

1           **A transgenerational mutational signature from ionizing radiation exposure.**

2  
3           Fabian Brand<sup>1</sup>, Hannah Klinkhammer<sup>1,3</sup>, Alexej Knaus<sup>1</sup>, Manuel Holtgrewe<sup>2</sup>, Leonie  
4           Weinhold<sup>3</sup>, Dieter Beule<sup>2</sup>, Kerstin Ludwig<sup>4</sup>, Prachi Kothiyal<sup>5</sup>, George Maxwell<sup>5</sup>, Markus  
5           Noethen<sup>4</sup>, Matthias Schmid<sup>3</sup>, Karl Sperling<sup>6</sup>, Peter Krawitz<sup>1</sup>

6  
7           <sup>1</sup>Institute of Genomics Statistics and Bioinformatics, School of Medicine, University  
8           Hospital Bonn & University of Bonn, Germany

9           <sup>2</sup>Core Unit Bioinformatics, Berlin Institute of Health, Berlin, Germany

10          <sup>3</sup>Institute of Medical Biometry, Informatics and Epidemiology, University Hospital Bonn,  
11          Bonn, Germany

12          <sup>4</sup>Institute of Human Genetics, School of Medicine, University Hospital Bonn & University of  
13          Bonn, Germany

14          <sup>5</sup>Inova Translational Medicine Institute, Inova Health System, Falls Church, VA, USA

15          <sup>6</sup>Institute of Medical and Human Genetics, Charité-Universitätsmedizin Berlin, Germany

16  
17  
18          **Abstract**

19  
20          Effects of accidental radiation exposure on the human germline remain a topic of medical  
21          interest. Sequencing technology now allows the accurate analysis of the entire human  
22          genome in potentially exposed cohorts to search for possible signatures of ionizing radiation  
23          exposure. Clustered *de novo* mutations (cDNMs), that are multiple DNA lesions within 20 bp  
24          of each other, have been hypothesized to be a signature for paternal exposure to ionizing  
25          radiation. We analyzed whole genome trio data of 240 children and their likely exposed  
26          parents as well as 1,275 offspring from unexposed families. The cohort of exposed children  
27          consists of 130 offspring of Chernobyl cleanup workers (CRU, exposure range = 0-4,080  
28          mSv) and 110 offspring of former soldiers from both German armies which have likely been  
29          irradiated during their service (Radar cohort, exposure range = 0-353 mSv). For the Radar  
30          cohort, a retrospective dosage estimation was conducted based on the service records from  
31          each soldier, but it remains a challenging task to retrospectively estimate received dose  
32          accurately. In agreement with previous epidemiological data, no transgenerational effects  
33          could be detected for the class of isolated *de novo* mutations, which cause the majority of  
34          genetic disorders. We found that the number of isolated *de novo* mutations increases by 2%  
35          per year of age of the father at conception, which is also in line with previous works.  
36          On average, we observed 2.65 cDNMs per offspring in the CRU cohort, 1.48 in the Radar  
37          cohort and 0.88 in the control cohort, with a median of two clustered mutations per genome  
38          in offspring of irradiated fathers. This is a significant increase ( $p < 0.005$ ) in the number of  
39          cDNMs compared to a set of age-matched controls. Furthermore, we show that the cDNM  
40          rates scale with paternal exposure to ionizing radiation ( $p < 0.001$ ). Our findings  
41          corroborate that clustered *de novo* mutations represent a transgenerational biomarker for  
42          paternal exposure to ionizing radiation.

## 44 **Introduction**

45

46 Somatic mutational signatures can point to endogenous or exogenous mutagen exposure and  
47 have become an important biomarker to determine the cause and underlying processes in  
48 human cancers<sup>1</sup>. Exposure to ultraviolet light, for instance, results in pyrimidine dimers that  
49 can be removed by nucleotide excision repair on the transcribed strand. Consequently, a  
50 higher prevalence of C>T mutations on the untranscribed strand can be observed in e.g.  
51 malignant melanoma. Some signatures also result from anticancer drugs that are mutagens.  
52 For instance, systemic treatment with platinum-based drugs also affects the germline and can  
53 result in hypermutation in offspring<sup>2</sup>.

54 While ultraviolet light is too weak to penetrate into cells of the germline, during oogenesis  
55 and spermatogenesis higher energy radiation could potentially induce DNA damage that is  
56 passed on to the next generation. Changes to human DNA in the form of chromosomal  
57 inversions and translocations have previously been found after long duration space flight  
58 onboard the International Space Station where astronauts were subjected to a dose of 146.34  
59 mSv of high-linear energy transfer (LET) space radiation<sup>3</sup>. This environment has also been  
60 found to increase the cancer risk for astronauts following prolonged stays in orbit<sup>4,5</sup>. On  
61 earth, gamma rays with energies ranging from a few kiloelectronvolts (keV) to  
62 mega-electronvolts (MeV) occur naturally from various atmospheric interactions with cosmic  
63 particles, or from the decay of radioisotopes and accumulate to an annual whole body dose of  
64 1-2 mSv<sup>6</sup>. Singular medical radiation exposure e.g. during computer tomography is in the  
65 same order of magnitude in the target region, but the gonads are usually shielded<sup>6</sup>.  
66 Unintentional, high exposure to man-made sources of gamma rays can happen during nuclear  
67 disasters or in certain military occupations (e.g. operating high powered radar devices) where  
68 stray radiation is emitted, and personnel is insufficiently protected<sup>7</sup>. However, since  
69 retrospective dosage estimations are difficult, biomarkers indicating prior radiation injury  
70 have been searched for decades with the respective methods available at that time.

71 While high-energy particles can cause double-strand breaks through direct impacts on the  
72 DNA, the primary damage is often a result from the generation of reactive oxygen species  
73 (ROS) following ionization of nearby water molecules in the cell<sup>8-10</sup>. Within a few helix  
74 turns of the DNA, this ionization can result in multiple single-strand breaks, abasic sites, and  
75 oxidized purine or pyrimidine bases. In contrast to damage caused by endogenous oxidative  
76 stress, clustered DNA lesions from ionizing radiation are less likely to be repaired due to their  
77 complexity and are thus more likely to persist after cell division<sup>11</sup>. Acute exposure to high  
78 doses (> 4Gy) of ionizing radiation has been studied extensively in mice, where an increase  
79 in *de novo* mutations within a few helix turns of DNA have been found to be associated with  
80 parental exposure to ionizing radiation, alongside structural changes to the genome<sup>12,13</sup>.  
81 Recently, Matsuda et al. demonstrated a dose-dependency of the increase of clustered DNA  
82 damage and structural changes to the genome by analyzing data from mouse hematopoietic  
83 stem cells following exposure to 3.8, 7.5 and 7.7 Gy<sup>14</sup>. Multiple damaged sites affecting  
84 nucleotides in close vicinity are thus a characteristic signature from ionizing radiation that  
85 could be passed on to the next generation when occurring in the germline<sup>12</sup>.

