Modeling age-dependent radiation-induced second cancer risks and estimation of mutation rate: an evolutionary approach

Kamran Kaveh · Venkata S. K. Manem · Mohammad Kohandel · Siv Sivaloganathan

Abstract Although the survival rate of cancer patients has significantly increased due to advances in anti-cancer therapeutics, one of the major side effects of these therapies, particularly radiotherapy, is the potential manifestation of radiation-induced secondary malignancies. In this work, a novel evolutionary stochastic model is introduced that couples short-term formalism (during radiotherapy) and long-term formalism (post-treatment). This framework is used to estimate the risks of second cancer as a function of spontaneous background and radiation-induced mutation rates of normal and pre-malignant cells. By fitting the model to available clinical data for spontaneous background risk together with data of Hodgkin’s lymphoma survivors (for various organs), the second cancer mutation rate is estimated. The model predicts a significant increase in mutation rate for some cancer types, which may be a sign of genomic instability. Finally, it is shown that the model results are in agreement with the measured results for excess relative risk (ERR) as a function of exposure age and that the model predicts a negative correlation of ERR with increase in attained age. This novel approach can be used to analyze several radiotherapy protocols in current clinical practice and to forecast the second cancer risks over time for individual patients.

Keywords Radiation-induced cancer · Second cancer · Excess relative risk (ERR) · Mutation rate · Evolutionary dynamics · Genomic instability

Introduction

The evolution of radiation medicine coupled with various technological advances, in therapeutic radiation treatments has dramatically improved the survival rate of cancer patients over the last several decades. Current treatment techniques deliver doses to the gross tumor volume with great precision. However, this comes at the cost of damaging healthy tissue in the vicinity of the tumor, leading to the possible manifestation of a secondary tumor post-treatment in many cancer survivors, which may occur sometimes several decades after treatment (Xu et al. 2008). Due to the higher survival rate of cancer patients in recent years, the investigation of secondary malignancy risks due to radiation therapy has now become of greater importance than ever. The fact that the survival rate of patients (for example Hodgkin’s lymphoma, leukemia and sarcomas) has improved to several more decades suggests that one needs to construct a model that can predict the age and time-dependent risks of second cancers. Radiation-induced cancer risks have been well documented in atomic bomb survivors mortality studies (Preston et al. 2007; Little 2009). There is also evidence from epidemiological studies that the mortality rate of Hodgkin’s lymphoma survivors is high post-radiation treatment (Alman et al. 2003; Bhatia et al. 2003). Several cohort and case-control studies have indicated that there is an increased risk of second malignancies with young Hodgkin’s lymphoma survivors post-irradiation (Bhatia and Sklar 2002; Hodgson et al. 2007b). Recently, several
clinical and modeling studies have described the risk of second malignancies for various neighboring organs (Sachs and Brenner 2005; Xu et al. 2008; Moteabbed et al. 2014; Paganetti et al. 2012; Athar and Paganetti 2011; Hall 2006; Manem et al. 2014a). Some of these clinical studies have suggested that there is a high incidence rate of breast cancer among young Hodgkin’s lymphoma survivors compared to women treated in middle age (Travis et al. 2005). Children treated with radiotherapy are particularly vulnerable to radiation-induced carcinogenesis (Bhatia and Sklar 2002). Several clinical findings indicate that second cancer risks change with respect to age at exposure and also with respect to time since exposure (Travis et al. 2005; Dores et al. 2002). Some authors studied the efficacy of dose escalation on secondary cancer induction (Manem et al. 2014b; Schneider et al. 2007). It has also been noted that the latency period is typically between 10 and 20 years for solid tumors, and around 5 years for leukemia post-irradiation (Zhang et al. 2012; Wang et al. 2014; Schneider et al. 2014; van Leeuwen et al. 2003; Yeoh and Mikhaeel 2011; Yahalom 2009; Hodgson et al. 2007a). As a result of increasing prevalence of secondary malignancies, it is clearly of great importance to investigate the risks associated with different treatment regimens as well as their possible variation with age (by incorporating age at exposure and time since exposure, Hodgson et al. 2007b). Site-specific dose-response relationships for cancer induction from the combined Japanese A-bomb and Hodgkin cohorts for doses on a par with radiation bomb survivors have been actively studied in the literature (Preston et al. 1994, 2007). The approaches taken are phenomenological models assuming the temporal pattern of risk to be an exponential power law (Preston et al. 2007). With respect to the baseline rates, (Preston et al. 2007) showed that these are proportional to the attained age raised to the power of 5.5, and decrease around the age of 90. Another point to take into consideration is that gender appears to play a major role in defining the baseline rates. This type of analysis of cancer incidence among A-bomb survivors indicates that the cancer risk (excess relative risk) is a function of time and decreases with age, thus leading to the conclusion that, for long-term age-dependent models, both exposure age and attained age are important factors to incorporate. At the same time, however, age-dependent factors are difficult to model in practice. Radiation-induced second cancers are thought to arise as a result of multiple mutations of healthy cells (or pre-cancerous, pre-malignant cells) post-irradiation of the primary tumor (Sachs and Brenner 2005; Lindsay et al. 2001; Schneider 2009). Pre-malignant cells are those cells that have the potential to turn into malignant cells. The process is very complex to model due to variations in various biological mechanisms in patients (that depend on habits, or environmental exposure). As discussed in the literature on primary cancer age incidence modeling (Armitage 1985; Armitage and Doll 1954), the manifestation of a second cancer is considered here to be the result of a two-hit process. It is assumed that anti-cancer therapies, such as radiation treatment, introduce a first hit to some of the normal cells in the vicinity of the tumor. This introduces a sub-population of radiation-induced pre-malignant cells in the system (see Sachs and Brenner 2005 and references therein). The detailed bio-physical mechanisms of pre-malignant cells resulting from radiation exposure are still not well understood. Additionally, radiation therapy has a destructive effect on the tissue architecture, leading to some changes in the spatial structure of the tissue (see for example Yorke et al. 1993). It has also been suggested that radiation therapy can introduce chromosomal instabilities and thus make the second hit mutations more effective and unpredictable in nature (Huang et al. 2003). In the presence of genomic instability, the possibility of having a hit on a particular locus is broadened into a range of hits and changes across a pair of chromosomes. However, there is no direct evidence for genomic instability as a result of radiation exposure, and there are several ongoing debates on some intriguing questions related to the conditions required for genomic instability (Bhatia and Sklar 2002; Sigurdson and Jones 2003). On the other hand, the progression of cancer and evolutionary dynamics of two-stage models of cancer progression has been discussed in the existing literature. The latter were developed to quantify the dynamics of how an advantageous mutant (pre-malignant cell) can outgrow a background population of normal stem cells and upon a second (malignant) mutation resulting in a malignant niche and eventually resulting to a secondary malignancy. These models have also been successfully applied to describe dynamics of chromosomal instability by different authors (Komarova et al. 2003; Nowak et al. 2002). It has also been proposed that second cancers may arise due to the creation of residual pre-malignant cells (cells due to radiation treatment) which can proliferate in between therapies and eventually grow afterward (Sachs and Brenner 2005). In this paper, a mathematical model is constructed to predict the excess relative risk (ERR) of second cancers due to radiation exposure during treatment of a primary malignancy. The present framework captures the stochastic nature of such complex evolutionary systems. The parameters of the model are as follows: (1) the mutation rate after radiation exposure which is assumed to be different from the background value of the mutation rate and (2) the proliferation strength of the pre-malignant cells. The dependence of the ERR on attained age and age at exposure of the patients is predicted by the model.
Materials and methods

