Towards the characterization of neutron carcinogenesis through direct action simulations of clustered DNA damage

Logan Montgomery[1], Christopher M Lund[1], Anthony Landry[2,3], John Kildea[1]

[1]Medical Physics Unit, McGill University, Montreal, QC, H4A3J1, Canada

[2]Prince Edward Island Cancer Treatment Centre, Charlottetown, PE, C1A8T5, Canada

[3]Department of Radiation Oncology, Dalhousie University, Halifax, NS, B3H4RZ, Canada

E-mail: logan.montgomery@mail.mcgill.ca

Abstract.

Neutron exposure poses a unique radiation protection concern because neutrons have a large, energy-dependent relative biological effectiveness (RBE) for stochastic effects. Recent computational studies on the microdosimetric properties of neutron dose deposition have implicated clustered DNA damage as a likely contributor to this marked energy dependence. So far, publications have focused solely on neutron RBE for inducing clusters of DNA damage containing two or more DNA double strand breaks (DSBs). In this study, we have conducted a novel assessment of neutron RBE for inducing all types of clustered DNA damage that contain two or more lesions, stratified by whether the clusters contain DSBs (complex DSB clusters) or not (non-DSB clusters). This assessment was conducted for eighteen initial neutron energies between 1 eV and 10 MeV as well as a reference radiation of 250 keV x-rays. We also examined the energy dependence of cluster length and cluster complexity because these factors are believed to impact the DNA repair process. To carry out our investigation, we developed a user-friendly TOPAS-nBio application that includes a custom nuclear DNA model and a novel algorithm for recording clustered DNA damage. We found that neutron RBE for inducing complex DSB clusters exhibited similar energy dependence to the canonical neutron RBE for stochastic radiobiological effects, at multiple depths in human tissue. Qualitatively similar results were obtained for non-DSB clusters, although the quantitative agreement was lower. Additionally we identified a significant neutron energy dependence in the average length and complexity of clustered lesions. These results support the idea that many types of clustered DNA damage contribute to the energy dependence of neutron RBE for stochastic radiobiological effects and imply that the size and constituent lesions of individual clusters should be taken into account when modeling DNA repair. Our results were qualitatively consistent for (i) multiple radiation doses (including a low dose 0.1 Gy irradiation), (ii) variations in the maximal lesion separation distance used to define a cluster, and (iii) two distinct collections of physics models used to govern particle transport. Our complete TOPAS-nBio application has been released under an open source license to enable others to independently validate our work and to expand upon it.

1 Introduction

Reducing the long-term stochastic risk of radiation carcinogenesis is one of the primary objectives of radiation protection. Of particular concern are radiation scenarios involving neutron exposure, which include high-energy ($\gtrsim 8\,\mathrm{MeV}$) radiation therapy (Howell *et al* 2006, Maglieri *et al* 2015), nuclear incidents (Shuryak *et al* 2020), and space travel (Benton *et al* 2001, Koshiishi *et al* 2007). Neutron exposure must be assessed separately from exposure to other forms of radiation because neutrons have a comparatively high and energy-dependent relative biological effectiveness (RBE) for stochastic effects. This energy dependence is encapsulated in the neutron weighting factors (w_R) promulgated by the International Commission on Radiological Protection (ICRP 2003, 2007) and the neutron quality factors (Q) published by the United States Nuclear Regulatory Commission (US NRC 2021). Both sets of factors convey a marked energy dependence of neutron RBE with a peak value occurring around 1 MeV. Accordingly, radiation therapy scenarios warrant particular attention because the fluence spectrum of secondary neutrons that is generated in such environments has a fast neutron peak around 1 MeV (Howell *et al* 2006, Maglieri *et al* 2015), coincident with maximal neutron RBE. While the secondary neutron absorbed dose associated with radiation therapy is low compared to the therapeutic dose (Kry *et al* 2017), it is generally believed that stochastic radiobiological effects can occur at any dose.

Although the w_R and Q factors exhibit qualitatively similar energy dependences, their magnitudes are highly discrepant. This follows from the considerable variety of neutron RBE values reported by the epidemiological investigations and radiobiological experiments on which these factors are based (ICRP 2007, US NRC 2021). In light of these discrepancies, it is desirable to trace the origin of neutron RBE to fundamental biophysical principles, as recently asserted by Baiocco *et al* (2016). To this end, consideration must be given to the mechanisms by which ionizing radiation is generally believed to cause cancer.

The somatic mutation theory of carcinogenesis (Vaux 2011, Hanahan and Weinberg 2000, 2011) posits that genomic mutations can lead to carcinogenesis and is perhaps the most longstanding mechanistic theory of carcinogenesis. Nowadays, it is generally believed that radiation can induce mutagenesis via several pathways, including both nuclear DNA damage and non-targeted effects (Little 2000, Iyer and Lehnert 2000). Thus, it is of interest to characterize direct radiation-induced DNA damage as one piece of the larger puzzle of radiation-induced carcinogenesis.

The propensity of radiation to induce clusters of DNA damage, which are difficult to repair and rarely occur endogenously, has been widely theorized as the primary mechanism by which radiation induces mutagenesis (Goodhead 1994, Ward 1995, Magnander and Elmroth 2012, Georgakilas *et al* 2013, Sage and Shikazono 2017). Therefore, one approach to evaluate the mutagenic potential of neutrons is to model their relative ability to induce clustered DNA damage.

Track-structure Monte Carlo (TSMC) techniques that utilize low-energy physics models to simulate event-by-event particle interactions can be used to model radiation-induced damage on the nanoscopic level (Nikjoo *et al* 1998, Tajik-Mansoury *et al* 2017). Our group recently used the open-source Geant4-DNA toolkit (Incerti *et al* 2010a,b, Bernal *et al* 2015, Incerti *et al* 2018) to analyze neutron RBE for the microdosimetric endpoint dose-mean lineal energy at various depths

in human tissue (Lund *et al* 2020). The methodology we employed was similar to an earlier study by Baiocco *et al* (2016) who used PHITS (Sato *et al* 2013) and PARTRAC (Friedland *et al* 2011). The energy dependence of the microdosimetric neutron RBE values obtained in both studies was qualitatively similar to the ICRP's w_R factors and the US NRC's Q factors. While these microdosimetric studies provide information on the spatially-clustered nature of energy deposition by neutrons and their secondary particles, they do not explicitly provide information on clustered DNA damage. To do so, a geometric model of nuclear DNA must be incorporated.

In a parallel study, Baiocco *et al* (2016) incorporated a geometric DNA model and evaluated neutron RBE for the direct induction of clusters of DNA lesions containing at least two double strand breaks (DSBs) within 25 bp (sometimes labeled DSB++ (Friedland *et al* 2011)). These results also exhibited similar qualitative energy dependence to the neutron w_R and Q factors. However, DSB++ lesions represent only a small subset of an infinite variety of clustered DNA damage lesions, each of which may have mutagenic potential. Recent experimental review papers (Sage and Shikazono 2017, Georgakilas *et al* 2013, Magnander and Elmroth 2012) have discussed the mutagenic potential of clustered lesions both with and without DSBs (i.e. complex-DSB clusters and non-DSB clusters). These papers also highlighted how DNA repair is heavily influenced by the specific types and number of lesions comprising a cluster as well as the number of base pairs between them.

In this work we performed a novel assessment of neutron RBE for stochastic effects by simulating neutron RBE for the direct induction of both complex DSB clusters and non-DSB clusters. This manuscript describes how we used the TOPAS-nBio framework (Schuemann *et al* 2019a) to create a new open-source Monte Carlo application that includes a custom nuclear DNA model and a novel algorithm that records clustered DNA damage. We used our application to determine neutron RBE for inducing both types of clustered DNA damage as a function of neutron energy at multiple depths in human tissue. We also analyzed the length and number of lesions comprising each cluster in consideration of the variety of lesions encompassed by the complex DSB and non-DSB cluster types, as well as the possible variations in their mutagenic consequences. Finally, we performed repeat simulations with variations in particular simulation parameters in order to assess the validity of key assumptions that we made.

2 Methods

2.1 Overview

Our overall methodology included four steps, which are summarized schematically in figure 1.

In step 1, condensed-history Monte Carlo simulations were used to determine the energy spectra and relative dose contributions of the secondary particles liberated in human tissue by uniform fluences of monoenergetic neutrons and a reference radiation of 250 keV x-rays. These simulations were performed in Geant4 v10.04.p02 and used the ICRU-4 sphere (White *et al* 1989) as a human tissue phantom. Similar to the approach taken by Baiocco *et al* (2016), data were recorded in three scoring volumes of increasing depth. Step 1 was conducted in our previous work (Lund *et al*

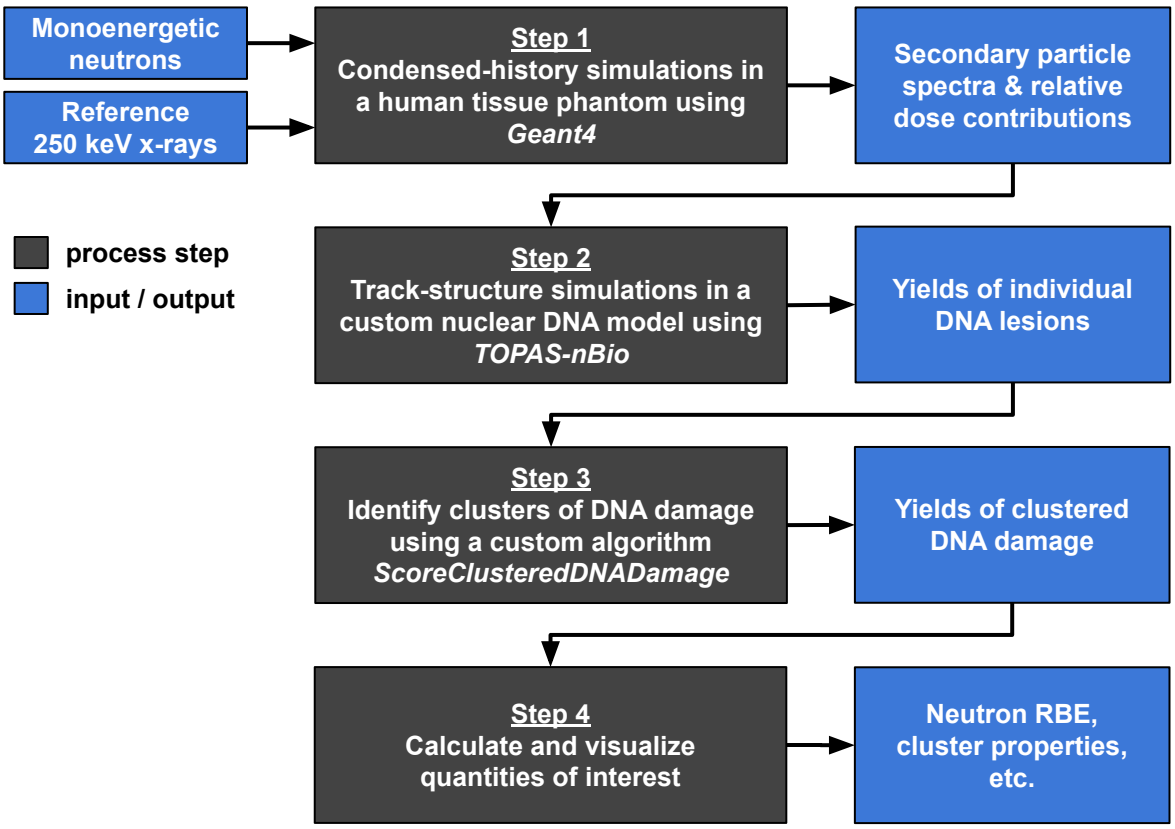


Figure 1: Schematic overview of our Monte Carlo simulation and analysis pipeline designed to determine direct neutron-induced DNA damage and compare with a reference radiation of 250 keV x-rays.

2020) and the resulting data were reused in this work.

In step 2, TSMC methods were used to determine the amount of DNA damage induced by the secondary particles identified in step 1. Particle tracks were stochastically sampled from the secondary particle energy spectra and used to irradiate a geometric model of nuclear DNA that we developed. Energy depositions that occurred in the sensitive nucleotide volumes were empirically converted to DNA damage lesions. These simulations were carried out using TOPAS v3.3.1 and the TOPAS-nBio 1.0 beta, which are both based on Geant4 v10.05.p01.