86

87

## 88 **Effects of ionizing radiation on human DNA**

89

90 Until now, such studies have not been replicated with human data, mainly because of the lack  
91 of cohorts that are suited to trio whole genome sequencing analysis and have been exposed to  
92 elevated doses of ionizing radiation. One notable event that caused widespread irradiation in  
93 humans and animals is the Chernobyl nuclear accident. The first analyses after the Chernobyl

94 accident focused on mutation rates for minisatellites, microsatellites, and tandem repeat  
95 expansions and found a two-fold increase<sup>15,16</sup>. However, due to the small cohort size,  
96 potential confounders, and the small number of loci that could be investigated, the statistical  
97 significance of these results is still debated<sup>17</sup>. With the availability of high-throughput  
98 sequencing, further analyses focused on genome-wide screens of single nucleotide *de novo*  
99 mutations (DNMs). Yeager, *et al.* collected and sequenced a cohort of 130 offspring and their  
100 105 parents that were exposed to ionizing radiation after the 1986 Chernobyl nuclear  
101 accident, where the parents were inhabitants of the town of Pripyat or liquidators guarding or  
102 cleaning up the accident site<sup>18</sup>. Doses prior to conception ranged from 0 to 4,080 mSv  
103 (Mean: 365.42 mSv, Std.: 684.55 mSv, Median: 29 mSv), but no elevated mutation rates for  
104 isolated DNMs were detected<sup>18,19</sup>. They also did not detect any enrichment of C>T mutations  
105 within 47 kb intervals, which has been hypothesized to be an indicator for hyper-mutability  
106 of single-strand intermediates during repair of double-strand breaks<sup>20,21</sup>.

107 We reanalyzed the cohort of Yeager, *et al.* comprising 130 offspring, based on their data  
108 accessed under dbGAP accession number phs001163.v1.p1. Now, we recruited and analyzed  
109 an extended cohort of 110 offspring from 80 soldiers who served in German radar units<sup>22</sup>.  
110 This second cohort of exposed parents and their children was collected from former radar  
111 operators of both German armies (Bundeswehr and Nationale Volksarmee, NVA) for this  
112 study, which are known to have been subjected to high doses of ionizing radiation when  
113 servicing radar ground stations during live operations<sup>7</sup>. Even though the German government  
114 spent significant efforts in investigating health effects and occupational risks following long  
115 time service with radar units, and despite the fact that soldiers serving at unprotected radar  
116 installations have a higher risk to develop certain cancers, reliable data on the damage caused  
117 by the stray radiation from these devices is very limited<sup>7,23</sup>. Herein, we examine whether the  
118 signature of radiation-induced clustered DNA lesions is detectable in children born to parents  
119 with a high likelihood of exposure to ionizing radiation.

120 For the radar soldiers, the gonadal doses have been estimated by the radiation measurement  
121 facility (*Strahlenmessstelle*) of the Bundeswehr and yielded a lower mean exposure than in  
122 the Chernobyl (CRU) cohort (9.2 mSv vs 365.42 mSv, Figure 1b). Dosage estimations were  
123 based on sparse measurements taken during the active service of common radar devices or on  
124 emissions of devices reconstructed especially for retrospective assessments. Due to these  
125 complications and because some soldiers served in military roles that probably did not result  
126 in elevated levels of exposure, the dosage estimation remains inconclusive for the majority of  
127 soldiers (n exposed= 22, n no exposure = 55, n no documents=3, Supplementary Figure 1, 2,  
128 Supplementary Bericht S 209/20).

129 As a control cohort we chose the Inova study group consisting of 1,214 family trios and  
130 quadros, for which no exposure beyond naturally occurring ionizing radiation has been  
131 recorded. This cohort was also used to analyze the parental age effect on isolated DNMs in  
132 the germline<sup>24,25</sup>. All sequencing data were reprocessed using the NVIDIA Parabricks and  
133 Illumina DRAGEN toolkits for short-read mapping and variant calling on the GRCh37  
134 reference genome, joint variant calling was performed on a total of 4,337 whole genomes  
135 using GLNexus v1.3.1<sup>26</sup>. All data was subjected to a rigorous quality control pipeline,  
136 including checks for mean coverage, contamination with foreign DNA, the expected pedigree  
137 and more (Supplemental Material, Section 1.4.3), which resulted in the exclusion of 14  
138 samples, mostly due to contamination of the saliva samples we received from these study  
139 participants. The resulting variant data was subsequently filtered for DNMs, with one of the  
140 criteria being that each *de novo* allele only occurs once in the entire call set (Supplemental  
141 Material, Section 1.4.2, Supplementary Table 2). In agreement with the literature, we could  
142 confirm that the number of isolated DNMs increased by 2% per year of paternal age (Figure  
143 2, Supplementary Table 5) which translates to an accumulation of 1-2 mutations per year of

144 age of the father<sup>24,27,28</sup>. To account for the significant age effect and age differences between  
145 the parents in the three cohorts, we computed an age-matched subset of the control cohort  
146 which was used for the subsequent statistical analysis. This age-matching minimizes the age  
147 difference between the control or CRU and the Radar cohort (Supplemental Material, Section  
148 1.6.1). For the age-matched subsets, we found a rate of 72.67 (18.15, median = 79) isolated  
149 DNMs per generation for the Inova cohort, 65.43 (13.57, median = 65) isolated DNMs for the  
150 CRU cohort, and 67.95 (17.25, median=64) isolated DNMs for the Radar cohort  
151 (Supplemental tables 1 and 3). These rates are comparable to the values reported in the  
152 original works and did not differ significantly in between the cohorts (Supplemental Material,  
153 Section 1.6.3)<sup>18,24</sup>.