In this section, the mathematical model used to estimate the ERR due to radiation exposure is briefly discussed. The mathematical details of the model are presented in “Appendix 1”. The framework used is built on the evolutionary model of cancer initiation/progression of two-hit mutation models of tumor suppressor gene inactivation (TSG), and the evolutionary dynamics of chromosomal instability developed by Komarova et al. (2003) and Nowak et al. (2002).

In modeling, the risks of secondary malignancies post-irradiation, a new time scale (age at exposure \( t_s \)) is introduced to capture the effect of radiation-induced pre-malignant cell initiation. It is assumed that radiation-induced secondary cancers are initiated by mutated stem cells. In a two-hit mechanism, it is further assumed that normal stem cells can be mutated into pre-malignant cells (also known as type-1, or one-hit cells). This random event occurs either by spontaneous mutation or as a result of a radiation-induced initiation mechanism during treatment. Each pre-malignant cell can then undergo a second spontaneous mutation and transform into a malignant phenotype (also known as a type-2, or two-hit cell). Malignant cells can eventually initiate tumor clones. The two-hit hypothesis here is a reflection of the fact that one is focusing on the malignancies induced by tumor suppressor gene inactivation (TSG). In the case of TSG inactivation, a single mutation does not lead to loss-of-function for a cell, due to which it is assumed that pre-malignant cells have a very weak selection advantage/disadvantage over normal stem cells; while malignant cells are expected to have a significantly higher proliferation potential.

The premise is that stem cells populate stem cell niche areas in the tissue which are homeostatically regulated. The proliferation/selection of the three types of stem cells—normal, pre-malignant, malignant—determine the phenotype that will eventually take over the corresponding niche. It is assumed that the selection mechanism is governed by a stochastic birth–death model, (i.e., Moran 1962; Durrett 2008). The form of the growth term of advantaged phenotypes (normally type-2 and also marginally type-1 cells) is reminiscent of logistic growth with the carrying capacity being the niche size. It is assumed that the total number of stem cells (either normal, pre-malignant, or malignant phenotypes) is a constant, \( N \) (within each niche). This number can vary between 10 and \( 10^6 \) (depending on the organ type). The total number of the niches is denoted by \( N_0 \), which is generally a large number (\( \sim 10^7 \)). The first and second spontaneous (background) mutation rates are denoted by \( u_1 \) and \( u_2 \). The second mutation rate for the irradiated (normal) tissues is distinguished by \( u_r \) (rather than \( u_2 \)) to indicate the secondary effects of radiation on pre-malignant cells. One of the probable secondary effects of ionizing radiation may be radiation-induced genomic instability which has been suggested in the literature (Siggurdson and Jones 2003; Komarova et al. 2003; Nowak et al. 2002). Moreover, the number of malignant stem cells created by radiation is ignored due to their small number relative to the pre-malignant cells.

The effects of ionizing radiation in the proposed framework can be summarized as follows:

- Radiation transforms normal cells into pre-malignant cells. These additional pre-malignant cells are distributed among the niches. As stated previously, the contribution of direct two-hit transformations due to radiation is ignored as the event is assumed to be a rare process.
- There is an effective second hit mutation rate, \( u_r \) for all the pre-malignant cells that is different from \( u_2 \). The approximate value of \( u_r \) can be estimated from the clinical data for cancer risk estimates.
- Due to the small numbers of pre-malignant cells relative to the total number of niche areas, only a small subgroup of niches will contain pre-malignant cells after treatment.

Figure 1 depicts a caricature of a two-hit process for background cancer and radiation-induced cancer. The dynamics of transformation from normal to pre-malignant tissue is displayed in Fig. 2.

In the absence of radiation, the probability of having at least one malignant cell (resulting through the sequential mutations \( u_1, u_2 \)) is calculated in a niche before time \( t \), i.e., attained age, denoted by \( f_0(t, u_1, u_2, r, N) \). For large niche sizes (typically \( N \geq 10 \)), a branching process approach is

\[ A^\text{(+t)} \rightarrow A^\text{(+r)} \rightarrow A^\text{(+r)} \]

\[ A^\text{(+t)} \rightarrow A^\text{(+t)} \rightarrow A^\text{(+t)} \]

\[ A^\text{(+t)} \rightarrow A^\text{(+t)} \rightarrow A^\text{(+t)} \]

\[ \gamma \rightarrow A^\text{(+t)} \rightarrow A^\text{(+t)} \]

\[ \gamma \rightarrow A^\text{(+t)} \rightarrow A^\text{(+t)} \]

\[ \gamma \rightarrow A^\text{(+t)} \rightarrow A^\text{(+t)} \]

\[ \gamma \rightarrow A^\text{(+t)} \rightarrow A^\text{(+t)} \]

\[ \gamma \rightarrow A^\text{(+t)} \rightarrow A^\text{(+t)} \]