In step 3, the map of individual DNA lesions was processed using a novel clustering algorithm that we developed to aggregate lesions in close proximity. Clusters were then stratified according to whether at least one DSB was present (i.e. complex DSB clusters) or not (i.e. non-DSB clusters). The yields of both types of clustered DNA damage were recorded.

Finally, in step 4, we calculated neutron RBE for inducing both types of clustered DNA damage. In consideration of other cluster parameters that may impact DNA repair, we also analyzed the length, complexity, and density of each cluster.

2.2 Step 1: Condensed-history simulations

We have previously published our approach to condensed-history simulations (Lund *et al* 2020). In short, Geant4 was used to simulate the irradiation of the ICRU-4 soft tissue-equivalent sphere (White *et al* 1989), which contained three sensitive scoring volumes at increasing depth, as shown in figure 2(a). The scoring volumes of radius 1.5 cm were centred along a single axis at radial distances of 0 cm (inner scoring volume), 7.5 cm (intermediate scoring volume), and 13.5 cm (outer scoring volume) from the centre of the phantom. A separate irradiation was performed for (i) 18 different neutron energies ranging from 1 eV to 10 MeV (i.e. the energy range of interest in linac-based radiation therapy) and (ii) a reference radiation of 250 keV x-rays. Each irradiation consisted of a uniform fluence of 1×10^{10} primary particles.

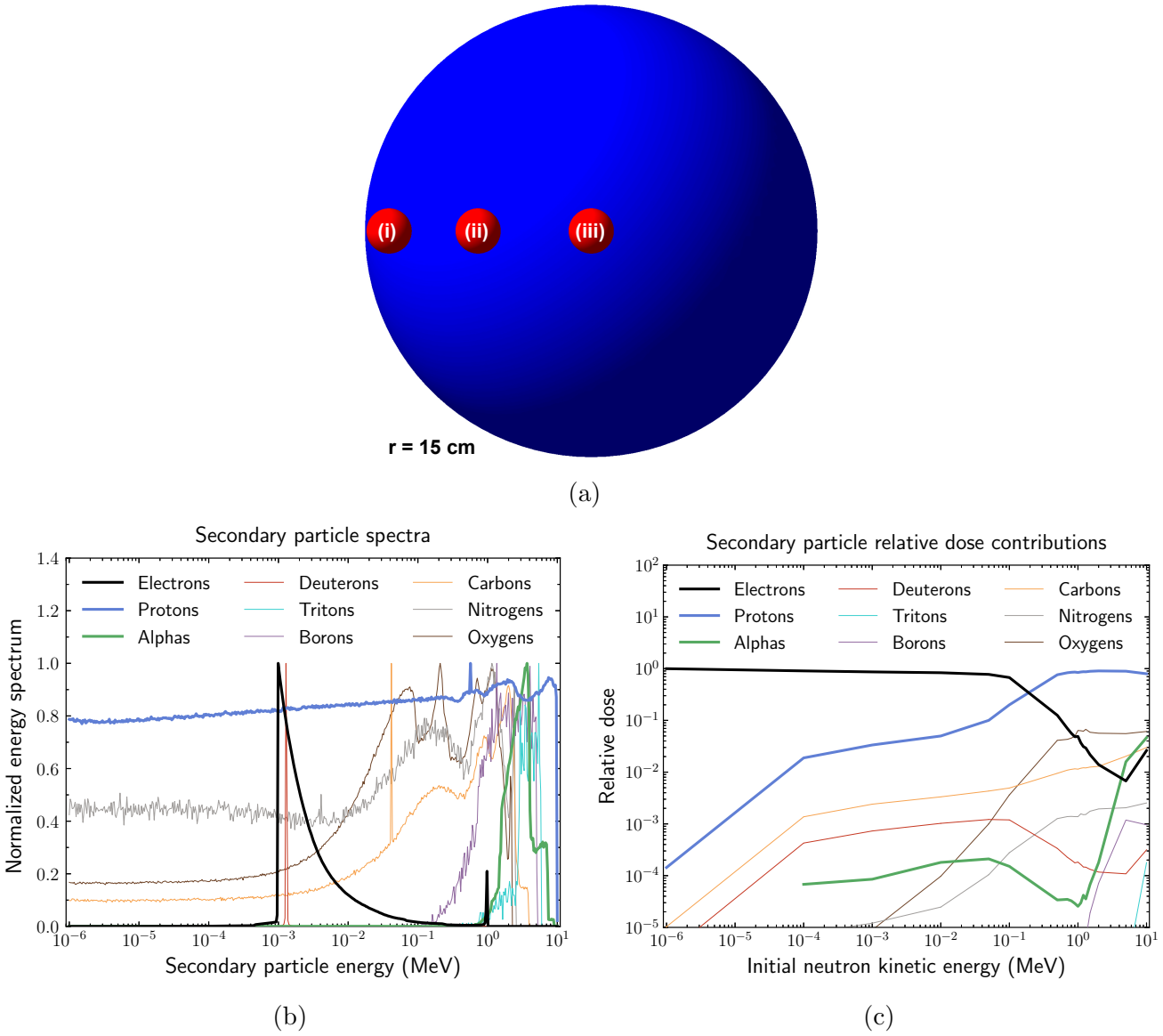
During these simulated irradiations, neutrons interacted with the tissue phantom and liberated a variety of secondary charged particles. Among the most relevant neutron interactions were (i) elastic scatter with hydrogen, (ii) neutron capture reactions involving hydrogen and nitrogen, and (iii) inelastic reactions involving carbon, nitrogen, and oxygen (Lund *et al* 2020). We assumed that charged particle equilibrium existed in the scoring volumes given their relatively large size compared to the range of the secondary charged particles. This assumption permitted use of the local approximation, wherein secondary particles were killed at their point of generation and their kinetic energy was considered to be deposited locally. Using the local approximation, the energy spectra and relative dose contributions of all secondary particle species were determined in each scoring volume for each initial particle energy. The energy spectrum of each species was simply the normalized distribution of initial kinetic energies among particles of that species. The relative dose contribution of each species was the sum of initial kinetic energies for that species relative to the sum for all species combined. Representative secondary particle energy spectra and relative dose contributions obtained in the intermediate scoring volume are shown in figure 2.

Special consideration was given to electrons generated above 1 MeV since Geant4-DNA does not currently handle their transport (Incerti *et al* 2018, Lund *et al* 2020). Instead of being killed, these electrons were allowed to propagate further and were tracked down to 1 MeV. The resulting 1 MeV electron and all other higher-order electron tracks produced during the slowing down process were recorded as independent tracks. Additionally, in simulation step 2, we considered only three secondary charged particle species: (i) electrons, (ii) protons, and (iii) alpha particles, because Geant4-DNA does not currently include complete physics models to handle the transport of heavier ions at the energies of interest (Incerti *et al* 2016).

2.3 Step 2: Track-structure simulations

2.3.1 Nuclear DNA model

Despite an abundance of nuclear DNA models that are described in the literature (Friedland *et al* 2011, Meylan *et al* 2017, Lampe *et al* 2018, Tang *et al* 2019, Zabihi *et al* 2020, Zhu *et al* 2020a, Sakata *et al* 2020), there were no complete open-source models available at the time of this study. Thus, we constructed a custom nuclear DNA model using the TOPAS extensions framework. A graphical overview of our DNA model is presented in figure 3 and a summary of the geometric



(a) (b) (c)

Figure 2: Irradiation geometry and representative results from the condensed-history Monte Carlo simulations performed in our previous work (Lund *et al* 2020). (a) Our irradiation target, the ICRU-4 soft tissue phantom (White *et al* 1989) and three scoring volumes shown in red: (i) outer, (ii) intermediate, and (iii) inner. (b) Normalized secondary particle energy spectra produced by initial 10 MeV neutrons in the intermediate scoring volume. (c) Relative dose contribution of secondary particles in the intermediate scoring volume as a function of initial neutron energy. Particle species represented with thicker lines (i.e. electrons, protons, and alpha particles) were the only species considered in this work.

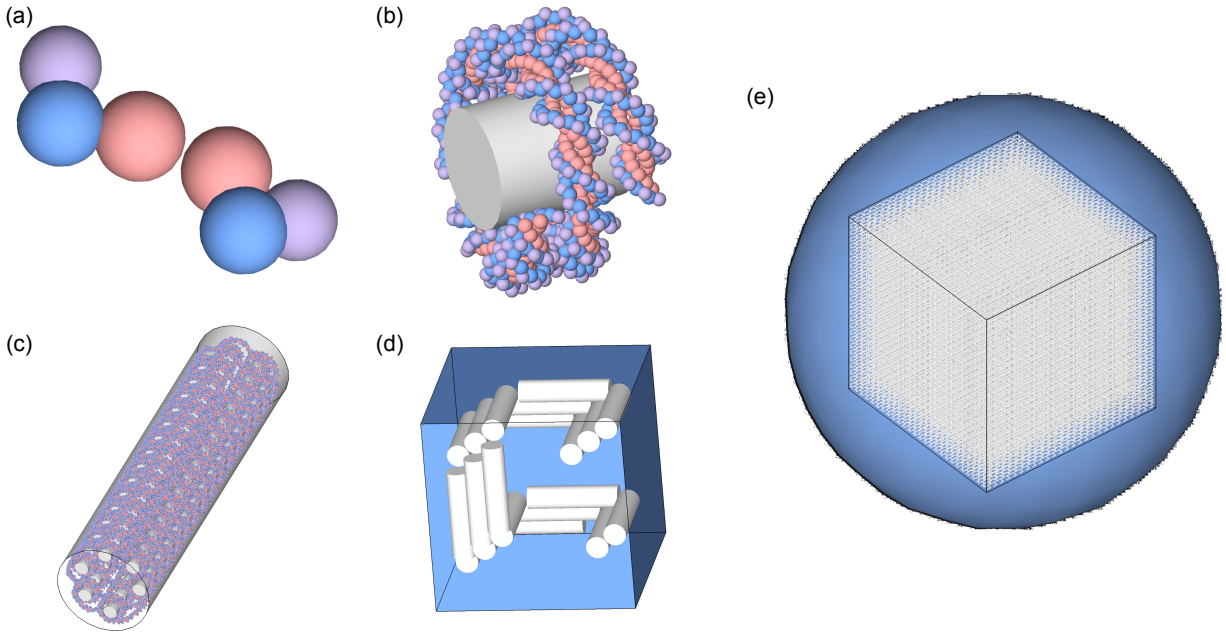


Figure 3: Our nuclear DNA model. (a) Nucleotide base pair containing two nitrogenous bases (red), two deoxyribose molecules (blue) and two phosphate groups (purple). (b) Nucleosome containing 154 base pairs arranged in a double helix and wrapped around a cylindrical volume representing a histone protein complex. (c) Chromatin fibre containing 90 nucleosomes arranged in a helical pattern and each connected via 46 base pairs of linker DNA. (d) Voxel containing 20 chromatin fibres arranged in a fractal pattern. (e) Cubic human fibroblast nucleus containing 6.3 Gbp of DNA, enclosed in a spherical cell volume.

parameters describing it is provided in table 1.

Our chromatin fibre model is a re-implementation of a model that was developed by Villagrasa *et al* (2017), the source code of which was included in the TOPAS-nBio beta release (McNamara *et al* 2018, Schuemann *et al* 2019b). We ported this source code from Geant4 to TOPAS format to allow parameters to be read in from a TOPAS parameter file. In our model, nucleotide base pairs were created using six spheres to represent two nitrogenous bases, two deoxyribose molecules, and two phosphate groups (figure 3(a)). These spheres were cut in some locations to prevent geometric overlap. Nucleosomes were formed by wrapping a cylindrical histone complex with 154 nucleotide base pairs arranged in a double helix (figure 3(b)). Ninety nucleosomes were then linked together in helical fashion to fill out a cylindrical chromatin fibre (figure 3(c)).

We constructed a nuclear model by creating multiple copies of our chromatin fibre model using a voxelized approach, similar to the method described by Zhu *et al* (2020a). Each voxel was filled with 20 chromatin fibres that were arranged in the fractal pattern shown in figure 3(d). Using a fractal arrangement of chromatin fibres is justified by experimental evidence on the nature of chromatin folding reported by Lieberman-Aiden *et al* (2009) and is now commonplace in DNA modeling (Lampe *et al* 2018, McNamara *et al* 2018, Zhu *et al* 2020a, Sakata *et al* 2020).

Table 1: Parameters describing our geometric model of nuclear DNA.