154

155

### 156 **Analysis of clustered *de novo* mutations**

157

158 We continued our analysis by filtering for loci with multiple lesions (clustered DNMs,  
159 cDNMs). These sites are defined as genomic regions where at least two *de novo* mutations  
160 occur within 20 bp distance of each other. If there are multiple successive pairs of DNMs that  
161 fulfill this criteria, they are always grouped into the same cluster, even if the first and last  
162 lesion in the cluster might exceed the 20 bp distance threshold. In total, we observed 1,989  
163 clustered mutations in 1,515 offspring. The median number of clustered DNMs was two in  
164 offspring from the Chernobyl cohort, which is twice as many as in an age-matched subset of  
165 the unexposed controls. Using a negative binomial regression model we confirmed that the  
166 estimated number of cDNMs differ between either exposed and the control cohort ( $n = 110$ ,  $p$   
167 Radar= 0.005,  $p$ Chernobyl<  $1 \cdot 10^{-3}$ ; Figure 3, Supplemental Table 6-7). These differences are  
168 more prominent when the Radar cohort is split into a subcohort of children with either  
169 exposed fathers (estimation > 0 mSv) or unexposed fathers, where the children of exposed  
170 parents show a higher number (exposed=1.72, not exposed = 1.39) of cDNMs on average  
171 (Supplemental Material, Section 1.6.8).

172 In contrast to isolated DNMs, the false positive rate is higher for clustered DNMs<sup>29</sup>. We,  
173 therefore, established criteria for the assessment of alignments from short read data in the  
174 IGV browser (Supplemental Material, Section 1.5). This was an iterative process involving  
175 validating different call sets of clusters from the Radar cohort, which is the only of the three  
176 cohorts where DNA for validation with secondary sequencing methods was available by  
177 Sanger and PacBio sequencing (Supplementary Figure 6). Our validation efforts resulted in  
178 an estimate for the positive predictive value (PPV), i.e. the probability of a discovered cDNM  
179 being true positive, of 0.23 (95% Clopper-Pearson confidence interval 0.17 - 0.30). It is  
180 noteworthy that in some individuals none of the detected cDNMs could be validated,  
181 including the outlier with 14 cDNMs shown in Figure 3. However, the negative binomial  
182 regression model is not substantially affected by this as it models the median count of  
183 cDNMs per offspring, which is robust against outliers. Additionally, simulations accounting  
184 for this PPV did not affect the significance of the test results (Supplemental Material, Section  
185 1.6.6). All true positive cDNMs identified in the Radar cohort were analyzed with respect to  
186 their likely effect on diseases or the coding region in general, but none was found to have any  
187 implications with any genetic conditions in the study participants (Supplemental Material,  
188 Section 1.5.1).

189 We used read-based phasing of DNMs with informative variants to determine the parental  
190 gamete of origin<sup>30</sup>. The proportion of DNMs that can be phased varied between the three  
191 cohorts because of technical and stochastic reasons. The influencing factors are the distance  
192 between DNM and informative SNP, the coverage in that region, the length of the sequencing  
193 reads (100bp in the control cohort vs 150bp in both case cohorts), and the distribution of

194 fragment sizes. We did not observe changes in the number of isolated DNMs that were  
195 attributed to the paternal or maternal alleles between the three cohorts in this study (Chi-  
196 Squared-Test, Supplemental Material, Section 1.6.7). Since clustered DNMs can extend over  
197 several base pairs, not all lesions in a cluster can necessarily be phased. If read-based phasing  
198 suggested the paternal or maternal germline based on the information from at least one lesion  
199 we assumed this origin for the whole cluster. We did not observe any clusters with  
200 contradicting phasing information regarding the parental origin. In addition to read-based  
201 phasing, we also determined the parental origin with Sanger and PacBio long-read  
202 sequencing for a subset of cDNMs in the Radar cohort during our validation efforts. By  
203 resequencing the phase informative single nucleotide polymorphism (SNP) alongside each  
204 cluster, we validated the parental origin of 26 clusters in the Radar cohort, where 17 clusters  
205 of paternal and 9 of maternal origin were present. Due to the shorter read length of 100bp in  
206 the Inova cohort, a reliable estimate for this ratio in the population could not be computed.  
207 In addition to a general increase in the amount of cDNMs per sample in the cohorts that were  
208 exposed to ionizing radiation, we found a positive association between the estimated dose and  
209 the number of cDNMs. Using a negative binomial regression model, we see a significant ( $p <$   
210  $0.001$ ) increase of clustered DNMs with paternal radiation exposure at  $f(n) = 1.55 \cdot e^{0.0005n}$   
211 mutations per  $n$  mSv, when combining both exposed cohorts (Figure 4, Supplementary Table  
212 12). However, we could not assert statistical significance for this model in the Radar cohort  
213 alone. The average number of lesions per cluster, which we term cluster size, is expected to  
214 increase with LET and the largest clusters that we observed in each of the three cohorts is  
215 eight (Inova), nine (Radar), and eleven (CRU) lesions respectively (Supplemental Figure 10).  
216 Analyzing the distribution of cluster sizes in all three cohorts did not yield a statistically  
217 significant shift (Supplemental Material, Section 1.6.9).

218  
219  
220

## 221 Discussion

222

223 The question of whether there is a transgenerational effect of ionizing radiation on the human  
224 genome has been studied ever since large human cohorts were subjected to increased doses of  
225 ionizing radiation. The early epidemiologic studies on the offspring of atomic bomb survivors  
226 now date back more than 70 years and the disadvantage of such analyses is that they require a  
227 readout on the phenotype-level such as malformations<sup>31</sup>. More recent studies indicate that  
228 larger quantities of environmental radiation lead to a higher incidence of cancer and birth  
229 defects, emphasizing that there are subtle effects of ionizing radiation on the human germline  
230 that have not been captured by earlier studies<sup>32</sup>. If an analysis is based on phenotypic data  
231 alone an increased rate of *de novo* mutations or other subtle changes may remain unobserved,  
232 since most *de novo* mutations are rare and occur in the non-coding part of the DNA.  
233 Therefore, the unique capabilities of whole genome sequencing can yield much deeper  
234 insights into the consequences of prolonged ionizing radiation exposure on the human  
235 genome than possible with earlier technologies. Yeager, *et al.* already studied isolated *de*  
236 *nov* mutations and clusters on the scale of kb, which are associated with the repair of  
237 double-strand breaks, and did not find a significant increase in mutation rates<sup>18,24</sup>. In mice  
238 and somatic cells however, clusters on the single-bp scale were previously implied as a  
239 consequence of the secondary effect of high-energy particles interacting with DNA<sup>12-14</sup>. We  
240 reanalyzed the cohort of Yeager *et al.*, as well as an additional cohort with potential exposure  
241 to ionizing radiation, and found *de novo* mutation clusters to be significantly increased in