Fig. 1 Two-hit model of background and radiation-induced second cancer. a In the background two-hits on the wild-type allele A, occur randomly with rates \( u_1 \) and \( u_2 \). b For the case of second cancer, the first mutation happens during the radiation treatment for a small population of normal stem cells (with rate \( \gamma \)), because of the possible after effects such as genetic instability induced by radiation the second cancer mutation rate, \( u_r \), is assumed to be different from the background rate, \( u_2 \).
deemed appropriate for the two-hit model. An explicit form for \( f_0(t, u_1, u_2, r, N) \) can be obtained (see “Appendix 1” and Iwasa et al. 2005),

\[
f_0(t, u_1, u_2, r, N) = 1 - \exp\left[-Nu_1 \cdot \left(\frac{a + b}{c} \cdot \ln\left(\frac{e^{rt} + a/b}{1 + a/b}\right) - br\right)\right],
\]

where coefficients \( a, b \) and \( c \) are functions of the second mutation rate \( u_2 \) and proliferation/regeneration rate \( r \) given in “Appendix 1”. Upon the appearance of a malignant stem cell in a niche, due to its high fitness advantage, there is a high probability for it to take over the whole niche and form a malignant niche. The overall probability of a malignant clone arising at time \( t \) is given by:

\[
\rho(\tilde{r}, N) f_0(t, u_1, u_2, r, N),
\]

where \( \rho(\tilde{r}, N) \) is the fixation probability of malignant cells with a proliferation strength \( \tilde{r} \) (the fixation probability is the probability that one malignant cell takes over the whole niche). We assume a much higher proliferation rate for malignant cells than for pre-malignant or normal cells and thus the average time to fixation is much shorter relative to the time scales of the latency time of secondary cancer, and given by

\[
t_{\text{lag}} = \Delta t_{\text{fixation}} \sim \frac{1}{r - 1} \ln N \ll \frac{1}{u_2}.
\]

The left-hand-side of Eq. (2) is the approximate value of the fixation time (Durrett 2008). The fixation time for a malignant phenotype is treated as a free parameter that indicates the lag time from the appearance of the first malignant stem cell in a niche until it fixes the system.

As a result of radiation exposure, extra numbers of premalignant cells are initiated during radiation treatment. This number is on the order of one premalignant cell per million normal cells (if distributed uniformly in all the niches) (Sachs and Brenner 2005). Hence, it is assumed that the radiation-induced as well as the spontaneously created premalignant cells are evenly distributed among a subpopulation of niches, i.e., there is a small number of stem cells niches which each have one type-1 cell and \( N - 1 \) type-0 cells, soon after radiotherapy, while the majority of niches are filled with type-0 cells. The dynamics of these two initial conditions are inherently different. The probability of having at least one malignant, type-2, cell arise at time \( t \) after radiation exposure (in a niche or compartment), is denoted by \( f_1(t, u_1, u_2, r, N) \). Notice that the value of the second mutation rate is now changed to \( u_2 \). Under a similar branching process approximation, the functional form of \( f_1(t, u_1, u_2, r, N) \) is taken to be:

\[
f_1(t, u_1, u_2, r, N) = \frac{e^{(r(1-u_2)(a+b)t)}}{(1/a) e^{(r(1-u_2)(a+b)t)} + (1/b)},
\]

Similarly, the probability of having a malignant niche after time \( t \) is \( \rho(\tilde{r}, N) f_1(t, u_1, u_2, r, N) \). To estimate the risk of cancer initiation, it is assumed that the probability of cancer initiation is proportional to the number of malignant niches. It is also assumed that in both background and radiation-exposed cases the number of malignant cells created before the exposure age is negligible, but there can be a number of pre-malignant cells that underwent spontaneous initiation, and these are denoted by \( M_0 \), while the number of radiation-induced pre-malignant cells after radiation is denoted by \( M_1 \). Thus, after radiation exposure, there are \( M_0 + M_1 \) niches with one pre-malignant cell distributed per niche; and \( N - M_0 - M_1 \) niches with zero pre-malignant cells. Similarly, in the background case, the same scenario is assumed, but with only \( M_0 \) niches containing a pre-malignant cell. This can be summarized as follows:

\[
\begin{align*}
\text{Number of malignant cells (background)}
&= N \tilde{\rho} \{ M_0 f_1(t, u_1, u_2, r, N) \\
&\quad + (N - M_0) f_0(t, u_1, u_2, r, N) \}, \\
\text{number of malignant cells (irradiated)}
&= N \tilde{\rho} \{ (M_0 + M_1) f_1(t, u_1, u_2, r, N) \\
&\quad + (N - M_0 - M_1) f_0(t, u_1, u_2, r, N) \},
\end{align*}
\]
where \( \hat{\rho} = \rho(\hat{r}, N) \) and \( t = t_a - t_e - t_{lag} \) is the time since exposure, while \( t_a, t_e \) and \( t_{lag} \) are attained age, age at exposure and lag time as discussed in previous section. The ratio of the above two equations gives a good approximation to the ERR for second cancers. Using the definition of relative risk (RR) and excessive relative risk (ERR),

\[
RR = 1 + ERR = \frac{\text{(number of malignant cells (irradiated))}}{\text{(number of malignant cells (background))}}.
\]

(5)

now, the number of initiated pre-malignant cells during the irradiation time \( M_1 \), can be estimated using the initiation–inactivation–proliferation model proposed by Sachs and Brenner (2005) (see also Sachs et al. 2007) for a given radiotherapy protocol. \( N_{tot}(=N_N) \) denotes the total number of normal cells, \( \gamma(Gy^{-1}) \) denotes the initiation rate of normal cells to pre-malignant cells, then a constant dose delivery rate (radiation strength) leads to a simple form for \( M_1 \) given by:

\[
M_1 \approx N_{tot}(e^{\gamma D} - 1),
\]

(6)

It is again assumed that the number of malignant cells prior to radiation treatment is negligible so that, the number of pre-malignant cells \( M_0 \) that result from spontaneous mutations before the radiation exposure time, \( M_0 \), can be approximated from a simple evolutionary Moran model with two populations of normal and pre-malignant cells. This leads to the forms for \( M_1 \) (as in Sachs and Brenner 2005), and

\[
M_0 \approx N(1 - e^{-\mu N a t_e}),
\]

(7)

where \( t_e \) is the age at exposure. Notice that similar forms are used for the incidence rates in the presence and absence of radiation. The effects of radiation only appear in the altered value of the second mutation rate with an additional number of pre-malignant cells. The fact that the same form Eq. (4) can be used for the background incidence rate relies on the assumption that the number of malignant cells developed before the radiation exposure time is negligible. This will lead to a number of new incidences after the radiation exposure age. Therefore, these approximations limit the ERR estimates to decades after the radiation exposure. However, this is a well-known fact from epidemiological studies of second cancers (Preston et al. 2007).