Component	Parameter	Value
DNA base pair	Nitrogenous base radius	0.30 nm
	Deoxyribose radius	0.29 nm
	Phosphate group radius	0.27 nm
Nucleosome	Histone complex dimensions	2.4 nm radius, 5.72 nm height
	Number of bp per nucleosome	154 bp (+ 46 bp of linker DNA)
Chromatin fibre	DNA content per chromatin fibre	90 nucleosomes (18 000 bp)
	Chromatin fibre radius	17 nm
	Chromatin fibre length	136 nm
Voxel	Number of fibres per voxel	20 fibres
	Voxel dimensions	0.3 µm × 0.3 µm × 0.3 µm
Nucleus	Number of voxels	17 576 voxels (26 × 26 × 26 grid)
	Volume	475 μm^3
	Number of base pairs	6.3 Gbp
	Density of DNA	13.3 Mbp/μm^3
Cell	Volume	2000 μm^3

Identical voxels were placed on a cubic grid using Geant4 replica volumes. This was a simple and computationally-efficient way to create a nucleus-sized volume with 6.3 Gbp of DNA at a density of 13.3 Mbp/μm^3, consistent with human nuclei (Ghosh and Jost 2018, Zhu *et al* 2020a).

The resulting 475 μm^3 nucleus model was placed in a 2000 μm^3 spherical cell volume to simulate a fibroblast in G_0/G_1 phase (figure 3(e)), consistent with the fibroblast dimensions described by Seaman *et al* (2015). The majority of recently published Monte Carlo investigations of radiation-induced DNA damage have used fibroblast models (Baiocco *et al* 2016, Meylan *et al* 2017, Zhu *et al* 2020a, Sakata *et al* 2020). Fibroblasts are an appealing choice of cell to model because they are proliferative, are found throughout the human body, and are not overly specialized like a neuron or a red blood cell.

The entire cellular volume was treated as liquid water with a density of 1 g·cm^{-3} except in the nucleotide volumes where a density of 1.407 g·cm^{-3} was used (Śmiałek *et al* 2013, Zhu *et al* 2020a). In order to allow subsequent identification of clustered DNA damage, unique ID numbers were used to identify each: (i) voxel, (ii) chromatin fibre, (iii) DNA strand, (iv) nucleotide, and (v) molecule within each nucleotide.

To benchmark our DNA model against previously published results, we irradiated it with monoenergetic protons (500 keV, 1 MeV, and 10 MeV) and recorded the resulting direct DNA

single strand breaks (SSBs) and double strand breaks (DSBs), as defined in section 2.4. The total number of strand breaks (SBs) was calculated as the number of SSBs plus twice the number of DSBs (two strand breaks per DSB) to enable comparison with the literature. We compared our direct SB and DSB yields with analogous results obtained by Zhu *et al* (2020a), Meylan *et al* (2017), and Sakata *et al* (2020), each of whom used their own nuclear DNA model. Although the irradiation conditions varied between authors, this exercise allowed us to assess the consistency of our results with published data. Our setup most closely emulated the setup of Zhu *et al* (2020a) by placing the initial protons at random locations on the nuclear surface with a random inward orientation and by using the G4EmDNAPhysics_option2 physics constructor (Incerti *et al* 2018) to govern particle transport.

2.3.2 Physics settings

In Geant4-DNA (and thus TOPAS-nBio), there are a variety of competing physics models that describe the physical processes governing electron transport in liquid water (Incerti *et al* 2018). These processes include ionization, electronic excitation, vibrational excitation, elastic scattering, molecular attachment, and Auger electron emission. Geant4-DNA offers a variety of physics constructors, i.e. collections of physics models, from which users can choose. Among these are the G4EmDNAPhysics_option2 (opt2) and the G4EmDNAPhysics_option4 (opt4) constructors. The physics models included in the opt4 constructor are more recent and sophisticated than the models included in opt2. However, the opt4 physics models can only be used for electrons between 10 eV and 10 keV. On the other hand, the physics models included in opt2 can be used from 10 eV up to 1 MeV.

In our previous work (Lund *et al* 2020), we developed a custom Geant4-DNA physics constructor, labeled G4EmDNAPhysics_hybrid2and4, in order to extract the best features of the opt2 and opt4 constructors. Our hybrid constructor uses the physics models from opt4 up to 10 keV and the models from opt2 at higher energies up to 1 MeV. In this work, we imported G4EmDNAPhysics_hybrid2and4 into TOPAS-nBio as a custom physics module and used it in our simulations. Consistent with recommendations by the TOPAS collaboration (Perl *et al* 2016), electron tracks with kinetic energy less than 10 eV were killed and their energy was deposited locally.

Only a single set of physics models is provided for protons and for alpha particles in Geant4-DNA (Incerti *et al* 2018) and thus there is no variation across physics constructors, including our hybrid constructor.

2.3.3 Irradiation setup

As indicated schematically in figure 1, source particle energies for our track-structure simulations were obtained by stochastically sampling the neutron secondary particle spectra described in section 2.2. Particles were placed randomly throughout the cell volume (including the nucleus) with random orientation. This approach was chosen to emulate the manner in which secondary particles were generated and recorded in the upstream condensed-history simulations and is consistent with

the work of Baiocco *et al* (2016).

For each scoring volume k and initial neutron or x-ray energy E, particles were simulated until 1 Gy of dose was delivered to the nucleus. A target dose D_0 of 1 Gy was chosen for consistency with recent literature (Baiocco *et al* 2016, Zabihi *et al* 2020, Tang *et al* 2019, Zhu *et al* 2020a). Additional simulations were performed with alternative D_0 values to assess dose dependence, as described in section 2.6. A delivered dose equal to D_0 was achieved by running three distinct simulations, one for each secondary particle species i and a corresponding species-specific target dose $[D_i(E)]_k$. This species-specific target dose was set as D_0 scaled by the relative dose contribution $[r_i(E)]_k$ for a particular species i, at a particular neutron energy E, and in a particular scoring volume k:

$$[D_i(E)]_k = [r_i(E)]_k D_0$$

Each simulation was terminated once the species-specific target dose was delivered, but only after complete processing of the current event (i.e. the current source particle track and any secondary tracks). Thus, the target dose $[D_i(E)]_k$ was slightly exceeded in every simulation and accounted for by recording the actual delivered dose $[d_i(E)]_k$.

Every simulated irradiation was repeated 100 times using pseudorandom seed values to obtain statistically independent results. Simulations were performed on our internal computer cluster that contains 212 threads across 106 cores. While the specifics varied between configurations, each simulation required no more than a few thousand source particles and ten minutes of simulation time.

2.4 Step 3: DNA damage clustering algorithm

Energy depositions in the sensitive DNA volumes (i.e. the nitrogenous bases, deoxyribose molecules, and phosphates) were recorded for each irradiation described in section 2.3. A custom TOPAS scorer was developed to process these energy depositions and to record five types of DNA damage:

(i) Single strand breaks (SSBs)

(ii) Base lesions

(iii) Double strand breaks (DSBs)

(iv) Complex DSB clusters

(v) Non-DSB clusters

Schematic examples of each of these damage types are shown in figure 4. The definition of each type of damage and the associated rationale is described in the remainder of this section.

An SSB was recorded when the cumulative energy deposited in the sugar-phosphate molecules of a nucleotide (i.e. the backbone) exceeded 17.5 eV. This energy threshold is based on the findings of Charlton and Humm (1988) who modeled the experimental work of Martin and Haseltine (1981) to

analyze SSB induction by Auger electrons emitted by iodine-125. Use of this threshold is standard in the field (Nikjoo *et al* 1997, Watanabe *et al* 2015, Meylan *et al* 2017, Lampe *et al* 2018, Mokari *et al* 2020, Zhu *et al* 2020a).

A base lesion was recorded when the cumulative energy deposited in a nitrogenous base exceeded 17.5 eV. Compared to SSBs, fewer simulation studies have considered base damage and there is no consensus as to which interactions lead to a base damage or the optimal energy threshold at which it occurs (Meylan *et al* 2017). Indeed, radiation-induced base damage encompasses a diverse array of lesions including apurinic sites, apyrimidic sites, and oxidized bases (Sage and Shikazono 2017). Despite the lack of consensus, there is precedent for generically scoring base lesions and applying the same threshold as used for SSBs (Nikjoo *et al* 1997, 2001, Watanabe *et al* 2015).

A DSB was recorded when two SSBs occurred within 10 bp of each other on opposing strands of the DNA double helix (i.e. within approximately one turn of the DNA double helix), as is standard in the field (Nikjoo *et al* 1997, Watanabe *et al* 2015, Meylan *et al* 2017, Lampe *et al* 2018, Mokari *et al* 2020, Zhu *et al* 2020a, Sakata *et al* 2020). This maximum separation distance is based on experimental evidence by Van der Schans (1978) who evaluated the maximum SSB separation above which DSBs would not occur in bacteriophage DNA.

Complex DSB clusters and non-DSB clusters were identified by using our custom DNA damage clustering algorithm to process the list of recorded SSBs, base lesions, and DSBs. Any two or more lesions were included in a cluster if they occurred within 40 bp of each other, i.e. within a few turns of the DNA double helix and close enough to impact DNA repair (Magnander and Elmroth 2012, Georgakilas *et al* 2013, Sage and Shikazono 2017). If the resulting cluster contained one or more DSBs we labeled it as a complex DSB cluster, otherwise we labeled it as a non-DSB cluster.

A summary of our "default" simulation parameters is provided table 2. Since all research involving modeling requires a variety of assumptions that can impact the results, we designed our code such that the simulation parameters can be readily modified by the user using a TOPAS-style parameter file. We used this functionality to perform additional simulations with variations in some of these parameters in order to assess their impact on our results, as described in section 2.6.

2.5 Step 4: Quantities of interest

2.5.1 DNA damage yields

DNA damage yields $[Y_i^j(E)]_k$ were recorded for all five types of DNA damage j, and for each particle species i, initial neutron energy E, and scoring volume k (i.e. penetration depth).

The species-specific damage yields were combined via a weighted sum to get the total neutron-induced yield for a particular type of DNA damage, initial neutron energy, and scoring volume as follows:

$$[Y_N^j(E)]_k = \sum_{i=1}^{I} \frac{[Y_i^j(E)]_k [D_i(E)]_k}{[d_i(E)]_k}$$

Dose weighting was necessary to normalize the yields to the same target dose D_0, which facilitated

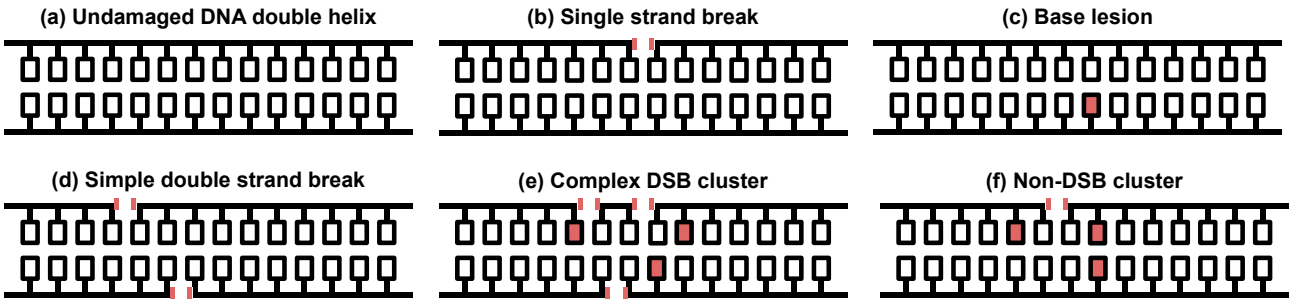


Figure 4: Schematics examples of the types of DNA damage considered in our simulations. (a) An undamaged DNA double helix wherein each square represents a nitrogenous base attached to the sugar-phosphate backbone. (b) A single strand break (SSB) depicted as a red separation in the backbone. (c) A generic base lesion depicted as a red base. (d) A double strand break (DSB) containing two SSBs on opposing strands within 10 base pairs of each other. (e) A complex DSB cluster containing two or more damage sites, including at least one DSB, each within 40 base pairs of each other. (f) A non-DSB cluster containing two or more SSBs or base lesions, each within 40 base pairs of each other.

unbiased comparison with the reference 250 keV x-rays.

A similar calculation was performed for the reference x-ray radiation, however in this case there was only one secondary particle species (electrons, denoted as e) and one initial energy, 250 keV.

$$[Y_X^j]_k = \frac{[Y_e^j]_k [D_e(E)]_k}{[d_e(E)]_k}$$

The mean neutron-induced and x-ray-induced yield for each simulation configuration was obtained by averaging 100 statistically independent simulations. The standard deviation of the mean yield (sometimes called the standard error of the mean) (Joint Committee for Guides in Metrology 2008) was determined by dividing the corresponding standard deviation by the square root of the number of simulations.