242 human DNA. Furthermore, our statistical models indicate that the number cDNMs in  
243 offspring increases with the dosage the paternal germline was exposed to.  
244 The potential of transmission of radiation-induced DNA changes to the next generation is of  
245 particular concern for parents that were potentially subjected to higher doses of ionizing  
246 radiation than deemed safe. To address the concerns articulated by former radar soldiers, the  
247 German Ministry of Defense initiated this study with the goal of investigating the  
248 transgenerational effects of ionizing radiation in offspring of exposed parents. Subsequently,  
249 we recruited retired radar operators, but even though we took great care when assessing the  
250 inclusion criteria for each individual, the available information about exposition towards  
251 ionizing radiation was limited. In practice that meant, radar devices that were in active  
252 service more than 50 years ago had to be made operational again, to measure scattered  
253 radiation profiles <sup>7</sup>. In addition, the service hours and proximity to the radar device during  
254 operation and maintenance were derived from a generalized service manual of the German  
255 armies based on rank, position and mission of the soldier, while actual practice from memory  
256 protocols of the study participants differed, suggesting a potential source of errors. There is  
257 anecdotal evidence of higher ranking soldiers helping with maintenance work, however, this  
258 is not part of the official records. Thus, some of the individuals that we list as unexposed,  
259 might actually have been irradiated (Figure 3 and 4) and dosage values given for members of  
260 the Radar cohort are likely to be underestimated. For instance, we had one offspring in the  
261 Radar cohort with 14 clustered DNMs whose parents were unexposed according to the  
262 estimation based on official records, yet no other causes of hypermutation were known <sup>2</sup>.  
263 Despite the discussed inaccuracies of dosage estimation we proceeded with dose-effect  
264 estimations, and excluded this individual and other offspring of allegedly not exposed fathers  
265 from the negative binomial regression models (Figure 4).  
266 Aside from difficulties in a-priori gauging potential radiation exposure scenarios for each  
267 soldier, there were several potential ascertainment biases in recruiting. First, individuals that  
268 were under the subjective impression that they had been exposed were more likely to  
269 participate (volunteer bias), and second, radar soldiers that were operating devices emitting  
270 the highest quantities of stray radiation, were in their eighties and had a high personal risk for  
271 diseases following their service (survivorship bias) <sup>7</sup>. However, also with these caveats, most  
272 of the statistical results indicate an influence of paternal radiation exposure on the number of  
273 cDNMs per offspring, even when considering the smaller and less exposed radar cohort  
274 alone. These findings are further supported by the increase of effect size and statistical  
275 significance in the more likely exposed subgroup.  
276 In all exposed cohorts, we observed that the number of cDNMs increased by one to two per  
277 genome in children of fathers that were subjected to elevated doses of ionizing radiation for a  
278 longer time. To derive a clinical interpretation from these statistical results, we can compare  
279 the number and impact of cDNMs with the disease burden due to all DNMs. The total  
280 number of clusters exceeds that of the general population by 0.6 for the Radar and 1.77 for  
281 the CRU cohort. In addition to the increase in the total number of DNMs per sample due to  
282 radiation exposure, it is also hypothesized that the functional impact of cDNMs is larger, if  
283 they fall within coding regions of the human genome compared to isolated DNMs. This  
284 increased impact could lead to pathogenicity or even embryonic lethality, in cases where  
285 cDNMs affect important parts of the exome. However, given the low overall rate of excess  
286 cDNMs following paternal exposure to ionizing radiation and the portion of the genome that  
287 is protein coding, the likelihood that a disease in an offspring of exposed parents is triggered  
288 by a DNMs cluster is minimal. Therefore, with the paternal age effect at roughly 1 additional  
289 DNMs per year of age of the father, and the expected number of 60 to 80 DNMs per  
290 generation on average, the paternal exposure to low dose ionizing radiation contributes less to  
291 an individual's risk for genetic diseases than age <sup>20,27,33</sup>. Thus, the excess risk attributable to

292 DNMs that arose after paternal exposure to ionizing radiation is negligible compared to the  
293 base risk for genetic diseases. While the expected clinical consequences of each clustered or  
294 isolated DNM is of comparable order, the consequences of a double strand break that is  
295 incorrectly repaired, is usually more severe. Translocations are most likely an indirect  
296 consequence of double-strand breaks, and have been observed with increased frequency in  
297 irradiated mice as well as in offspring of Radar soldiers<sup>14,22</sup>. However, in contrast to cDNMs,  
298 comparing mutation rates for structural variants is more error-prone when cohorts were  
299 sequenced with different short read lengths and we have therefore refrained from assessing  
300 them statistically.

301 Further studies with longer read lengths, and ideally larger cohorts and more accurate radio-  
302 dosimetry are required to better characterize the disease risk and dose-response relationships  
303 exhibited by transgenerational signatures of prolonged paternal exposure to low dose ionizing  
304 radiation, like cDNMs. Furthermore, to determine the paternal to maternal ratio in cDNMs in  
305 the general population, it is necessary to sequence an appropriate cohort deeply and with  
306 greater read length than what was possible in this study. Since the number of clusters with  
307 paternal and maternal origin is expected to differ due to the accumulation of repair errors, an  
308 accurate measurement of the paternal to maternal cDNM ratio in the general population is  
309 necessary to assess the influence of radiation exposure accurately<sup>34</sup>. A further shift towards  
310 more paternally inherited clusters in exposed cohorts would provide additional evidence for  
311 the correlation between ionizing radiation exposure of the fathers and cDNM rates in the  
312 offspring. Targeting a statistical analysis of structural variants and translocations of the  
313 general population compared to exposed cohorts would also become feasible after subjecting  
314 samples to long read sequencing. It is also interesting to further investigate the potential  
315 impact of the linear energy transfer on the cluster size. Naturally, one would expect the  
316 damage to increase proportionally with the linear energy transfer of the ionizing radiation,  
317 leading to larger clusters, or an increased number of structural variants, when individuals are  
318 subjected to ionizing radiation with higher LET. This effect could also explain differences in  
319 the number and nature of the clusters between the two exposed cohorts, because the gamma  
320 ray spectrum of the ionizing radiation the two cohorts were exposed to differs and a  
321 difference in the LET is one of the consequences thereof<sup>35</sup>.