Figure 3 displays sensitivity of the probability that one malignant cell arises before time \( t \) in the system filled with normal cells for various parameter values. a fixed mutation rates and variable proliferation rates \( r \), b Fixed proliferation rate \( r \) and variable mutation rates \( (\mu_2 = 10 \mu_1) \)

Results

Background age incidences

Using the model described above, (1) the time dependence of the excess relative risks of radiation for second cancer and (2) the effective mutation rate of pre-malignant cells resulting from radiation therapy were calculated. The model was fitted with the background incidences for colon, breast and lung cancers using data extracted from (Shuryak et al. 2009a, b). The value of the proliferation rate, \( r \), can be obtained from these fits. In this framework, the niche size
has a very weak impact on the value of ERR. It is assumed that \( N \sim 10 \) for the case of colorectal cancer as the stem cell niches are restricted to the bottom of each crypt with small populations. For other organs, a much larger value \( N \sim 10,000 \) is used. For the mutation rate values, an estimate from (Iwasa et al. 2005) is employed that approximates the mutation rates for tumor suppressor gene inactivation. A gene is approximately 10 kb long, and we assume a hit on 5% of bases (500 bases) can cause a change-of-function mutation, while each hit might occur with a probability of \( 10^{-10} \) per cell division (Michor et al. 2004; Kunkel 2004). This gives the effective rate of mutation on each gene \( u_1 \sim 10^{-7} \) while each generation is assumed to last about a week. Since the gene can affect either of the two alleles, for the second mutation rate, \( u_2 \), the probability of mutation is higher due to mechanisms of loss of heterozygosity (mitotic recombination and chromosome non-disjunction). Thus, \( u_2 \) is assumed to be on the order of \( 10^{-6} \). For the above estimates, a range of first mutation rates varying between \( 10^{-8} \) and \( 10^{-6} \) is considered, while assuming \( u_2 \) to be an order of magnitude larger than \( u_1 \). Using a lag time of 20 years, results for different cancer incidence models are fitted. This does match the values reported in the literature and is also of the same order of magnitude as that given by Eq. (2), which is the theoretical measure of the lag time.

The results from Surveillance, Epidemiology and End Results (SEER) database (http://seer.cancer.gov/data/) is plotted in Fig. 4 for different values of proliferation rates. For the figure, \( r = 1.05-1.1, N = 10 \) for colon cancer; \( r = 1.03-1.05, N = 10^4 \) for breast cancer and \( r = 1.07-1.1 \) with \( N = 10^4 \) for lung cancer was chosen.

These values give a crude estimate of the second mutation rate post-irradiation. In the following section, an estimate of \( u_2 \) taken from the atomic bomb survivors data is given. Figure 4 displays the model fit with the background incidence data for colon, lung, breast cancers using the above estimated values.

Dose–response in atomic bomb survivors: \( u_2 \) estimate

The dose–response relation for atomic bomb survivors has been reported in the literature (see Preston et al. 2007). The average dose received by survivors is about 0–3 Gy (Preston et al. 2007) (with average estimated to be 0.5 Gy). The dose–response has a linear functional form for the low-dose regions, and gradually bends for high doses (Fig. 5).
displays a plot for low-dose region—for all solid cancer incidences. The slope of the linear fit (which passes through the origin and is indicated by the blue line) is 0.46 per Gy, denotes the weighted absorbed dose, and encompasses the dose from neutrons (which is multiplied by an RBE value of 10 to make it equivalent to the gamma dose).

The dose–response relation reported uses the standardized ERR which is for age at exposure of 30 years and an attained age of 70 years. The dose–response graph can be fitted with a linear function for which the slope indicates the value of ERR per dose. Inserting the value for slope into the analytical form for the ERR, Eq. (4) and using the rough estimates for background mutation rates and proliferation strengths \((u_1 = 10^{-7}, u_2 = 10^{-6}, r = 1.05)\) as discussed above, we solve for the value of the radiation-induced mutation rate,

\[
u_r \simeq 1.0 \times 10^{-3}(\text{Gy}^{-1} \text{ generation}^{-1}). \tag{8}\]

Since the dose–response plot is a cumulative risk for all solid cancers, the above estimate is a good order of magnitude approximation. A much higher value of the mutation rate relative to the background value of \(u_2 \simeq 10^{-6}\) is noticeable. This is interpreted as due to the fact that the secondary effects of radiation probably results in the induction of genomic instability in the model, which effectively manifests itself in the value of the second hit probability.

Tables 1 and 2 show the range of values of the second mutation rates as \(u_1\) is varied between \(10^{-8}\) and \(10^{-6}\) (while keeping \(u_2 = 10u_1\) and \(r\) varies between 1.001 and 1.05).

ERR versus exposure age, Hodgkin’s lymphoma survivors

In order to obtain similar results for second cancers as a result of radiation therapies, the model is applied to the SEER data on second malignancies among Hodgkin’s lymphoma survivors. The estimates for the background mutation rates and the proliferation strength are further tested by fitting the ERR as a function of exposure age (attained age is fixed at 70 years). Here, the results for three types of cancers, namely, colon, breast and lung, are reported. Using averaged values of ERR over age of exposure for each cancer type, we solve for the values of \(u_r\), and use it in Eq. (5) to plot ERR as a function of age at exposure \((u_r)\). Fits were done for values of \(u_1\) in the range \(\sim 10^{-8} – 10^{-6}\), and for small proliferation strengths \(r \sim 1.01 – 1.1\). The values of the best fits are reported in Table 2, and the dependence of exposure age on the ERR is displayed in Fig. 6.

As can be seen, risk is an increasing function of the exposure age as it approaches the attained age of 70. This is in fact in agreement with the phenomenological studies of Preston et al. (2007) in which a Poisson regression analysis is used by assuming a log-quadratic form in their assumption for ERR exposure age dependence.

ERR versus attained age: Hodgkin’s lymphoma survivors

The above analysis will allow prediction of the risks of second cancer for different age groups and for different attained ages. The predicted ERR’s are reported in Fig. 7. Notice that the different age groups of 10, 30 and 50 years display different behaviors of the ERR. This is due to the fact that different values of ERR are used as the input to estimate \(u_r\) for each age group in addition to differences in the age at exposure that explicitly appears in Eq. (5). In other words, the effect of age at exposure on the second hit manifests itself in the value of the second hit probability.