2.5.2 Relative biological effectiveness (RBE)

Neutron RBE for inducing each type of DNA damage j was calculated as a function of neutron energy E by dividing the mean neutron-induced yield by the corresponding mean x-ray-induced yield, in each scoring volume k, as follows:

$$[\mathrm{RBE}^j(E)]_k = \frac{[Y_N^j(E)]_k}{[Y_X^j]_k}$$

The uncertainty in RBE was obtained by propagating the standard deviation of the mean neutron-induced and x-ray-induced yields using conventional uncertainty propagation rules.

Table 2: Default simulation parameters.

Parameter	Value
Target geometry	Nuclear DNA Model described in table 1.
Target material	Liquid water Density = 1.407 g·cm^{-3} in the sensitive DNA volumes, 1 g·cm^{-3} elsewhere.
Physics Module	G4EmDNAPhysics_hybrid2and4 Custom physics constructor that uses physics models from G4EmDNAPhysics_option4 between 10 eV–10 keV and physics models from G4EmDNAPhysics_option2 between 10 keV–1 MeV.
Source particles	Electrons, protons, or alpha particles. Energies stochastically sampled from neutron and x-ray secondary particle spectra.
Simulation cutoff	1 Gy Scaled by the relative dose contribution of the source particle species being simulated (1 Gy total across all species for each initial neutron or x-ray energy in each scoring volume).
Number of histories	Variable Between 1–10 000 histories per simulation.
Number of repeated simulations	100
Induction of SSB	17.5 eV Cumulative energy deposit in the sugar-phosphate molecules comprising a nucleotide.
Induction of base lesion	17.5 eV Cumulative energy deposit in a nitrogenous base.
Induction of DSB	Two SSBs within 10 bp on opposing strands.
Induction of clustered DNA damage	Aggregation of individual DNA lesions within 40 bp of each other. If cluster contains a DSB, labeled as a complex DSB cluster. If not, labeled as a non-DSB cluster.

2.5.3 Cluster length

The length of every recorded DNA damage cluster was obtained by calculating the number of base pairs separating the damage lesions at either end of the cluster (including the endpoints). The

mean cluster length was calculated separately for complex DSB clusters and non-DSB clusters in all three scoring volumes and for each initial neutron or x-ray energy. The corresponding standard deviation of the mean was calculated accordingly.

2.5.4 Cluster complexity

The complexity of every recorded DNA damage cluster was calculated as the number of individual lesions (SSBs or base lesions) within each cluster. A DSB was simply interpreted as two SSBs for the purpose of this analysis. The mean cluster complexity was calculated separately for complex DSB clusters and non-DSB clusters in all three scoring volumes and for each initial neutron or x-ray energy. The corresponding standard deviation of the mean was calculated accordingly.

2.5.5 Cluster density

The density of every recorded DNA damage cluster was calculated as the cluster complexity divided by the cluster length. The mean cluster density was calculated separately for complex DSB clusters and non-DSB clusters in all three scoring volume and for each initial neutron or x-ray energy. The corresponding standard deviation of the mean was calculated accordingly.

2.6 Dose and parameter sensitivity analysis

Motivated by previous investigations by Pater *et al* (2014) and Zhu *et al* (2020b), we conducted a dose and parameter sensitivity analysis to evaluate the impact of various assumptions that we made in our simulations. We identified the target dose D_0, the DNA damage clustering distance, and the physics constructor as the parameters that were the most crucial in interpreting our results and that had not been investigated previously. We repeated our full set of simulations with variations in these parameters in order to gain further insight into our results and to assess their robustness.

2.6.1 Dose

Because a primary motivation for this work lies in understanding radiation-induced effects leading to carcinogenesis at low doses, it was necessary to repeat our analysis in the low-dose regime (in addition to the standard 1 Gy used for consistency with the literature). While a precise definition of the low-dose regime is lacking, its upper limit is often set at approximately 0.1 Gy (ICRP 2007, Mullenders *et al* 2009, Harbron 2012, Morgan and Bair 2013). Therefore we performed another set of simulations with the target dose D_0 set at 0.1 Gy. To get a sense of the linearity of our results with respect to dose, we also performed a set of simulations with a D_0 of 2 Gy.

2.6.2 DNA damage clustering distance

We identified a range of DNA damage clustering distances that were utilized in published Monte Carlo studies that considered clustered DNA damage in some capacity (Nikjoo *et al* 1997, Watanabe *et al* 2015, Baiocco *et al* 2016, Lampe *et al* 2018, Tang *et al* 2019). The minimum and

maximum values of this range were 10 bp and 100 bp, respectively. Our default 40 bp clustering distance represented an intermediate value in this range and was consistent with the general belief that adjacent lesions within a few turns of the DNA double helix can impact the DNA repair process (Magnander and Elmroth 2012, Georgakilas *et al* 2013, Sage and Shikazono 2017). To assess the sensitivity of our results to this choice, we repeated our simulations using the 10 bp minimum and 100 bp maximum clustering distances obtained from our literature review.

2.6.3 Physics constructor

Finally, we evaluated the impact of our choice to use the opt4 physics models at low electron energies (below 10 keV) instead of the opt2 physics models. We did so by simply repeating our simulations with the opt2 constructor instead of our hybrid constructor. Note that we did not compare with the opt4 constructor alone because it does not handle inelastic interactions over the full range of electron energies that we considered.

3 Results and discussion

In this section we present and discuss a representative subset of the results obtained in our simulations. The results that we felt did not offer additional insight are provided in the supplementary materials for completeness.

3.1 Benchmarking our nuclear DNA model

Total SB ($Y^{\mathrm{SSB}} + 2Y^{\mathrm{DSB}}$) and DSB yields induced by monoenergetic protons via direct effects are plotted alongside other published results in figure 5. We have only compared DSB yields with Zhu *et al* (2020a) because the other studies did not explicitly distinguish directly-induced DSBs from indirectly-induced DSBs. Although every study used a different nuclear DNA model and different irradiation conditions, all the results are of the same order of magnitude and do not vary significantly with energy. The differences in magnitude are consistent with model variations in both the size of the sensitive DNA volumes and the overall density of DNA in the nucleus. For example, both Zhu *et al* (2020a) and Meylan *et al* (2017) modeled a hydration shell (Swarts *et al* 1992) around their nucleotide base pairs, in which energy depositions were accumulated with those in the molecules. This resulted in a larger effective volume of their sugar-phosphate backbone, and corresponded to higher SB yields. Additionally, the model used by Zhu *et al* (2020a) had the highest DNA density of 14.4 Mbp/µm^3 while the model used by Sakata *et al* (2020) had the lowest DNA density of 12 Mbp/µm^3. Overall, our results are in reasonable agreement with published work, which serves to benchmark our nuclear DNA model.

3.2 DNA damage yields

Yields of all five types of DNA damage obtained in the intermediate scoring volume are plotted as a function of initial neutron energy in figure 6(a). Without consideration of indirect effects, we cannot compare absolute damage yields with results obtained from radiobiological experiments. Nevertheless, the relative difference in yields of each type of DNA damage can be assessed.

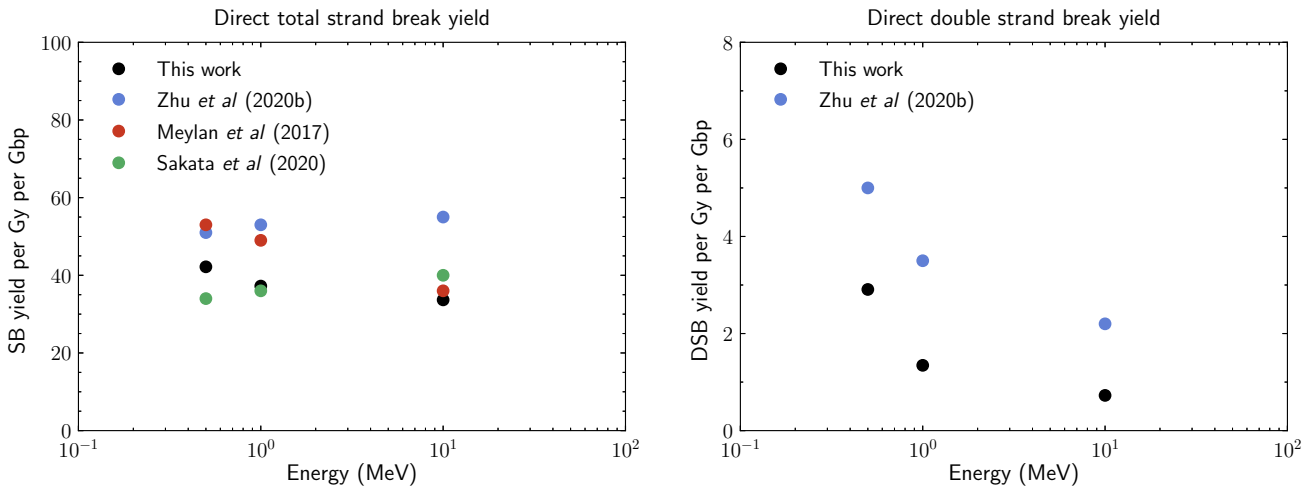


Figure 5: Comparison of proton-induced DNA damage yields obtained using our nuclear DNA model with published results obtained using other nuclear DNA models. (a) Total strand break yield. (b) Double strand break yield.

Approximately two times more SSBs were induced than base lesions, which is accounted for by the fact that the combined volume of the sugar-phosphate backbone is approximately twice that of the nucleotide bases in our model. We observed an increase in aggregate damage yields (DSBs, complex DSB clusters, and non-DSB clusters) beginning around 100 keV and a corresponding decrease in the yield of isolated lesions (SSBs and base lesions). As described in microdosimetric terms in our previous work (Lund *et al* 2020), this trend occurs due to a change in dominance of the relative dose contribution from electrons to protons around 100 keV (intersection of the black and blue curves in figure 2(c)). The yield of non-DSB clusters was several times greater than the yield of complex DSB clusters, which agrees well with predictions by Magnander and Elmroth (2012) and Nikitaki *et al* (2016).

Figure 6(b) and figure 6(c) demonstrate the depth dependence of clustered DNA damage yields by comparing the results obtained in each scoring volume. The peak height was relatively consistent across all scoring volume depths for both types of clustered DNA damage likely due to the fact that the shape of the secondary particle spectra did not drastically change with depth (Lund *et al* 2020). However, we observed that this peak began at higher initial neutron energies with increasing depth. This shift coincided with a shift in the neutron energy at which the relative dose contribution becomes dominated by protons with increasing depth (Lund *et al* 2020). Ultimately, this result can be traced back to increased neutron moderation with increased penetration depth in human tissue.