322 While there are some remaining insecurities about the nature and quantity of the radiation  
323 involved, we could provide strong evidence for the existence of a transgenerational effect of  
324 prolonged paternal exposure to low-dose ionizing radiation on the human genome for the first  
325 time. Luckily for those affected families, the additional risk due to radiation induced cDNMs  
326 on the scale of single base-pairs is very low. In addition, this analysis gives rise to several  
327 further promising research avenues to characterize potential additional transgenerational  
328 signatures of ionizing radiation on the human genome, f.e. analyzing structural changes such  
329 as translocations, which are even more complicated to detect than cDNMs.

330

## 331 **Figures**

332

### 333 **Figure 1: Study Cohorts**

334

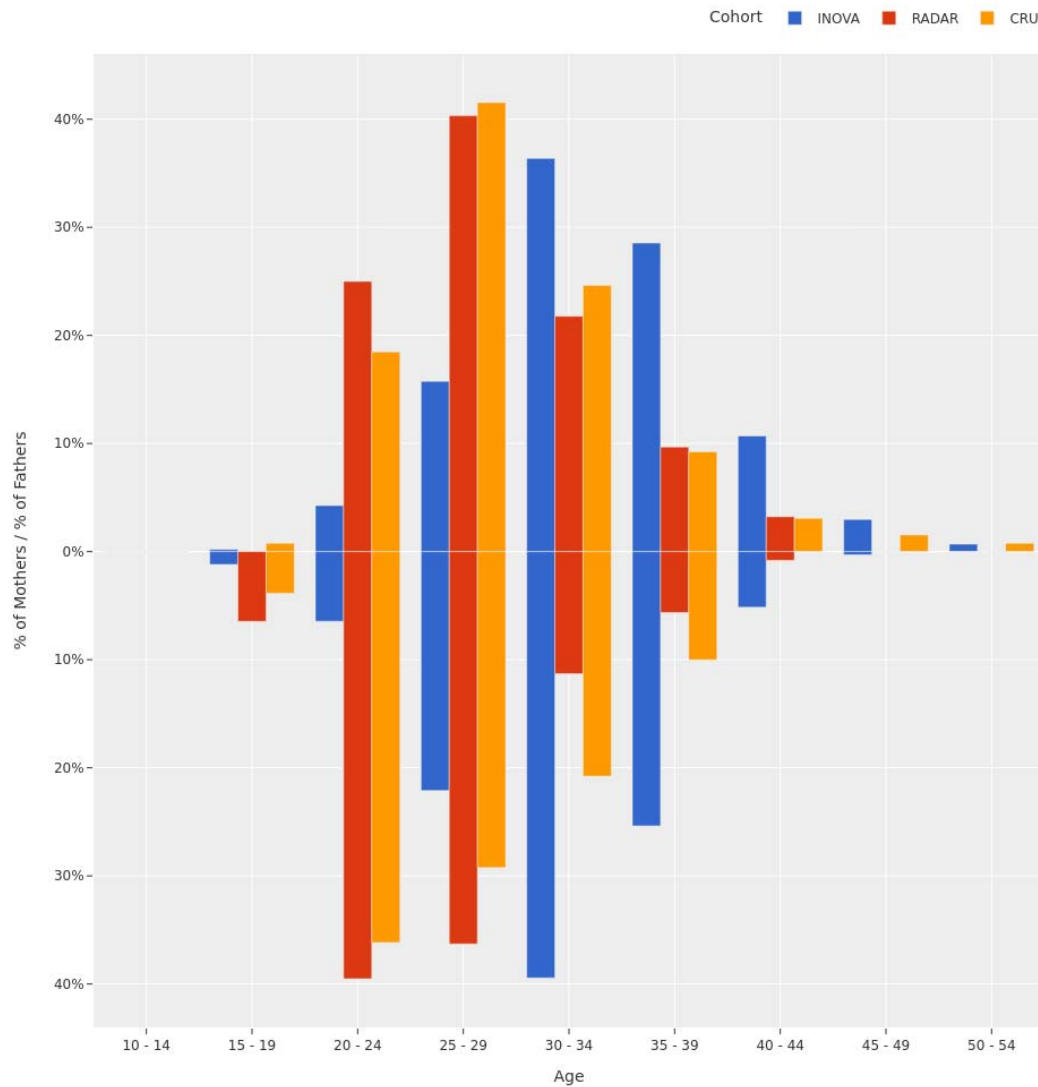
335 **a)** Age distribution of the three study cohorts. Due to the large difference in cohort size,  
336 Y-Axis is given in percentage of the total cohort size. Negative values on the Y-Axis  
337 show the distribution of maternal age, positive values that of paternal age. On  
338 average, fathers are >5 years older in the Inova cohort compared to the case cohorts  
339 and mothers are >5.5 years older on average.

340 **b)** Distribution of paternal exposure for the Radar and CRU cohorts. The maximum  
341 exposure observed in the Radar cohort is 353 mSv, and 4,079 mSv in the CRU cohort.  
342 In the Radar cohort, 107 samples are recorded with an exposure estimation of exactly  
343 0, meaning that their fathers are unlikely to have been exposed or no estimation could  
344 be made.

345 **c)** Distribution of sex of the children in each of the three cohorts.