![Fig. 5 ERR versus dose for total solid cancer incidences, age at exposure of 30 years and attained age of 70 years (data points extracted from Preston et al. 2007). The slope of the linear fit is used to estimate the mutation rates in a two-stage model of cancer initiation for atomic bomb survivors (see Eq. 8).](image)

**Table 1** Best fit mutation rates for background incidence curves, along with the post-radiation-induced mutation rate

| \(N\) | \(r\) | \(u_1\) | \(u_2\) | \(u_r\) | \(u_r/u_2\) |
|-----|-----|-----|-----|-----|-----|
| \(10^5\) | 1.01 | \(10^{-8}\) | \(10^{-7}\) | \(2.0 \times 10^{-7}\) | \(2.0 \times 10^0\) |
| \(10^4\) | 1.01 | \(10^{-7}\) | \(10^{-6}\) | \(8.8 \times 10^{-6}\) | \(8.8 \times 10^0\) |
| \(10^4\) | 1.01 | \(10^{-6}\) | \(10^{-5}\) | \(5.9 \times 10^{-5}\) | \(5.9 \times 10^0\) |
| \(10^4\) | 1.03 | \(10^{-8}\) | \(10^{-7}\) | \(4.4 \times 10^{-5}\) | \(4.4 \times 10^2\) |
| \(10^4\) | 1.03 | \(10^{-7}\) | \(10^{-6}\) | \(2.3 \times 10^{-4}\) | \(2.3 \times 10^2\) |
| \(10^4\) | 1.03 | \(10^{-6}\) | \(10^{-5}\) | \(5.5 \times 10^{-4}\) | \(5.5 \times 10^1\) |
| \(10^4\) | 1.05 | \(10^{-8}\) | \(10^{-7}\) | \(3.6 \times 10^{-4}\) | \(3.6 \times 10^1\) |
| \(10^4\) | 1.05 | \(10^{-7}\) | \(10^{-6}\) | \(1.0 \times 10^{-3}\) | \(1.0 \times 10^3\) |
| \(10^4\) | 1.05 | \(10^{-6}\) | \(10^{-5}\) | \(2.2 \times 10^{-3}\) | \(2.2 \times 10^2\)|
cancer risk at a given age (attained age) is twofold: (1) ERR explicitly depends on the age at exposure in Eq. (5). (2) The value of the second mutation rate, $u_r$, depends on the age at exposure and thus should be determined separately for each age group. This is done by using the results of ERR versus age at exposure discussed in the previous section. While the mutation rate, $u_r$, was estimated for standardized attained age of 70 years old in the previous section, exposure ages of 10, 30 and 50 years old are used from Fig. 6 for ERR to estimate a new mutation rate value for each exposure age group. Unsurprisingly, this secondary effect is not very significant as the ERR values do not change much for different exposure ages (with the exception of lung cancer). The above results can be compared with the phenomenological findings of Preston et al. (2007) and Shuryak et al. (2009a, b). The importance of the present result is that one can confirm the behavior of the ERR on a more general grounds based on two-stage mutation model, while much of the details of the model such as number of stem cell niches, $N$, and niche size, $N$, proliferation rate, $r$ and mutation rates $u_{1,2}$ and $u_r$ denote the mutation rates.

It is worth mentioning that for organs with strong hormonal production such as thyroid and breast, one does expect a deviation of the predicted ERR around certain ages where hormone production can aggravate the risk. This can be seen as a slight deviation in Fig. 6 (for breast) around an exposure age of 15–20. The reader is referred to Shuryak et al. (2009b) for a similar deviation in thyroid secondary primary cancers from Hodgkin’s lymphoma survivors. Similar deviations can be observed in the fits for SEER breast cancer incidences in Fig. 4 after age 60, which we propose might be influenced by a similar hormonal mechanism. However, it is suggested that the overall framework especially the estimate of second cancer mutation rate and possibility of genomic instability due to radiation is independent of such details. Further investigation of hormonal effects and its interplay with other mechanisms of radiation-induced cancers remains a subject for further investigation.

### Discussion

In the present paper, a biologically motivated evolutionary model is described to forecast radiation-induced second cancer risks over time, post-irradiation. The mathematical framework adopts an evolutionary dynamics approach, particularly with respect to the two-hit mechanism used to model cancer initiation. This two-hit process accounts for two types of mutations that occur in normal cells, similar to that of tumor suppressor gene inactivation. The assumption is that normal cells can mutate to pre-malignant cells during radiation treatment, which in turn transform to malignant cells over a period of time. The second mutation can be correlated with either genomic instability (which is considered to be one of the secondary effects of radiation) or to spontaneous mutation (due to the effects of several carcinogens from, for example, food, background radiation, etc.).

The current unified framework is an improvement over other models found in the literature. The short-term model allows one to consider various radiation therapy protocols used in current clinical practice. The output of this biological model is then used as an input to track the number of malignant niches over time to predict the changes in ERR over time, using a simplified evolutionary model. One of the major improvements over existing models is that a stochastic formalism is used to predict ERR over time. This seems a more realistic approach for applications to clinical data since mutations appear to occur randomly post-irradiation. Another major advantage of the presented model is that it has fewer parameters compared to other existing models.

This framework takes into consideration the short-term biological mechanism of pre-malignant cell initiation, along with a long-term formalism resulting in the manifestation of a secondary tumor. The short-term model adopts the initiation–inactivation–proliferation model of Sachs and Brenner (2005) to estimate the number of radiation-induced pre-malignant cells at the end of treatment. The long-term formalism relies on an evolutionary model of the two-hit processes similar to the tumor suppressor gene inactivation model. An evolutionary model is developed based on the underlying biological assumptions. The two important parameters in the model are proliferation strengths and mutation rates which are estimated from the atomic bomb survivors data. The results from the presented model are consistent with those of the work of Preston et al. (2007).