3.3 Neutron relative biological effectiveness (RBE)

Neutron RBE values for inducing all five types of DNA damage in the intermediate scoring volume are plotted as a function of initial neutron energy in figure 7(a). We observed that neutron RBE

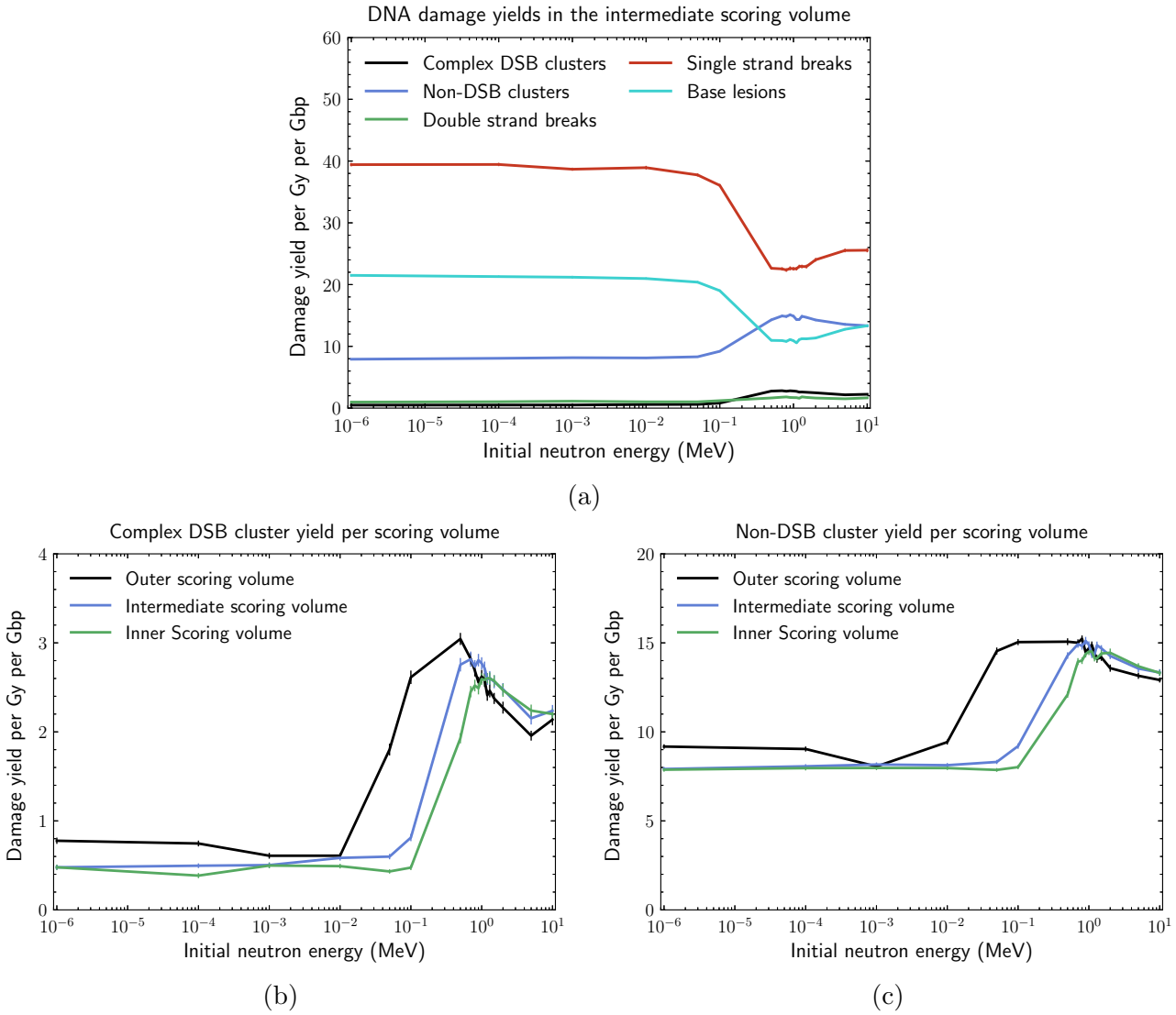


Figure 6: DNA damage yields per Gy of dose delivered per Gbp as a function of initial neutron energy. (a) DNA damage yield for all five types of DNA damage in the intermediate scoring volume. (b) Complex DSB cluster yield in each of the three scoring volumes. (c) Non-DSB cluster yield in each of the three scoring volumes. Plotted values are the mean values obtained over 100 statistically independent simulations. Error bars represent the standard deviation of the mean and are too small to be seen in some cases.

for inducing complex DSB clusters exhibited a sharp increase around 100 keV with a peak value of 8.0 ± 0.6 at 700 keV. Neutron RBE for inducing non-DSB clusters and simple DSBs similarly exhibited an increase around 100 keV but with smaller peak values of 2.28 ± 0.05 and 2.2 ± 0.1 respectively. Associated with the increase in neutron RBE for aggregate damage lesions was a decrease in neutron RBE for isolated SSBs and base lesions. These trends follow logically from the observed trends in DNA damage yield (section 3.2).

Figure 7(b) shows how neutron RBE for inducing complex DSB clusters varied with scoring depth. As with the complex DSB cluster yields shown in figure 6(b), we saw that the characteristic

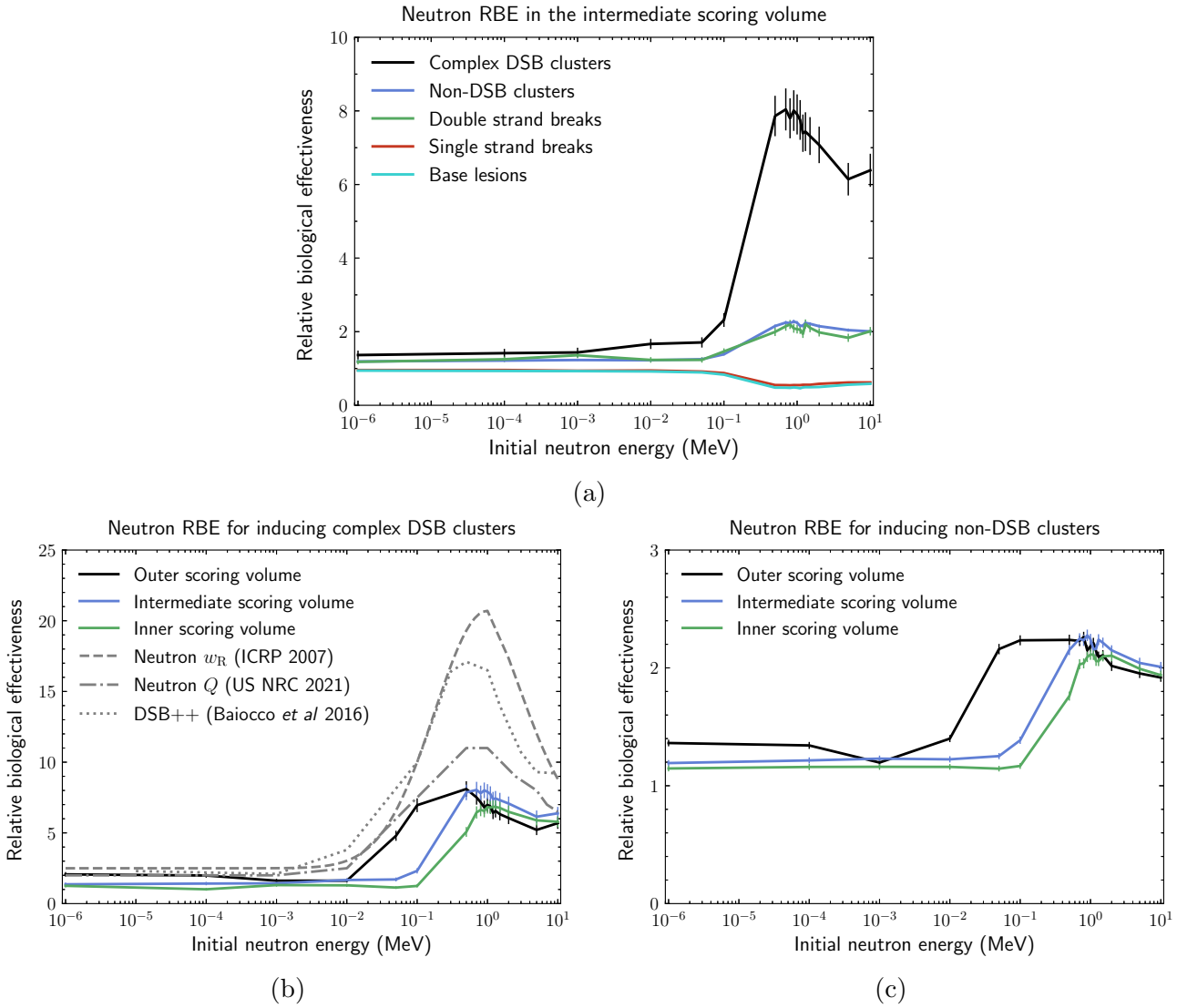


Figure 7: Neutron RBE for inducing DNA damage, as a function of initial neutron energy. (a) Neutron RBE for inducing five types of DNA damage in the intermediate scoring volume. (b) Neutron RBE for inducing complex DSB clusters in three scoring volumes of increasing depth. Results are compared with the ICRP neutron weighting factors (ICRP 2007), the US NRC neutron quality factors (US NRC 2021), and neutron RBE for DSB++ induction as obtained by Baiocco *et al* (2016). (c) Neutron RBE for inducing non-DSB clusters in three scoring volumes of increasing depth. Plotted values are the mean values obtained over 100 statistically independent simulations. Error bars represent the standard deviation of the mean.

increase in neutron RBE began at higher neutron energies with increasing depth. The peak neutron RBE value was relatively consistent across all depths within the allotted uncertainties, having a maximum value of 8.1 ± 0.6 in the outer scoring and a minimum value of 6.9 ± 0.5 in the inner scoring volume. These findings are qualitatively consistent with our previous microdosimetric results (Lund *et al* 2020). Figure 7(c) shows the analogous results obtained for non-DSB clusters for which a comparable depth dependence was observed. While the peak values are significantly lower than for complex DSB clusters, they exhibit similar energy dependence and have an RBE value significantly larger than 1.

In Figure 7(b) we also compared our results for inducing complex DSB clusters with the more specific DSB++ result obtained by Baiocco *et al* (2016), as well as the ICRP w_{R} (ICRP 2007) and US NRC Q factors (US NRC 2021). This comparison was not explicitly included on the non-DSB cluster plot (Figure 7(c)) so that the axes could be magnified and the depth dependence more clearly seen. Quantitative comparison with Baiocco *et al* (2016) highlights the fact that neutrons have variable propensity for inducing specific types of clustered lesions relative to x-rays. One could review the literature to identify other clustered lesions of interest that are believed to have high mutagenic potential, and isolate their results from the broader classifications that we employed here.

Overall, we observed good qualitative agreement between our results and the reference w_{R} and Q factors. Our findings, and those of Baiocco *et al* (2016), suggest that the propensity of neutrons to induce a variety of direct clustered DNA damage lesions, both with and without DSBs, is a promising mechanism to explain the energy dependence of neutron RBE for stochastic radiobiological effects.

One discrepancy between our simulated data and the reference factors is that the peak in our data did not exhibit the same rapid falloff above 1 MeV as in the reference factors. We hypothesize that this may be due to the fact that we have not yet considered indirect effects in our model. Direct action is more dominant than indirect action for the high LET secondary protons and alphas, which begin to dominate the relative dose contribution over the low LET electrons at these higher neutron energies (Lund *et al* 2020). Thus, we expect that incorporating indirect effects into our model would increase the yield of x-ray-induced clustered damage relatively more than neutron-induced damage, thereby reducing RBE at these high energies. Additionally, we have not modeled DNA repair following irradiation nor other carcinogenic pathways (like non-targeted effects) that are inherently built into the radiobiological data on which the w_{R} and Q factors are based.

3.4 Cluster length, complexity, and density

In addition to overall damage yields, we considered cluster length, cluster complexity, and cluster density as factors that may impact the DNA repair process and thus relate to stochastic effects. Figure 8 demonstrates the dependence of these quantities on neutron energy and compares them with the results obtained using the reference 250 keV x-ray radiation. Again, only results obtained in the intermediate scoring volume are shown here, as we found no significant depth dependence.