346

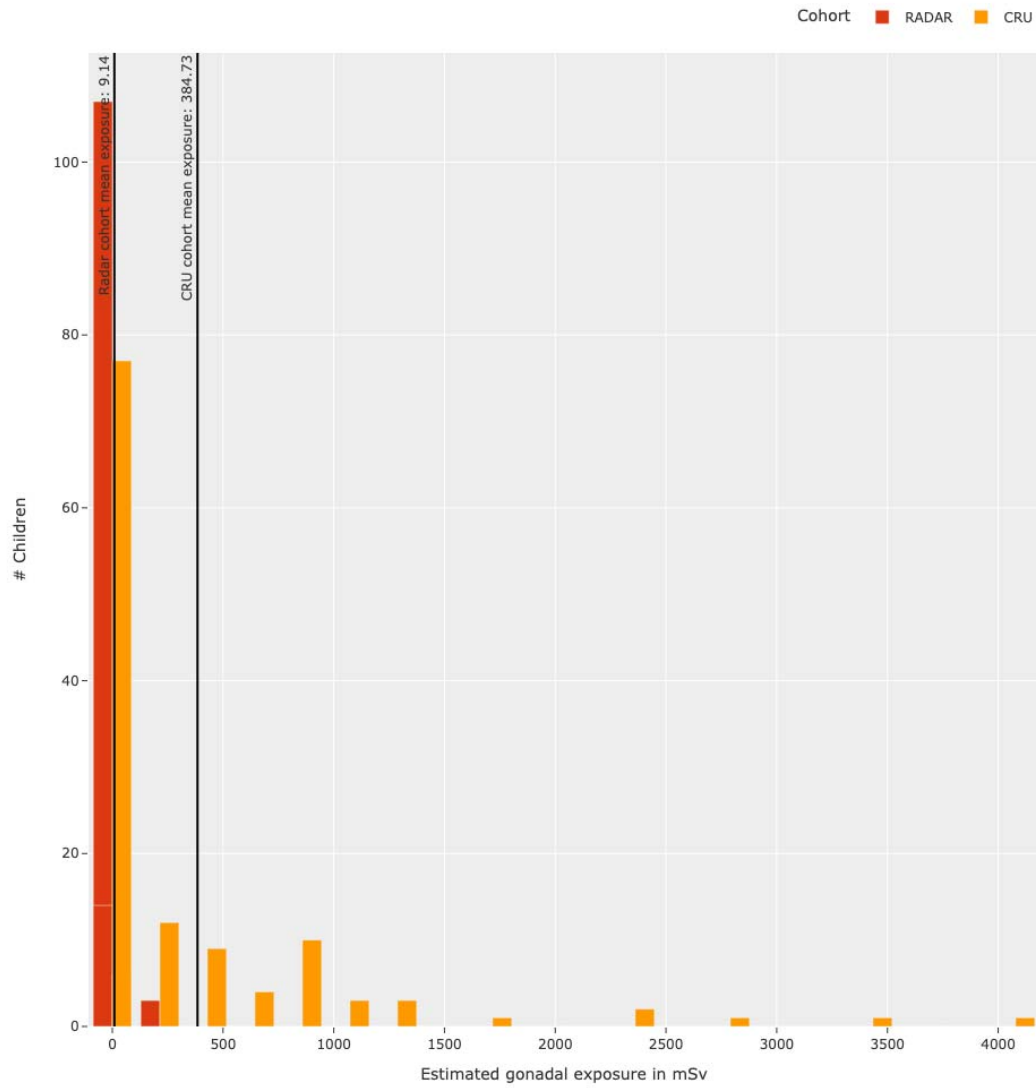
347 **a)**



348

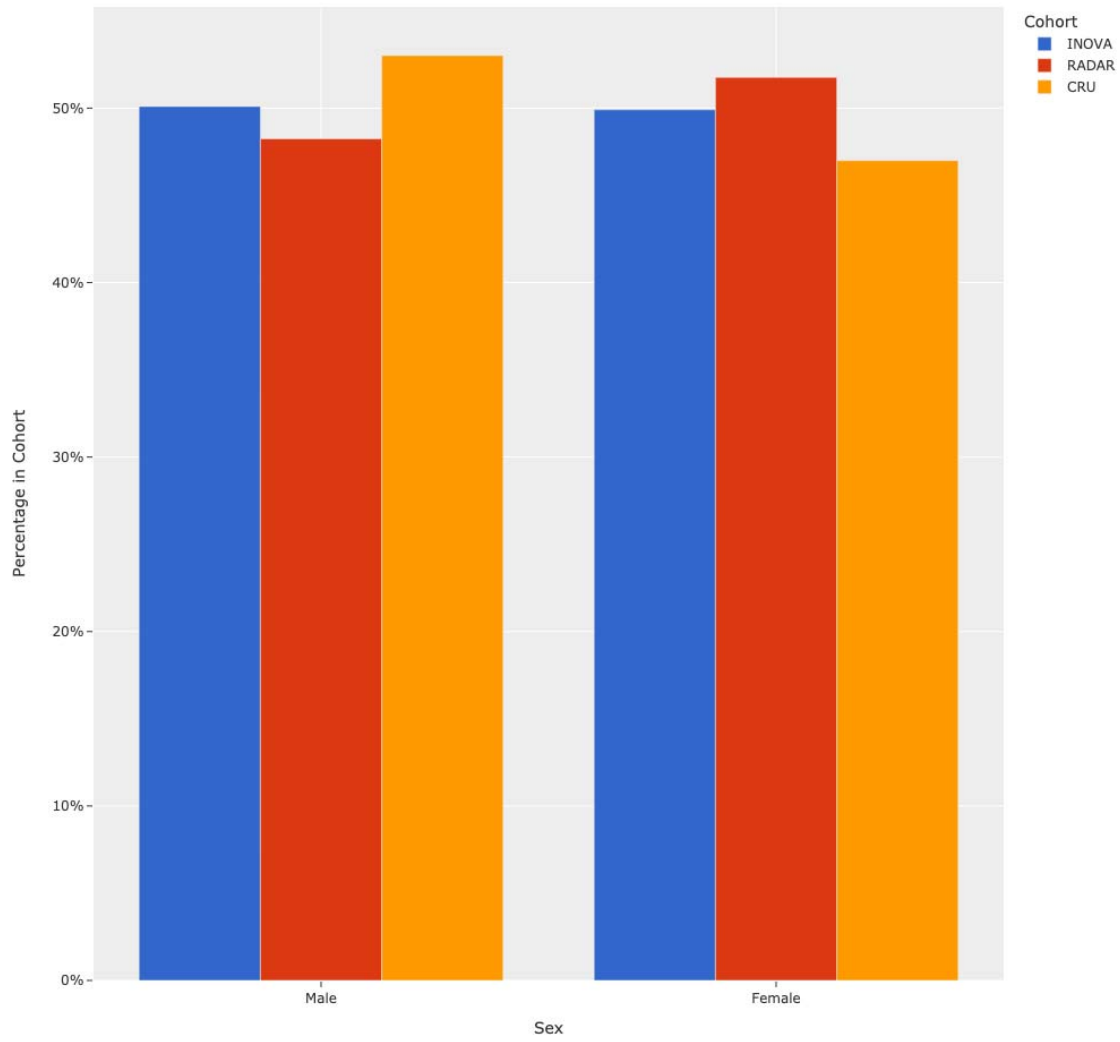


349 b)



350

351 c)



352

353

354 **Figure 2: Paternal Age Effect**

355 Paternal age effects computed by a negative binomial regression (NBR) model estimating the  
356 number of DNMs given the paternal age at conception of the child for each cohort. When  
357 fitting this model, no age matching was applied to the data, so on average the parents are  
358 older in the Inova cohort. Nevertheless, the paternal age effect is roughly 2% for each of the  
359 cohorts, which results in an increase of ~1 DNM per year of age of the father.