In the present paper, the model was applied to patient survivors of Hodgkin’s lymphoma. For this, the

### Table 2. Second cancer mutation rate estimates for colon, breast and lungs

|                | Ave ERR (Gy$^{-1}$) | $u_1$ (first hit, background) | $u_2$ (second hit, background) | $u_r$ (second hit after irradiation) | $u_r/u_2$ |
|----------------|---------------------|-------------------------------|-------------------------------|-------------------------------------|-----------|
| Colon          | 0.735               | $10^{-7}$                     | $10^{-6}$                     | $2.8 \times 10^{-3}$               | $2.8 \times 10^{-3}$ |
| Breast         | 0.883               | $10^{-8}$                     | $10^{-7}$                     | $3.5 \times 10^{-3}$               | $3.5 \times 10^{-3}$ |
| Lung           | 0.316               | $10^{-6.5}$                   | $10^{-5.5}$                   | $4 \times 10^{-3}$                 | $1.3 \times 10^{-3}$ |
proliferation strengths and mutation rate parameters were extracted for various organs (in particular: colon, breast and lungs) from the background incidences (taken from the SEER database). This was done by fitting the developed

Fig. 6 ERR versus exposure age for colon, breast and lungs among Hodgkins lymphoma survivors from SEER data (http://seer.cancer.gov/data/ see also Shuryak et al. 2009b). Solid lines are the best fit from the presented model using background parameter values from SEER background incidence rates

Fig. 7 Predicted values for ERR vs. attained age for various age groups for colon, breast and lung cancers. The parameters in Eq. (5) are fixed using SEER background data (http://seer.cancer.gov/data/). Different segments of each graph correspond to exposure ages of 10, 30 and 50 years old. A similar distinction is done in Preston et al. (2007) (see discussion in the text)

proliferation strengths and mutation rate parameters were extracted for various organs (in particular: colon, breast and lungs) from the background incidences (taken from the SEER database). This was done by fitting the developed
analytical functions to the background incidence curves, in order to obtain mutation rates and proliferation strengths. Then, these values were used to predict ERR’s for various attained and exposure ages. The predicted age-dependent ERR values are consistent with epidemiological findings and clinical studies. An important point to highlight from the present calculations is that the mutation rate post-irradiation \( (u_r) \), is greater than that of the second mutation rate \( (u_2) \) for background incidence. It is also observed that for low proliferation rates and for low risk, the corresponding \( u_r \) estimate is similar to \( u_2 \).

In future work, we intend to extend our framework to include therapeutic doses using dose–volume histograms (DVHs) and validate this on existing second cancer data sets. It is also intended to extend this work to compare more contemporary protocols for various radiotherapy treatment techniques to predict the risk of second cancers for individual patients. There are several ongoing debates on the effects of radiation, (e.g., concerning genomic instability). It is planned to modify the model to investigate and understand the effects of radiation on tissue architecture as well as on genomic instability.

**Acknowledgments** M. Kohandel and S. Sivaloganathan are supported by an NSERC/CIHR Collaborative Health Research grant. The authors thank DC Hodgson from the Princess Margaret Hospital, Toronto for fruitful discussions.

**Appendix 1: Evolutionary dynamics of two-hit process**

Consider a population of \( N \) cells (inside a niche) governed by the Moran process (types—0, 1, 2 indicating normal, pre-malignant and malignant phenotypes). At each time step, a cell is randomly chosen, based on its fitness, to reproduce and another cell is randomly chosen to die. In the presence of mutations, such as in the model presented in this paper, at each time step, there is an alternative possibility (instead of death–birth) that a normal cell may transform into a pre-malignant cell or alternatively a pre-malignant cell may transform into a malignant cell. The mutation rate for the first hit is \( u_1 \) and the rate for the second hit (pre-malignant into malignant) is \( u_2 \) (or \( u_r \) after radiation). As discussed in the text, let \( f_0(t, u_1, u_2, r, N) \) be the probability that a first malignant cell has emerged \( \) before time \( t \) given that the initial population consists of all normal cells in the niche. Similarly, let \( f_1(t, u_1, u_2, r, N) \) be the probability that a malignant cell emerges in a niche of size \( N \) before time \( t \), beginning with one pre-malignant cell and \( N - 1 \) normal cells. It is straightforward to show that the two probabilities satisfy the continuous time limit of the Kolmogorov equations,

\[
\frac{df_0(t)}{dt} = Nu_1 f_1(t)(1 - f_0(t)),
\]

\[
\frac{df_1(t)}{dt} = ru_2 - (1 - r + 2ru_2)f_1(t) - r(1 - u_2)^2(t),
\]

with the initial condition \( f_0(0) = 0 \). The solutions of the above set of equations can be directly obtained under a branching process approximation, i.e., independence of two lineages starting from two different cells, one can then obtain an analytical solution for these probabilities,

\[
f_1(t, u_1, u_2, r, N) = \exp\left\{\frac{(r(1 - u_2)(a + b)t) - 1}{(1/a) \exp[r(1 - u_2)(a + b)t] + (1/b)}\right\},
\]

\[
f_0(t, u_1, u_2, r, N) = 1 - \exp\left\{-N u_1 \left\{\frac{a + b}{c} \ln \left[\frac{e^{u_1} + a/b}{1 + a/b}\right] - b t\right\}\right\},
\]

with \( a, b, c \) given as,

\[
a = \frac{1}{2(1 - u_2)} \left[\frac{1 - r + 2u_2}{r} + \sqrt{\left(\frac{1 - r + 2u_2}{r}\right)^2 + 4u_2}\right],
\]

\[
b = \frac{1}{2(1 - u_2)} \left[\frac{1 - r + 2u_2}{r} + \sqrt{\left(\frac{1 - r + 2u_2}{r}\right)^2 + 4u_2}\right],
\]

\[
c = r(1 - u_2)(a + b).
\]

As discussed in the main body of the paper, both the cumulative age incidence of different cancer types and excess relative risk of cancer (second cancer in this case) is expressed in terms of functions \( f_0(t, u_1, u_2, r, N) \) and \( f_1(t, u_1, u_2, r, N) \), which lead to a simple form for the ERR. As is obvious, there is no dependence on the total number of niches \( N \), and the dependence on the niche size, \( N \), is very weak. However, for more realistic fits, the value of the proliferation strength \( r \) is not extremely close to unity and thus the above approximation might not be appropriate. However, it is easy to see from the above approximate results that any quantity of interest (i.e., number of background age incidences or ERR) is represented by very weak functions of the niche size, \( N \). Thus, the present rough estimates for the value of \( N \) do not change any of the important conclusions resulting from this work.