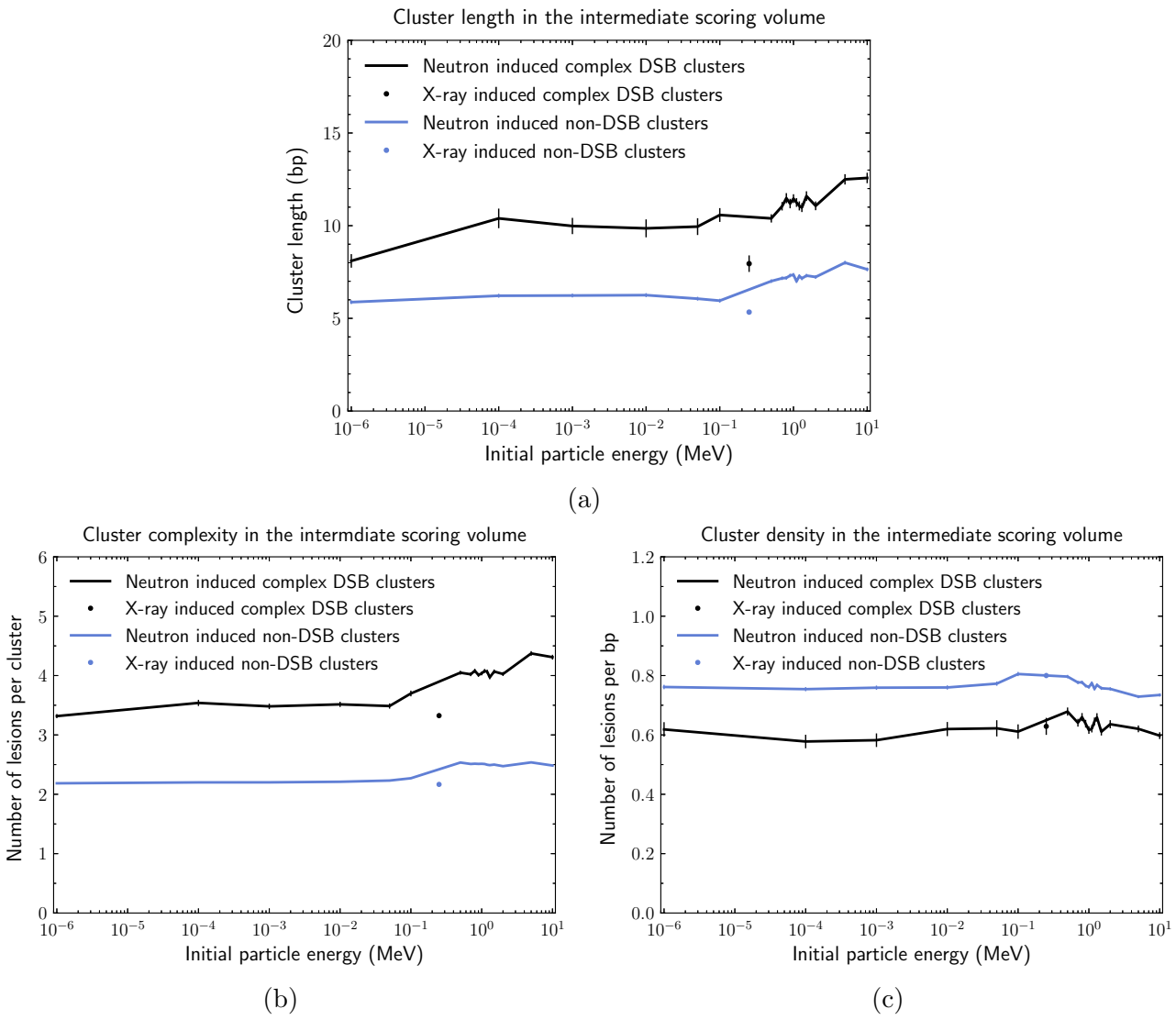


Figure 8: Clustered DNA damage properties of interest as a function of initial particle energy in the intermediate scoring volume: (a) cluster length, (b) cluster complexity, and (c) damage density per cluster. Plotted values are the mean values obtained over 100 statistically-independent simulations. Results for 18 initial neutron energies are plotted as a line alongside the individual result for the reference 250 keV x-ray radiation. Results for complex DSB clusters are plotted in black, while results for non-DSB clusters are plotted in blue. Error bars represent the standard deviation of the mean.

At low neutron energies below about 100 keV, there was no significant difference in the average complex DSB cluster length between neutrons and the reference x-rays. However, the average cluster length increased by about 30–50 %, from 8 bp at 1 eV to a maximum of ∼12 bp at 10 MeV. Similarly, the average complex DSB cluster complexity increased by about 30 % from ∼3.5 lesions per cluster at 1 eV to a maximum of ∼4.5 lesions per cluster at 5 MeV. These findings were anticipated because of the onset of the dominance of proton relative dose contributions for neutrons above 100 keV and their higher LET than electrons. Given that the results for cluster length and

complexity exhibited similar energy dependence, the lack of significant trend in average damage density per complex DSB was expected.

On average, complex DSB clusters were 30–50 % longer than non-DSB clusters across all neutron energies. Similarly, complex DSB clusters contained 40–70 % more lesions than non-DSB clusters on average. This result was expected and can largely be explained by considering that complex DSB clusters had by definition a minimum of three lesions whereas non-DSB clusters had a minimum of two. Non-DSB clusters were approximately 25 % more dense than complex DSB clusters, which follows from the relative trends in cluster length and complexity.

In summary, although neutrons above 100 keV tend to produce more clusters, longer clusters, and more complex clusters than the reference x-ray radiation, they do not produce clusters with a higher density of lesions. There is a clear neutron energy dependence in cluster length and complexity relative to x-rays, even though the trend does not precisely match the characteristic energy dependence of neutron RBE for stochastic effects. These parameters should thus be investigated further when modeling the repair of clustered DNA lesions.

3.5 Dose and parameter sensitivity analysis

3.5.1 Dose

The impact of varying the target delivered dose D_0 on the results obtained for complex DSB clusters in the intermediate scoring volume is shown in figure 9. The yield of neutron-induced complex DSB clusters was found to increase linearly with dose, as evidenced by the overlap of the dose-normalized curves in figure 9(a). Correspondingly, there was no dose dependence in neutron RBE for inducing complex DSB clusters.

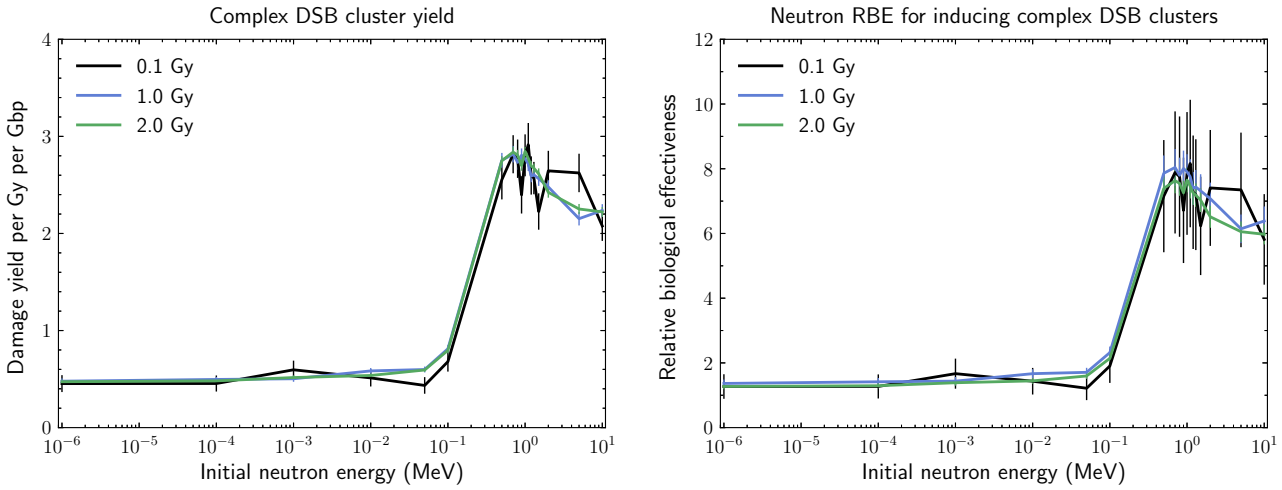
These findings indicate that a given cluster of DNA damage is typically caused by single-event action (i.e. a single primary track and its secondary tracks) within this dose range. Thus, our results should hold qualitatively in the low-dose regime because single event action scales linearly with dose. Our results at 0.1 Gy, nominally the upper limit of the low-dose regime, are demonstrative of this expectation.

3.5.2 Clustering distance

The impact of varying the clustering distance on the results for complex DSB clusters in the intermediate scoring volume is shown in figure 10.

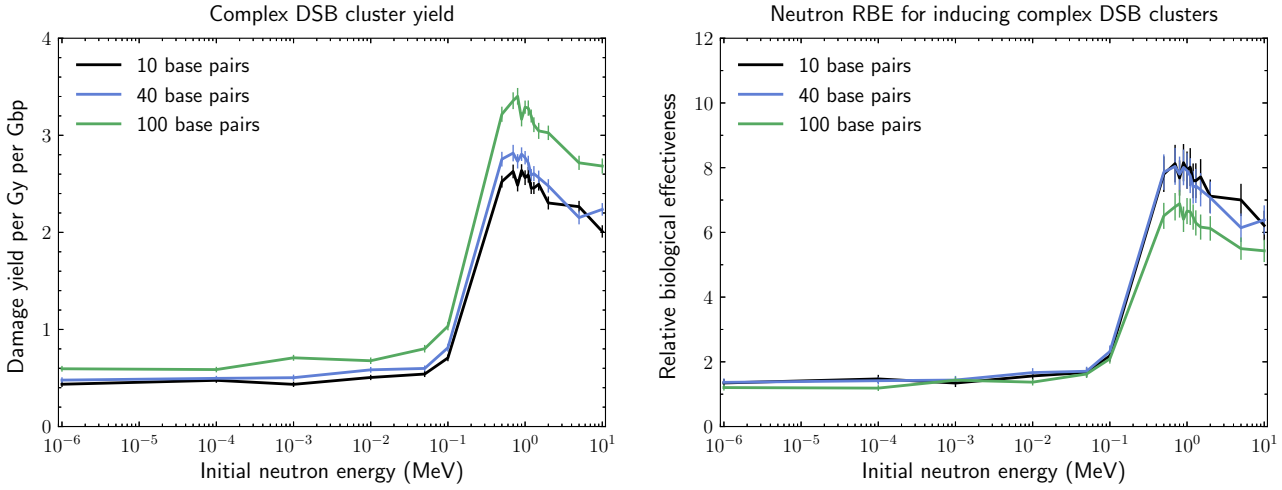
Comparison of the DNA damage yields in figure 10(a) revealed no significant difference for clustering distances of 10 and 40 bp. This was expected, given that the average complex DSB cluster length was found to be in the vicinity of 10 bp for all neutron energies (figure 8(a)). A small but noticeable increase in yield was obtained by increasing the clustering distance to 100 bp. There are two competing effects to consider when increasing the clustering distance:

(i) Formation of new clusters by combining lesions that would otherwise be treated as isolated lesions.



(a) (b)

Figure 9: Impact of varying the total delivered dose D_0 on (a) neutron-induced complex DSB cluster yield and (b) neutron RBE for inducing complex DSB clusters. Plotted values are the mean values obtained over 100 statistically independent simulations. Error bars represent the standard deviation of the mean.



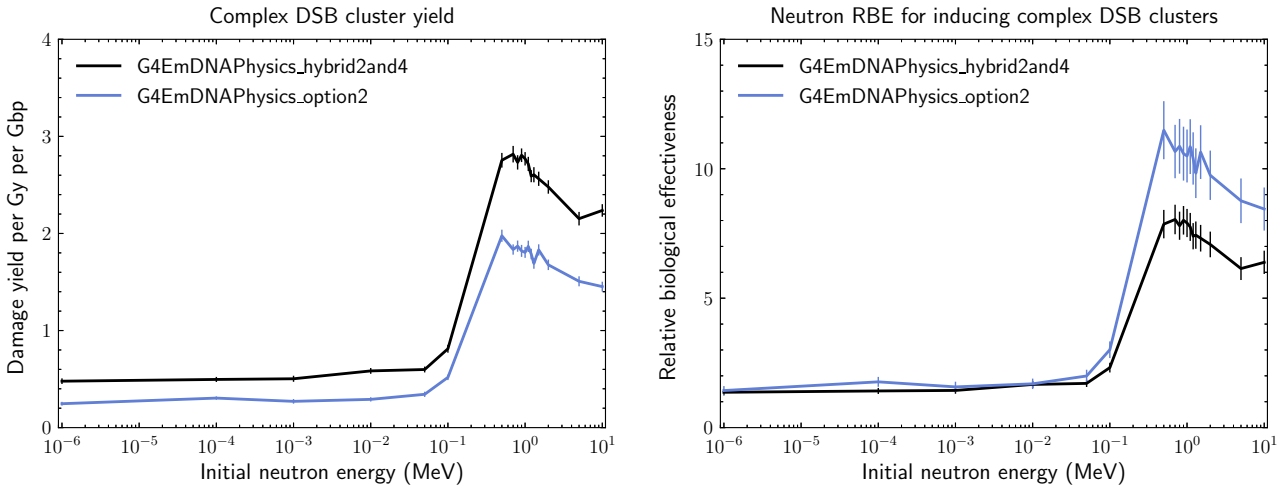
(a) (b)

Figure 10: Impact of varying the DNA damage clustering distance on (a) neutron-induced complex DSB cluster yield and (a) neutron RBE for inducing complex DSB clusters. Plotted values are the mean values obtained over 100 statistically independent simulations. Error bars represent the standard deviation of the mean.

(ii) Merging of adjacent clusters that would otherwise be treated as independent clusters.

Given that the clustered damage yields increased, the predominant effect must be the former.