360

361



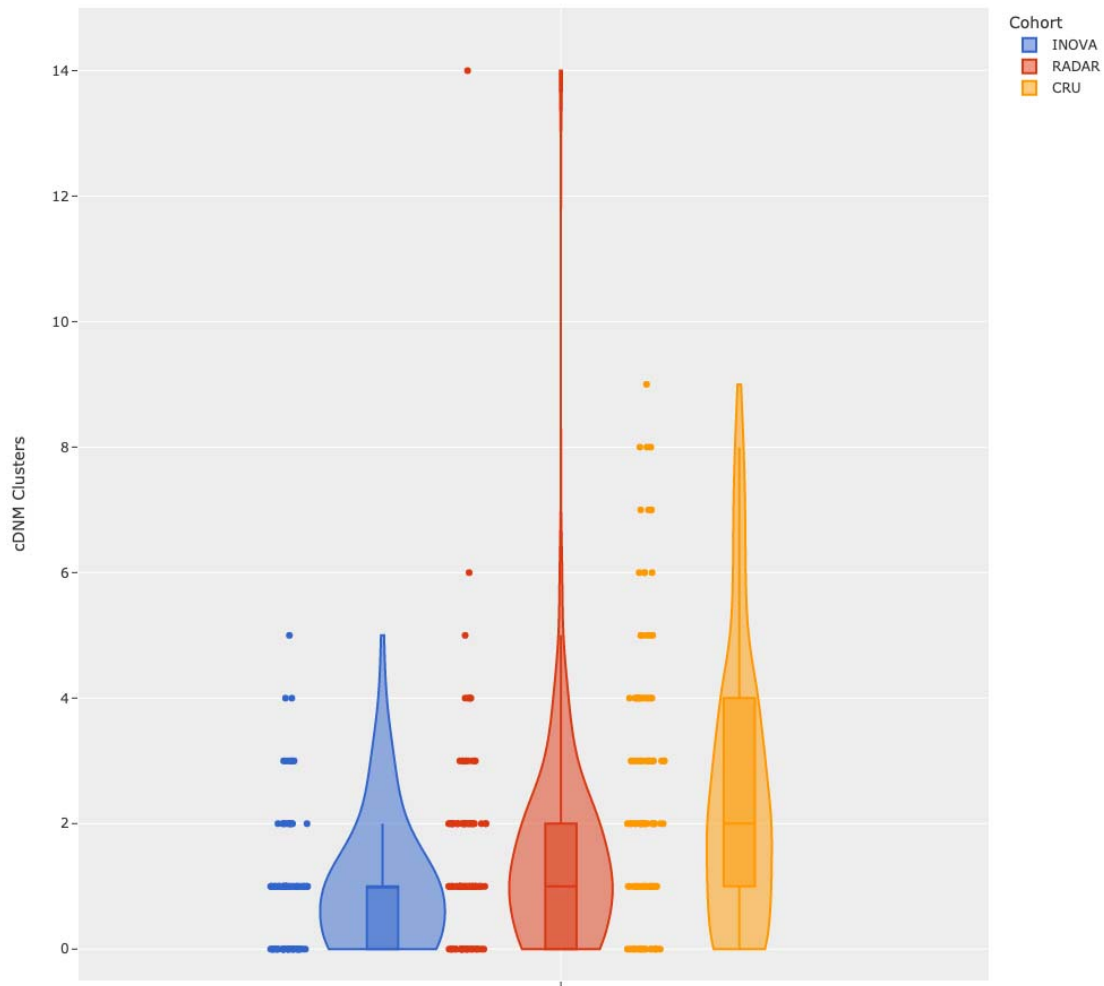
362

363

364 **Figure 3: Number of cDNMs per sample**

365 Violin plot of the number of cDNMs per sample grouped by cohort. The width of the violin at  
366 each integer value of the y-axis visualizes the number of samples and their respective count  
367 of cDNM clusters. The box plot for each cohort is included inside the respective violin to  
368 display the quartile ranges and median number of cDNMs per sample in the given cohort. On  
369 average, we detected [redacted] cDNMs in the Inova cohort,  
370 [redacted] in the Radar cohort and [redacted] in the  
371 CRU cohort on age matched data.

372  
373



374  
375

376 **Figure 4: Number of cDNMs per mSv of paternal exposure**

377 Estimation of the number of cDNMs depending on the paternal exposure in mSv as computed  
378 by a negative binomial regression (NBR) model with logarithmic link function modeling the  
379 number of clusters given the paternal exposure in mSv. Because the accuracy of the negative  
380 binomial regression model deteriorates rapidly with larger exposure estimates, the x-Axis has  
381 been cut off at 1.5k mSv, which means that five samples from the CRU cohort for which the  
382 estimation by the model is very inaccurate, are hidden. The fit shown in this image is  
383 conditional only on the number of mSv of paternal exposure, both exposed cohorts have been  
384 rolled into one supercohort for this analysis. Supplemental Figure 10 shows the analogous  
385 model also taking into account cohort membership for each sample. We only included the  
386 samples from the exposed subgroup of the radar cohort in this analysis, i.e. all offspring of  
387 fathers with an estimated exposure of >0 mSv. This restriction in the cohort leads to the  
388 exclusion of some children with many cDNMs, e.g. the outlier in the Radar cohort in Figure  
389 3.  
390



## 392 **Data Availability**

- 393     • Sequencing data can be accessed at EGA:
  - 394         ○ EGA: Study Accession EGAS00001007321
- 395     • Code can be accessed at GitHub and Zenodo:
  - 396         ○ GitHub: <https://github.com/brand-fabian/radarstudy>
  - 397         ○ Zenodo <https://doi.org/10.5281/zenodo.8431077>

## 398 **Acknowledgement**

399 The radar study was funded by the German Ministry of Defense and conducted by Deutsches  
400 Zentrum für Luft und Raumfahrt (DLR) under the supervision of an expert committee  
401 consisting of Michael Krawczak, Hajo Zeeb, Ivo Gut, André Reis and experts from the  
402 Bundeswehr (Reinhard Ullmann, Matthias Port). Dosages were estimated by Andreas  
403 Schirmer from Strahlenmessstelle der Bundeswehr. Recruitment was assisted by Dietmar  
404 Glaner, Josef Wiesner, and “Bund zur Unterstützung Radargeschädigter (BzUR e.V.)”. *De*  
405 *novo* mutation rates in unexposed controls were also discussed with Hákon Jónsson from  
406 DeCode. Data was generated by the NGS core facility at Universitätsklinikum Bonn, which is  
407 part of the West German Genome Center (WGGC) and processed by the Core Unit for  
408 Bioinformatics in Bonn. Epigenetic analysis was performed by Sascha Tierling and Jörn  
409 Walter of University Saarland.