**Appendix 2: Initiation–inactivation–repopulation model**

In this section, the initiation–inactivation–repopulation (Sachs and Brenner 2005) formalism is briefly reviewed to estimate the number of pre-malignant cells at the end of radiation treatment. The effect of radiation is simplified into two mechanism: (1) cell killing which is modeled by the linear–quadratic approximation and gives the number

\[
\frac{df_0(t)}{dt} = Nu_1 f_1(t)(1 - f_0(t)),
\]

\[
\frac{df_1(t)}{dt} = ru_2 - (1 - r + 2ru_2)f_1(t) - r(1 - u_2)^2(t),
\]

with the initial condition \( f_0(0) = 0 \). The solutions of the above set of equations can be directly obtained under a branching process approximation, i.e., independence of two lineages starting from two different cells, one can then obtain an analytical solution for these probabilities,
of either normal or pre-malignant cells killed due to radiation dose. (2) The cell initiation which is assumed to be a linear function of the dose with a small constant mutation rate per dose rate. Two populations of normal stem cells and pre-malignant stem cells are assumed. Normal stem cells grow logistically while they can die during the radiation exposure times, where cell death is given by linear–quadratic formula. They can also transform into pre-malignant cells during the exposure time. The population of pre-malignant cells has a similar growth form while its proliferation is regulated by the population of normal cells and has the same cell death rate due to radiation and also a positive rate due to normal stem cell initiation. The above can be written in the form of the following coupled system of ordinary differential equations.

\[
\frac{dn(t)}{dt} = r_0 n(t) \left( 1 - \frac{n(t)}{K} \right) - (\alpha + \beta D) \cdot \frac{dD}{dt} n(t) - \gamma \cdot \frac{dD}{dt} n(t),
\]

\[
\frac{dm(t)}{dt} = r_0 \lambda m(t) \left( 1 - \frac{n(t)}{K} \right) - (\alpha + \beta D) \cdot \frac{dD}{dt} m(t) + \gamma \cdot \frac{dD}{dt} n(t),
\]

where \( r_0 \) is the normal cell repopulation rate while \( r_0 \lambda \) is the pre-malignant repopulation rate. \( dD/dt \) is the dose delivery rate, \( \lambda \) represents the relative proliferation rate of pre-malignant cells to the normal cells and \( K \) is the carrying capacity of normal cells. \( \alpha \) and \( \beta \) are the cell killing rates from linear–quadratic dose dependence (LQ). The number of pre-malignant cells after the total radiation dose has been applied and is the quantity of importance here. This can be analytically calculated for the simplified case of an acute dose, \( dD/dt = \text{const.} \) applied in a finite interval of time

\[
M = m(\infty) = N(\exp(\gamma D) - 1),
\]

\[
\sim N^2 \gamma D,
\]

where \( D \) is the total dose. The total number of the cells \( N_{\text{tot}} \), is given by the number of niche times the niches size, \( NN \). For the values of the radiation-induced initiation rate \( \gamma \), this will result in \( M/N_{\text{tot}} \sim 10^{-6} \) (per Gy). Since we are interested on an estimate for number of radiation-induced pre-malignant cells, the details of cell killing mechanism in LQ formula does not appear in Eq. (13).

The above formalism can be expanded to incorporate fractionation therapy, i.e., non-constant dose–delivery rate which has been discussed in the literature. In the present work, there is no need to deal with these details. However, as a future direction of research, all the details of radiotherapy (fractionation protocol and the dose–volume histogram for each patient) can be used as input to estimate the ERR as a function of exposure age and attained age.

References

Aleman BM, van den Belt-Dusebout AW, Klokman WJ, van’t Veer MB, Bartelink H, van Leeuwen FE (2003) Long-term cause-specific mortality of patients treated for Hodgkin’s disease. J Clin Oncol 21(18):3431–3439

Armitage P (1985) Multistage models of carcinogenesis. Environ Health Perspect 63:195

Armitage P, Doll R (1954) The age distribution of cancer and a multistage theory of carcinogenesis. Br J Cancer 8(1):1

Athar BS, Paganetti H (2011) Comparison of second cancer risk due to out-of-field doses from 6-MV IMRT and proton therapy based on 6 pediatric patient treatment plans. Radiother Oncol 98(1):87–92

Bhatia S, Sklar C (2002) Second cancers in survivors of childhood cancer. Nat Rev Cancer 2(2):124–132

Bhatia S, Yasui Y, Robison LL, Birch JM, Bogue MK, Diller L, DeLaat C, Fossati-Bellani F, Morgan E, Oberlin O, Gregory Reaman JT, Frederick B, Ruymann, Meadows AT (2003) High risk of subsequent neoplasms continues with extended follow-up of childhood Hodgkin’s disease: report from the late effects study group. J Clin Oncol 21(23):4386–4394

Dores GM, Metayer C, Curtis RE, Lynch CF, Clarke EA, Glimelius B, Storm H, Pukkala E, van Leeuwen FE, Holowaty EJ, Andersson M, Wiklund T, Joensuu T, van’t Veer MB, Stovall M, Gospodarowicz M, Travis LB (2002) Second malignant neoplasms among long-term survivors of Hodgkin’s disease: a population-based evaluation over 25 years. J Clin Oncol 20(16):3484–3494

Durrett R (2008) Probability models for DNA sequence evolution, probability and its applications, vol 12, 2nd edn. Springer, Berlin

Hall EJ (2006) Intensity-modulated radiation therapy, protons, and the risk of second cancers. Int J Radiat Oncol Biol Phys 65(1):1–7

Hodgson DC, Gilbert ES, Dores GM, Schonfeld SJ, Lynch CF, Storm H, Hall P, Langmark F, Pukkala E, Andersson M et al (2007a) Long-term solid cancer risk among 5-year survivors of Hodgkin’s lymphoma. J Clin Oncol 25(12):1489–1497

Hodgson DC, Koh ES, Tran TH, Heydarian M, Tsang R, Pintilie M, Xu T, Huang L, Sachs RK, Brenner DJ (2007b) Individualized estimates of second cancer risks after contemporary radiation therapy for Hodgkin lymphoma. Cancer 110(11):2576–2586

Huang L, Snyder AR, Morgan WF (2003) Radiation-induced genomic instability and its implications for radiation carcinogenesis. Oncogene 22(37):5848–5854

Iwasa Y, Michor F, Komarova NL, Nowak MA (2005) Population genetics of tumor suppressor genes. J Theor Biol 233(1):15–23

