response meta-analyses [128] show little upward curvature in the relationship between ER and the amount smoked per day, a nonlinear dose-response has been claimed, with Doll and Peto [127] suggesting that risk is proportional to the $(cigs/day + 6)$ squared.

Thus, suppose that $F = 1$ corresponds to 20 cigarettes/day, so the ER is proportional to $26^2 - 6^2 = 640$, where $26^2$ is the dose factor in the Doll/Peto formula for smokers and $6^2$ is that for nonsmokers. Adjusted $F$ values, $F^*$, are then calculated by $F^* = ((20F + 6)^2 - 6^2)/640$, so that multiplying the ER for $F = 1$ by $F^*$ gives the quadratically interpolated curves. Thus unadjusted $F$ values of 0.75, 0.5, 0.25, and 0.1 become adjusted $F^*$ values of 0.633, 0.344, 0.133, and 0.044.

#### 6.2. Allowing for Variation in the Half-Life

As shown in Table 2 there is evidence that $H$ increases with age for lung
Table 2: Variation in $H$ by age\textsuperscript{a} for four smoking-related diseases as estimated in four recent publications.

<table>
<thead>
<tr>
<th>Disease</th>
<th>Age &lt; 50</th>
<th>Age 50–59</th>
<th>Age 60–69</th>
<th>Age 70–79</th>
<th>$p$\textsuperscript{b}</th>
</tr>
</thead>
<tbody>
<tr>
<td>IHD [3]</td>
<td>1.47 (0.91–2.39)</td>
<td>5.22 (3.79–7.20)</td>
<td>7.48 (4.56–12.28)</td>
<td>13.77 (4.27–44.35)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Stroke [6]</td>
<td>1.98 (0.41–9.65)</td>
<td>3.47 (0.79–15.19)</td>
<td>7.71 (2.12–27.97)</td>
<td>17.74 (1.46–215.7)</td>
<td>0.34</td>
</tr>
</tbody>
</table>

\textsuperscript{a}For prospective studies taken as age at baseline for IHD, COPD, and stroke but as age at midpoint of follow-up for lung cancer.

\textsuperscript{b}Heterogeneity between groups.

7. Discussion

Based on substantial data for IHD [3] and lung cancer [4] and more limited data for COPD [5] and for stroke [6], the decline in ER following quitting has been shown to fit well the simplest NEM version (formula (1)). Using more limited data, the decline in lung cancer risk following reducing cigarette consumption can be fitted well by an extended form (formula (4)).

The further extension (formula (6)) allows estimation of ER following multiple exposure changes. All it requires are estimates of $H$, the ER for continued smoking ($F = 1$), and the exposures ($F$) in each relevant period. The formula allows estimation of predicted risk patterns, not only for intermittent smoking periods and consumption changes, but also for changes in product smoked. Thus if conventional cigarette smoking has factor $F = 1$, and a MRTP [129] reduces exposure to relevant smoke constituents by 80%, $F$ would be set as 0.2. $F$ for dual users, using half conventional cigarettes and half MRTPs, could be set as 0.5 ($1 + 0.2$) = 0.6. Given data on switching rates from conventional cigarettes to MRTP and the frequency of dual use, the NEM can then predict changes in risk for smoking-related diseases.

Validating formula (6) is problematic, since published suitable epidemiological data on risk changes following multiple exposure changes are lacking. It is likely that, in fact, large epidemiological data sets exist which have recorded data on the extent of smoking at three or more time points. Providing that there are sufficient numbers of subjects showing differing patterns of changing exposure, it would be valuable to test the accuracy of fit to the NEM.

Another assumption inherent in formula (6) is that $H$ is invariant of age. Available results (see Table 2) suggest...
that $H$ increases with age for lung cancer, IHD, and stroke, though not for COPD. There is therefore a case for using age-dependent estimates of $H$, and we describe an adaptation of the NEM to allow for this. However, it may be difficult to obtain precise estimates. Those in Table 2 come from case-control and prospective studies. For case-control studies, age is when the cancer occurred, but for prospective studies it was taken as age at baseline (or for lung cancer, midpoint age of follow-up); such studies do not usually report risk by actual age.

Another issue is uncertainty in $H$, particularly for IHD and stroke, where substantial heterogeneity from different available data sets \([3, 6]\) renders the overall estimate less reliable. This creates uncertainty in ER estimates resulting from varying exposure histories, which might be addressed by presenting estimates using alternative $H$ values (e.g., $±1$ SE).

This paper does not attempt to further investigate the usefulness of the NEM by comparing its predictions with other models. There are two reasons for this. Firstly, the fit to more detailed models has only been studied for some of the diseases of interest. Second, for lung cancer, where various models have been tried, such as the multistage model (see e.g., \([1, 133–143]\)) or the two-stage clonal expansion model \([144–147]\), full comparisons of the predictions of the NEM and other models would be a major exercise, beyond the scope of this paper. We prefer, if possible, to test the goodness-of-fit of the NEM using available data sets. At the present time, we feel that the material presented suggests that the NEM is a valuable, simple tool meriting further investigation.

8. Conclusion

The NEM is a useful model with many applications to smoking and health data. The most general form not only allows a simple description of the time course of disease following quitting, but can predict changes in risk following single or multiple reductions or increases in exposure, whether resulting from changes in amount smoked or the type of product smoked, perhaps following switching to MRTPs.

Disclosure

Neither sponsor was involved in the planning, execution, or writing of the paper or the decision to submit it for publication.

Disclaimer

The opinions and conclusions of the authors are their own and do not necessarily reflect the position of Philip Morris Products S.A.

Conflict of Interests

Peter N. Lee, founder of P. N. Lee Statistics and Computing Ltd., is an independent consultant in statistics and an advisor in the fields of epidemiology and toxicology to a number of tobacco, pharmaceutical, and chemical companies. This includes Philip Morris Products S.A., the sponsor of this study. The other three authors are employees of P. N. Lee Statistics and Computing Ltd.

Acknowledgments

The authors thank Philip Morris Products S.A. that funded the work. They also thank Pauline Wassell, Diana Morris, and Yvonne Cooper for assistance in typing the various drafts of the paper and obtaining the relevant literature. The preparation of the paper and work on quitting smoking and multiple exposure changes was supported by Philip Morris Products S.A. The work on cigarette consumption reduction was supported by Altria Client Services Inc.

References


Advances in Epidemiology 13