Comparison of neutron RBE in figure 10(b) revealed that RBE decreased by increasing the clustering distance. This result may be explained by the relative LET properties of neutrons and



(a) (b)

Figure 11: Impact of varying the simulation physics constructor, and thus the underlying physics models, on (a) neutron-induced complex DSB cluster yields and (b) neutron RBE for inducing complex DSB clusters. Plotted values are the mean values obtained over 100 statistically independent simulations. Error bars represent the standard deviation of the mean.

x-rays. X-rays and their secondary electrons are low LET particles, which tend to produce isolated lesions. Meanwhile, neutrons are higher LET particles that tend to produce clusters of lesions. Thus, the previous observation that increasing the clustering distance tended to preferentially combine isolated lesions rather than adjacent clusters would result in relatively more x-ray-induced clusters than neutron-induced clusters. The net result is a decrease in neutron RBE, as we have observed in this work.

Overall, our results were found to be qualitatively robust over the range of DNA damage clustering distances identified in our literature review.

3.5.3 Physics constructor

The impact of varying the physics models that govern sub-10 keV electron transport is shown for complex DSB clusters in the intermediate scoring volume in figure 11. The yield was significantly larger using our hybrid constructor compared to the opt2 constructor for all neutron energies. This can be explained by expanding upon previous discussions by Kyriakou *et al* (2015), Bordes *et al* (2017), and Zhu *et al* (2020b). The physics models in opt2 are known to result in longer, more diffusive electron tracks than expected physically. This behaviour was "corrected" in the newer physics models included in the opt4 constructor (Kyriakou *et al* 2015) and our hybrid constructor. Thus, we expected a higher yield of clustered DNA damage induced by electrons modeled with opt4 compared to opt2, which is exactly what we observed using our hybrid constructor.

Comparison of neutron RBE in figure 11(b) revealed that the peak in RBE was higher for the opt2 constructor compared to our hybrid constructor. The underlying x-ray-induced complex DSB yields per Gy per Gbp for our hybrid constructor and opt2 were 0.37 ± 0.02 and 0.18 ± 0.02,

respectively. The relative difference in neutron-induced and x-ray-induced yields was larger for opt2 than our hybrid constructor, which resulted in the larger peak RBE.

Overall, this comparison of physics constructors indicated that the physics models used in the simulations have an impact on the quantitative accuracy of the results but not on the qualitative trends.

3.6 Limitations and future work

In this section we discuss some of the limitations of this work and, where appropriate, how they will be handled in future work.

3.6.1 Limitations of our nuclear DNA model

We have developed a new nuclear DNA model in a field where many models already exist and there is a lack of standardization (Schuemann *et al* 2019a). However, we attempted to atone for this by releasing our code under an open-source license.

Additionally, our nuclear model is cubic in shape, which does not align with the generally ellipsoidal shape of fibroblast nuclei (Seaman *et al* 2015). However, we believe our cubic geometry was justified given that the most important factors pertaining to DNA damage yields are (i) the overall density of DNA base pairs in the model and (ii) the size of the sensitive DNA volumes (Lampe *et al* 2018). Moreover, a cubic shape was the most computationally efficient to generate and the simplest to code. Ultimately, our benchmarking analysis indicated that our model was able to produce DNA damage yields that were consistent with previous studies.

Finally, the chromatin fibres in our model were not connected to each other, which limited our ability to combine lesions in adjacent fibres into a cluster. However, we assert that the impact of this limitation was small given that:

(i) Clusters were, on average, approximately 10 bp in length.

(ii) Only a few thousand DNA damage lesions were induced per Gy over the entire genome of ∼6 Gbp (figure 6(a)).

3.6.2 Limitations in physical modeling

Our inability to simulate secondary particle tracks heavier than alpha particles was a shortcoming of this work. Although previous studies have shown success in ignoring the effects of heavy ions for incident neutrons with energies up to 14 MeV (Zabihi *et al* 2020), we found that heavy ions contributed up to 14 % of the dose for 10 MeV neutrons in the outer scoring volume (Lund *et al* 2020). Nevertheless, we have already calculated the energy spectra and relative dose contributions of these particles in our previous work (Lund *et al* 2020). Our methodology described in this manuscript is robust enough to handle their inclusion once physical models that describe their transport have been incorporated into the Geant4-DNA and TOPAS-nBio frameworks.

Other physical factors that may have impacted our results include:

- The treatment of electrons with energies greater than 1 MeV as uncorrelated lower energy electron tracks.
- The treatment of the entire simulation volume as liquid water.

The impact of these assumptions is difficult to predict. However, our methodology and code can facilitate rapid reassessment of these assumptions as new features become available in Geant4-DNA and TOPAS-nBio.

3.6.3 Limitations in modeling radiation-induced biological effects

In this study we only considered the initial landscape of direct DNA damage induced by radiation. Our research group is currently developing an update to our custom TOPAS-nBio application that incorporates the indirect effects of radiation action and will conduct a follow-up study using these updates. It is also of interest to consider how DNA repair might affect our results in order to bridge the gap between DNA damage and mutagenesis. Other groups have made great strides in modeling such DNA repair, particularly the repair of SSBs and DSBs (Friedland *et al* 2010, McMahon *et al* 2016, Zhu *et al* 2020a). We aim to to expand on our work by applying repair models such as these to our simulated DNA damage and analyzing the impact on the resulting neutron RBE.

3.7 Open-source code release

Our complete Monte Carlo application was developed as a TOPAS extension (Perl *et al* 2012, Faddegon *et al* 2020) for use with TOPAS-nBio (Schuemann *et al* 2019b) and we have released it under an open-source license (Montgomery *et al* 2021). The following features are included:

- A nuclear DNA model implemented as a custom TOPAS geometry component.
- A physics constructor implemented as a custom TOPAS physics module.
- An algorithm for calculating clustered DNA damage yields implemented as a custom TOPAS scorer.
- A TOPAS-style parameter file to control the simulation.
- Our neutron and x-ray secondary particle energy spectra in TOPAS parameter file format.
- Our neutron and x-ray secondary particle relative dose contributions in TOPAS parameter file format.

Simulations may be readily run and configured by the user via the included parameter file. The custom components written in C++ have been extensively documented to improve readability and may be adopted or modified by users as needed. Installation requirements and instructions are provided with the code.

4 Conclusion

We have investigated the biophysical mechanisms underlying the energy dependence of neutron RBE for stochastic effects by simulating neutron-induced direct clustered DNA damage in a geometric DNA model. We found that neutron RBE for inducing clusters of DNA damage, both with and without DSBs, exhibited similar energy dependence to the ICRP's neutron radiation weighting factors and the US NRC's neutron quality factors. Our results support the hypothesis that a variety of clustered DNA damage lesions give rise to the energy dependence of neutron RBE for stochastic effects. We also identified an energy dependence in the average length and complexity of DNA damage clusters, indicating that these parameters should be considered when modeling mutagenic effects. Repeated simulations with variations in key parameters demonstrated the robustness of both our methodology and results, including their applicability to the low-dose regime. Our custom TOPAS-nBio application has been released under an open source license to enable others to independently validate our work and expand on it. In the future our aim is to incorporate indirect DNA damage effects and DNA repair models into our application.

5 Acknowledgements

This work was supported by the Natural Sciences and Engineering Research Council of Canada (Alexander Graham Bell Canada Graduate Scholarship-Doctoral: L. Montgomery, Discovery Grant: J. Kildea). Additional funding was provided to L. Montgomery from the McGill Faculty of Medicine to J. Kildea's lab from the Canada Foundation for Innovation John R. Evans Leaders Fund. We acknowledge the support of the Canadian Space Agency (CSA) (19FAMCGB25). Logistical support provided by the Prince Edward Island Cancer Treatment Centre and the Medical Physics Department at the McGill University Health Centre was invaluable for this work.

The authors would like to thank Joseph Perl for providing early access to the TOPAS source code, which informed various decisions when designing our code. We would also like to thank Nathanael Lampe for insightful discussion into the construction of complex geometry using Geant4.

All of the plots presented in this manuscript were generated using the open-source Matplotlib Python library (Hunter 2007).

References

Baiocco G *et al* 2016 The origin of neutron biological effectiveness as a function of energy *Scientific Reports* **6** 34033
URL http://dx.doi.org/10.1038/srep34033

Benton E R, Benton E V and Frank A L 2001 Neutron dosimetry in low-earth orbit using passive detectors *Radiation Measurements* **33**(3) 255–263
URL http://dx.doi.org/10.1016/S1350-4487(01)00047-6

Bernal M A, DeAlmeida C E, Incerti S, Champion C, Ivanchenko V and Francis Z 2015 The influence of DNA configuration on the direct strand break yield *Computational and Mathematical Methods in Medicine* 417501
URL http://dx.doi.org/10.1155/2015/417501

Bordes J, Incerti S, Lampe N, Bardiès M and Bordage M C 2017 Low-energy electron dose-point kernel simulations using new physics models implemented in Geant4-DNA *Nuclear Instruments and Methods in Physics Research, Section B: Beam Interactions with Materials and Atoms* **398** 13–20
URL http://dx.doi.org/10.1016/j.nimb.2017.02.044

Charlton D E and Humm J L 1988 A method of calculating initial DNA Strand breakage following the decay of incorporated 125i *International Journal of Radiation Biology* **53**(3) 353–365
URL http://dx.doi.org/10.1080/09553008814552501

Faddegon B, Ramos-Méndez J, Schuemann J, McNamara A, Shin J, Perl J and Paganetti H 2020 The TOPAS tool for particle simulation, a Monte Carlo simulation tool for physics, biology and clinical research *Physica Medica* **72** 114–121
URL http://dx.doi.org/10.1016/j.ejmp.2020.03.019

Friedland W, Dingfelder M, Kundrát P and Jacob P 2011 Track structures, DNA targets and radiation effects in the biophysical Monte Carlo simulation code PARTRAC *Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis* **711**(1-2) 28–40
URL http://dx.doi.org/10.1016/j.mrfmmm.2011.01.003

Friedland W, Jacob P and Kundrát P 2010 Stochastic simulation of DNA double-strand break repair by non-homologous end joining based on track structure calculations *Radiation Research* **173**(5) 677–688
URL http://dx.doi.org/10.1667/RR1965.1

Georgakilas A G, O'Neill P and Stewart R D 2013 Induction and Repair of Clustered DNA Lesions: What Do We Know So Far? *Radiation Research* **180**(1) 100–109
URL http://dx.doi.org/10.1667/rr3041.1

Ghosh S K and Jost D 2018 How epigenome drives chromatin folding and dynamics, insights from efficient coarse-grained models of chromosomes *PLoS Computational Biology* **14**(5) 1–26
URL http://dx.doi.org/10.1371/journal.pcbi.1006159

Goodhead D T 1994 Initial events in the cellular effects of ionizing radiations: Clustered damage in DNA *International Journal of Radiation Biology* **65**(1) 7–17
URL http://dx.doi.org/10.1080/09553009414550021

Hanahan D and Weinberg R A 2000 The Hallmarks of Cancer *Cell* **100** 57–70
URL http://dx.doi.org/10.1016/S0092-8674(00)81683-9

Hanahan D and Weinberg R A 2011 Hallmarks of cancer: The next generation *Cell* **144**(5) 646–674
URL http://dx.doi.org/10.1016/j.cell.2011.02.013

Harbron R W 2012 Cancer risks from low dose exposure to ionising radiation - Is the linear no-threshold model still relevant? *Radiography* **18**(1) 28–33
URL http://dx.doi.org/10.1016/j.radi.2011.07.003

Howell R M, Hertel N E, Wang Z, Hutchinson J and Fullerton G D 2006 Calculation of effective dose from measurements of secondary neutron spectra and scattered photon dose from dynamic MLC IMRT for 6 MV, 15 MV, and 18 MV beam energies *Medical Physics* **33**(2) 360–368
URL http://dx.doi.org/10.1118/1.2140119

Hunter J D 2007 Matplotlib: A 2D graphics environment *Computing in Science and Engineering* **9**(3) 90–95
URL http://dx.doi.org/10.1109/MCSE.2007.55

ICRP 2003 ICRP 92: Relative Biological Effectiveness (RBE), Quality Factor (Q), and Radiation Weighting Factor (wR) *Annals of the ICRP* **33**(4)

ICRP 2007 ICRP 103: The 2007 Recommendations of the ICRP *Annals of the ICRP* **37**(2-4)

Incerti S *et al* 2010a The Geant4-DNA project *International Journal of Modeling, Simulation, and Scientific Computing* **1**(2) 157
URL https://doi.org/10.1142/S1793962310000122

Incerti S, Douglass M, Penfold S, Guatelli S and Bezak E 2016 Review of Geant4-DNA applications for micro and nanoscale simulations *Physica Medica* **32**(10) 1187–1200
URL http://dx.doi.org/10.1016/j.ejmp.2016.09.007