Komarova NL, Sengupta A, Nowak MA (2003) Mutation-selection networks of cancer initiation: tumor suppressor genes and chromosomal instability. J Theor Biol 223(4):433–450

Kunkel TA (2004) DNA replication fidelity. J Biol Chem 279(17):16895–16898

Lindsay K, Wheldon E, Deehan C, Wheldon T (2001) Radiation carcinogenesis modelling for risk of treatment-related second tumours following radiotherapy. Br J Radiol 74(882):529–536

Little MP (2009) Cancer and non-cancer effects in Japanese atomic bomb survivors. J Radiol Prot 29(2A):A43

Manem V, Kohandel M, Hodgson D, Sharpe M, Sivaloganathan S (2014a) The effect of radiation quality on the risks of second malignancies. Int J Radiat Biol 1–39 [Epub ahead of print]

Manem VS, Dhawan A, Kohandel M, Sivaloganathan S (2014b) Efficacy of dose escalation on TCP, recurrence and second cancer risks: a mathematical study. Br J Radiol 87(1043):20140377

Michor F, Iwasa Y, Nowak MA (2004) Dynamics of cancer progression. Nat Rev Cancer 4(3):197–205
Moran PAP (1962) The statistical processes of evolutionary theory, 1st edn. Clarendon Press and Oxford University Press, Oxford
Moteabbed M, Yock TI, Pagani H (2014) The risk of radiation-induced second cancers in the high to medium dose region: a comparison between passive and scanned proton therapy, IMRT and VMAT for pediatric patients with brain tumors. Phys Med Biol 59(12):2883
Nowak MA, Komarova NL, Sengupta A, Jallepalli PV, Shih IM, Vogelstein B, Lengauer C (2002) The role of chromosomal instability in tumor initiation. Proc Natl Acad Sci 99(25):16226–16231
Paganetti H, Athar BS, Moteabbed M, Adams JA, Schneider U, Yock TI (2012) Assessment of radiation-induced second cancer risks in proton therapy and IMRT for organs inside the primary radiation field. Phys Med Biol 57(19):6047
Preston D, Ron E, Tokuyou S, Funamoto S, Nishi N, Soda M, Mabuchi K, Kodama K (2007) Solid cancer incidence in atomic bomb survivors: 1958–1998. Radiat Res 168(1):1–64
Preston DL, Kusumi S, Tomonaga M, Izumi S, Ron E, Kuramoto A, Kamada N, Dohy H, Matsui T, Nonaka H et al (1994) Cancer incidence in atomic bomb survivors. Part iii: leukemia, lymphoma and multiple myeloma, 1950–1987. Radiat Res 137(2s):S68–S97
Sachs RK, Brenner DJ (2005) Solid tumor risks after high doses of ionizing radiation. Proc Natl Acad Sci USA 102(37):13040–13045
Sachs RK, Shuryak I, Brenner D, Fakir H, Hlatky L, Hahnfeldt P (2007) Second cancers after fractionated radiotherapy: stochastic population dynamics effects. J Theor Biol 249(3):518–531
Schneider U (2009) Mechanistic model of radiation-induced cancer after fractionated radiotherapy using the linear–quadratic formula. Med Phys 36(4):1138–1143
Schneider U, Lomax A, Besserer J, Pemler P, Lombriser N, Kasera-Hotz B (2007) The impact of dose escalation on secondary cancer risk after radiotherapy of prostate cancer. Int J Radiat Oncol Biol Phys 68(3):892–897
Schneider U, Sumila M, Robotka J, Weber D, Gruber G (2014) Radiation-induced second malignancies after involved-node radiotherapy with deep-inspiration breath-hold technique for early stage Hodgkin lymphoma: a dosimetric study. Radiat Oncol 9(1):58
Shuryak I, Hahnfeldt P, Hlatky L, Sachs RK, Brenner DJ (2009a) A new view of radiation-induced cancer: integrating short-and long-term processes. Part i: approach. Radiat Environ Biophys 48(3):263–274
Shuryak I, Hahnfeldt P, Hlatky L, Sachs RK, Brenner DJ (2009b) A new view of radiation-induced cancer: integrating short-and long-term processes. Part ii: second cancer risk estimation. Radiat Environ Biophys 48(3):275–286
Sigurdson AJ, Jones IM (2003) Second cancers after radiotherapy: Any evidence for radiation-induced genomic instability? Oncogene 22(45):7018–7027
Travis LB, Hill D, Dores GM, Gospodarowicz M, van Leeuwen FE, Holowaty E, Glimelius B, Andersson M, Pukkala E, Lynch CF, Pee D, Smith SA, Veer MBV, Jonsu T, Storm H, Stovall M, JDB, Gilbert E, Gail MH (2005) Cumulative absolute breast cancer risk for young women treated for Hodgkin lymphoma. J Natl Cancer Inst 97(19):1428–1437
van Leeuwen FE, Klokman WJ, Stovall M, Dahler EC, van't Veer MB, No-ordijk EM, Crommelin MA, Aleman BMP, Broeks A, Gospodarowicz M, Travis LB, Russell NS (2003) Roles of radiation dose, chemotherapy, and hormonal factors in breast cancer following Hodgkin’s disease. J Natl Cancer Inst 95(13):971–980
Wang J, Pulido JS, O’Neill BP, Johnston PB (2014) Second malignancies in patients with primary central nervous system lymphoma. Neuro Oncol pii:nou105 [Epub ahead of print]
Xu XG, Bednarz B, Paganetti H (2008) A review of dosimetry studies on external-beam radiation treatment with respect to second cancer induction. Phys Med Biol 53(13):R193
Yahalom J (2009) Role of radiation therapy in Hodgkin’s lymphoma. Cancer J 15(2):155–160
Yeoh KW, Mihalek NG (2010) Role of radiotherapy in modern treatment of Hodgkin’s lymphoma. Adv Hematology 2011:258797
Yorke E, Kutcher G, Jackson A, Ling C (1993) Probability of radiation-induced complications in normal tissues with parallel architecture under conditions of uniform whole or partial organ irradiation. Radiotherapy Oncol 26(3):226–237
Zhang Y, Goddard K, Spinellii JJ, Gotay C, McBride ML (2012) Risk of late mortality and second malignant neoplasms among 5-year survivors of young adult cancer: a report of the childhood, adolescent, and young adult cancer survivors research program. J Cancer Epidemiol 2012:103032