Incerti S *et al* 2010b Comparison of GEANT4 very low energy cross section models with experimental data in water *Medical Physics* **37**(9) 4692–4708
URL http://dx.doi.org/10.1118/1.3476457

Incerti S *et al* 2018 Geant4-DNA example applications for track structure simulations in liquid water: A report from the Geant4-DNA Project *Medical Physics* **45**(8) e722–e739
URL http://dx.doi.org/10.1002/mp.13048

Iyer R and Lehnert B E 2000 Effects of ionizing radiation in targeted and nontargeted cells *Archives of Biochemistry and Biophysics* **376**(1) 14–25
URL http://dx.doi.org/10.1006/abbi.1999.1684

Joint Committee for Guides in Metrology 2008 JCGM 100: Evaluation of measurement data — Guide to the expression of uncertainty in measurement Technical report Working Group 1 of the Joint Committee for Guides in Metrology
URL http://www.bipm.org/en/publications/guides/gum.html

Koshiishi H, Matsumoto H, Chishiki A, Goka T and Omodaka T 2007 Evaluation of the neutron radiation environment inside the International Space Station based on the Bonner Ball Neutron Detector experiment *Radiation Measurements* **42**(9) 1510–1520
URL http://dx.doi.org/10.1016/j.radmeas.2007.02.072

Kry S F *et al* 2017 AAPM TG 158: Measurement and calculation of doses outside the treated volume from external-beam radiation therapy *Medical Physics* **44**(10) e391–e429
URL http://dx.doi.org/10.1002/mp.12462

Kyriakou I, Incerti S and Francis Z 2015 Technical note: Improvements in geant 4 energy-loss model and the effect on low-energy electron transport in liquid water *Medical Physics* **42**(7) 3870–3876
URL http://dx.doi.org/10.1118/1.4921613

Lampe N, Karamitros M, Breton V, Brown J M, Sakata D, Sarramia D and Incerti S 2018 Mechanistic DNA damage simulations in Geant4-DNA Part 2: Electron and proton damage in a bacterial cell *Physica Medica* **48** 146–155
URL http://dx.doi.org/10.1016/j.ejmp.2017.12.008

Lieberman-Aiden E *et al* 2009 Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome *Science* **326**(5850) 289–293
URL https://doi.org/10.1126/science.1181369

Little J B 2000 Radiation carcinogenesis *Carcinogenesis* **21**(3) 397–404
URL http://dx.doi.org/10.1093/carcin/21.3.397

Lund C M, Famulari G, Montgomery L and Kildea J 2020 A microdosimetric analysis of the interactions of mono-energetic neutrons with human tissue *Physica Medica* **73** 29–42
URL http://dx.doi.org/10.1016/j.ejmp.2020.04.001

Maglieri R, Licea A, Evans M, Seuntjens J and Kildea J 2015 Measuring neutron spectra in radiotherapy using the nested neutron spectrometer *Medical Physics* **42**(11) 6162–6169
URL http://dx.doi.org/10.1118/1.4931963

Magnander K and Elmroth K 2012 Biological consequences of formation and repair of complex DNA damage *Cancer Letters* **327**(1-2) 90–96
URL http://dx.doi.org/10.1016/j.canlet.2012.02.013

Martin R F and Haseltine W A 1981 Range of Radiochemical Damage to DNA with Decay of Iodine-125 *Science* **213**(4510) 896–898
URL http://dx.doi.org/10.1126/science.7256283

McMahon S J, Schuemann J, Paganetti H and Prise K M 2016 Mechanistic Modelling of DNA Repair and Cellular Survival Following Radiation-Induced DNA Damage *Scientific Reports* **6** 1–14
URL http://dx.doi.org/10.1038/srep33290

McNamara A L *et al* 2018 Geometrical structures for radiation biology research as implemented in the TOPAS-nBio toolkit *Physics in Medicine and Biology* **63**(17) 175018
URL http://dx.doi.org/10.1088/1361-6560/aad8eb

Meylan S, Incerti S, Karamitros M, Tang N, Bueno M, Clairand I and Villagrasa C 2017 Simulation of early DNA damage after the irradiation of a fibroblast cell nucleus using Geant4-DNA *Scientific Reports* **7** 11923
URL http://dx.doi.org/10.1038/s41598-017-11851-4

Mokari M, Moeini H, Soleimani M and Fereidouni E 2020 Calculation and comparison of the direct and indirect DNA damage induced by low energy electrons using default and CPA100

cross section models within Geant4-DNA *Nuclear Instruments and Methods in Physics Research, Section B: Beam Interactions with Materials and Atoms* **480** 56–66
URL http://dx.doi.org/10.1016/j.nimb.2020.08.011

Montgomery L, Lund C M, Manalad J, Landry A and Kildea J 2021 TOPAS Clustered DNA Damage
URL http://dx.doi.org/10.5281/zenodo.5090104

Morgan W F and Bair W J 2013 Issues in Low Dose Radiation Biology: The Controversy Continues. A Perspective *Radiation Research* **179**(5) 501–510
URL http://dx.doi.org/10.1667/rr3306.1

Mullenders L, Atkinson M, Paretzke H, Sabatier L and Bouffler S 2009 Assessing cancer risks of low-dose radiation *Nature Reviews Cancer* **9**(8) 596–604
URL http://dx.doi.org/10.1038/nrc2677

Nikitaki Z, Nikolov V, Mavragani I V, Plante I, Emfietzoglou D, Iliakis G and Georgakilas A G 2016 Non-DSB clustered DNA lesions. Does theory colocalize with the experiment? *Radiation Physics and Chemistry* **128** 26–35
URL http://dx.doi.org/10.1016/j.radphyschem.2016.06.020

Nikjoo H, O'Neill P, Goodhead D T and Terrissol M 1997 Computational modelling of low-energy electron-induced DNA damage by early physical and chemical events *International Journal of Radiation Biology* **71**(5) 467–483
URL doi.org/10.1080/095530097143798

Nikjoo H, O'Neill P, Wilson W E and Goodhead D T 2001 Computational approach for determining the spectrum of DNA damage induced by ionizing radiation *Radiation Research* **156** 577–583
URL http://dx.doi.org/10.1667/0033-7587(2001)156[0577:cafdts]2.0.co;2

Nikjoo H, Uehara S, Wilson W E, Hoshi M and Goodhead D T 1998 Track structure in radiation biology: theory and applications *International Journal of Radiation Biology* **73**(4) 355–364

Pater P, Seuntjens J, El Naqa I and Bernal M A 2014 On the consistency of Monte Carlo track structure DNA damage simulations *Medical Physics* **41**(12) 1–9
URL http://dx.doi.org/10.1118/1.4901555

Perl J, Hall D, Shin J, Schuemann J and Ramos-Méndez J 2016 Modular Physics Lists
URL https://topas.readthedocs.io/en/latest/parameters/physics/modular.html

Perl J, Shin J, Schümann J, Faddegon B and Paganetti H 2012 TOPAS: An innovative proton Monte Carlo platform for research and clinical applications *Medical Physics* **39**(11) 6818–6837
URL http://dx.doi.org/10.1118/1.4758060

Sage E and Shikazono N 2017 Radiation-induced clustered DNA lesions: Repair and mutagenesis *Free Radical Biology and Medicine* **107** 125–135
URL http://dx.doi.org/10.1016/j.freeradbiomed.2016.12.008

Sakata D *et al* 2020 Fully integrated Monte Carlo simulation for evaluating radiation induced DNA damage and subsequent repair using Geant4-DNA *Scientific Reports* **10**(1) 1–13
URL http://dx.doi.org/10.1038/s41598-020-75982-x

Sato T *et al* 2013 Particle and heavy ion transport code system, PHITS, version 2.52 *Journal of Nuclear Science and Technology* **50**(9)
URL http://dx.doi.org/10.1080/00223131.2013.814553

Schuemann J, McNamara A L, Ramos-Méndez J, Perl J, Held K D, Paganetti H, Incerti S and Faddegon B 2019a TOPAS-nBio: An Extension to the TOPAS Simulation Toolkit for Cellular and Sub-cellular Radiobiology *Radiation Research* **191**(2) 125–138
URL http://dx.doi.org/10.1667/rr15226.1

Schuemann J *et al* 2019b A New Standard DNA Damage (SDD) Data Format *Radiation Research* **191**(1) 76–92
URL http://dx.doi.org/10.1667/rr15209.1

Seaman L, Meixner W, Snyder J and Rajapakse I 2015 Periodicity of nuclear morphology in human fibroblasts *Nucleus* **6**(5) 408–416
URL http://dx.doi.org/10.1080/19491034.2015.1095432

Shuryak I *et al* 2020 A High Throughput Approach to Reconstruct Partial-Body and Neutron Radiation Exposures on an Individual Basis *Scientific Reports* **10**(1) 1–14
URL http://dx.doi.org/10.1038/s41598-020-59695-9

Śmiałek M A, Jones N C, Hoffmann S V and Mason N J 2013 Measuring the density of DNA films using ultraviolet-visible interferometry *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics* **87**(6) 1–4
URL http://dx.doi.org/10.1103/PhysRevE.87.060701

Swarts S G, Sevilla M D, Becker D, Tokar C J and Wheeler K T 1992 Radiation-Induced DNA Damage as a Function of Hydration. I. Release of Unaltered bases *Radiation Research* **129**(3) 333–344
URL https://doi.org/10.2307/3578034

Tajik-Mansoury M A, Rajabi H and Mozdarani H 2017 A comparison between track-structure, condensed-history Monte Carlo simulations and MIRD cellular S-values *Physics in Medicine and Biology* **62**(5) N90–N106
URL http://dx.doi.org/10.1088/1361-6560/62/5/N90

Tang N, Bueno M, Meylan S, Incerti S, Tran H N, Vaurijoux A, Gruel G and Villagrasa C 2019 Influence of chromatin compaction on simulated early radiation-induced DNA damage using Geant4-DNA *Medical Physics* **46**(3) 1501–1511
URL http://dx.doi.org/10.1002/mp.13405

US NRC 2021 Units of Radiation Dose *Code of Federal Regulations* **10**(20.1004) 298–300
URL https://www.nrc.gov/reading-rm/doc-collections/cfr/part020/part020-1004.html

Van der Schans G P 1978 Gamma-ray induced double-Strand breaks in DNA resulting from randomly-inflicted single-strand breaks: Temporal local denaturation, a new radiation phenomenon? *International Journal of Radiation Biology* **33**(2) 105–120
URL http://dx.doi.org/10.1080/09553007814550011

Vaux D L 2011 In defense of the somatic mutation theory of cancer *BioEssays* **33**(5) 341–343
URL http://dx.doi.org/10.1002/bies.201100022

Villagrasa C, Meylan S, Gonon G, Gruel G, Giesen U, Bueno M and Rabus H 2017 Geant4-DNA simulation of DNA damage caused by direct and indirect radiation effects and comparison with biological data. *EPJ Web of Conferences* **153** 1–6
URL http://dx.doi.org/10.1051/epjconf/201715304019

Ward J F 1995 Radiation Mutagenesis : The Initial DNA Lesions Responsible *Radiation Research* **142**(3) 362–368
URL http://dx.doi.org/10.2307/3579145

Watanabe R, Rahmanian S and Nikjoo H 2015 Spectrum of Radiation-Induced Clustered Non-DSB Damage – A Monte Carlo Track Structure Modeling and Calculations *Radiation Research* **183**(5) 525–540
URL http://dx.doi.org/10.1667/rr13902.1

White D R, Booz J, Griffith R V, Spokas J J and Wilson I J 1989 Report 44 *Journal of the International Commission on Radiation Units and Measurements* **os23**(1)
URL http://dx.doi.org/10.1093/jicru/os23.1.report44

Zabihi A, Tello J, Incerti S, Francis Z, Forozani G, Semsarha F, Moslehi A and Bernal M A 2020 Determination of fast neutron RBE using a fully mechanistic computational model *Applied Radiation and Isotopes* **156** 108952
URL http://dx.doi.org/10.1016/j.apradiso.2019.108952

Zhu H *et al* 2020a Cellular Response to Proton Irradiation: A Simulation Study with TOPAS-nBio *Radiation Research* **194**(1) 9–21
URL http://dx.doi.org/10.1667/rr15531.1

Zhu H *et al* 2020b A parameter sensitivity study for simulating DNA damage after proton irradiation using TOPAS-nBio. *Physics in medicine and biology* **65**(8) 085015
URL http://dx.doi.org/10.1088/1361-6560/ab7a6b