410

## 411 **Ethics Declaration**

412

413 All participants of the radarstudy were of legal age at time of recruitment and consented to  
414 the use of their data for research purposes. Ethics approval was granted by the  
415 “Ethikkommission der Medizinischen Fakultät Bonn” (Ethics committee of the medical  
416 faculty Bonn). All identifiers were pseudonymized prior to publication and the  
417 reidentification based on identifiers mentioned in the manuscript or supplemental material is  
418 only possible by members of the research group.

419

## 420 **Author contributions:**

421     Conceptualization: FB, HK, AK, MH, LW, DB, MN, MS, KS, PMK

422     Data curation: FB, AK

423     Formal analysis: FB, HK, LW

424     Funding acquisition: PMK, DB

425     Investigation: FB, HK, AK

426 Methodology: FB, HK, AK, LW, MS, PMK  
427 Project administration: PMK  
428 Resources: FB, AK, MH, KUL, PK, GM, DB, PMK  
429 Supervision: DB, MN, MS, KS, PMK  
430 Software: FB  
431 Visualization: FB  
432 Writing – original draft: FB, HK, AK, PMK  
433 Writing – review & editing: FB, HK, AK, PMK  
434

435

## 436 **Bibliography**

437

- 438 1. Alexandrov, L. B. *et al.* Signatures of mutational processes in human cancer. *Nature* **500**,  
439 415–421 (2013).
- 440 2. Kaplanis, J. *et al.* Genetic and chemotherapeutic influences on germline hypermutation.  
441 *Nature* **605**, 503–508 (2022).
- 442 3. Garrett-Bakelman, F. E. *et al.* The NASA Twins Study: A multidimensional analysis of a  
443 year-long human spaceflight. *Science* **364**, (2019).
- 444 4. Cucinotta, F. A. Space Radiation Risks for Astronauts on Multiple International Space  
445 Station Missions. *PLoS ONE* **9**, e96099 (2014).
- 446 5. George, K., Rhone, J., Beitman, A. & Cucinotta, F. A. Cytogenetic damage in the blood  
447 lymphocytes of astronauts: Effects of repeat long-duration space missions. *Mutat. Res.*  
448 *Toxicol. Environ. Mutagen.* **756**, 165–169 (2013).
- 449 6. Radiation, U. N. S. C. on the E. of A. & others. *Sources and effects of ionizing radiation,*  
450 *united nations scientific committee on the effects of atomic radiation (UNSCEAR) 2008*  
451 *report, volume I: Report to the general assembly, with scientific annexes A and B-*  
452 *sources.* (United Nations, 2010).
- 453 7. König, W. *et al.* Bericht der Expertenkommission zur Frage der Gefährdung durch  
454 Strahlung in früher Radareinrichtungen der Bundeswehr und der NVA

- 455 (Radarkommission)[Report of the expert group on the issue of the hazards of radiation in  
456 earlier types of radar devices and the German Radar Commission]. Berlin, German Radar  
457 Commission. (2003).
- 458 8. Frankenberg-Schwager, M. Induction, repair and biological relevance of radiation-  
459 induced DNA lesions in eukaryotic cells. *Radiat. Environ. Biophys.* **29**, 273–292 (1990).
- 460 9. Georgakilas, A. G., O’Neill, P. & Stewart, R. D. Induction and repair of clustered DNA  
461 lesions: what do we know so far? *Radiat. Res.* **180**, 100–109 (2013).
- 462 10. Sage, E. & Shikazono, N. Radiation-induced clustered DNA lesions: Repair and  
463 mutagenesis. *Free Radic. Biol. Med.* **107**, 125–135 (2017).
- 464 11. Sage, E. & Harrison, L. Clustered DNA lesion repair in eukaryotes: relevance to  
465 mutagenesis and cell survival. *Mutat. Res. Mol. Mech. Mutagen.* **711**, 123–133 (2011).
- 466 12. Adewoye, A. B., Lindsay, S. J., Dubrova, Y. E. & Hurles, M. E. The genome-wide  
467 effects of ionizing radiation on mutation induction in the mammalian germline. *Nat.*  
468 *Commun.* **6**, 6684 (2015).
- 469 13. Satoh, Y. *et al.* Characteristics of induced mutations in offspring derived from irradiated  
470 mouse spermatogonia and mature oocytes. *Sci. Rep.* **10**, 1–13 (2020).
- 471 14. Matsuda, Y. *et al.* Spectra and characteristics of somatic mutations induced by ionizing  
472 radiation in hematopoietic stem cells. *Proc. Natl. Acad. Sci.* **120**, e2216550120 (2023).
- 473 15. Dubrova, Y. E. *et al.* Human minisatellite mutation rate after the Chernobyl accident.  
474 *Nature* **380**, 683–686 (1996).
- 475 16. Slebos, R. J. *et al.* Mini- and microsatellite mutations in children from Chernobyl accident  
476 cleanup workers. *Mutat. Res. Toxicol. Environ. Mutagen.* **559**, 143–151 (2004).
- 477 17. Little, M. P., Goodhead, D. T., Bridges, B. A. & Bouffler, S. D. Evidence relevant to  
478 untargeted and transgenerational effects in the offspring of irradiated parents. *Mutat. Res.*  
479 *Mutat. Res.* **753**, 50–67 (2013).



- 480 18. Yeager, M. *et al.* Lack of transgenerational effects of ionizing radiation exposure from  
481 the Chernobyl accident. *Science* **372**, 725–729 (2021).
- 482 19. Chumak, V. *et al.* Estimation of radiation gonadal doses for the American–Ukrainian trio  
483 study of parental irradiation in Chornobyl cleanup workers and evacuees and germline  
484 mutations in their offspring. *J. Radiol. Prot.* **41**, 764–791 (2021).
- 485 20. Jónsson, H. *et al.* Parental influence on human germline de novo mutations in 1,548 trios  
486 from Iceland. *Nature* **549**, 519–522 (2017).
- 487 21. Roberts, S. A. & Gordenin, D. A. Hypermutation in human cancer genomes: footprints  
488 and mechanisms. *Nat. Rev. Cancer* **14**, 786–800 (2014).
- 489 22. Holtgrewe, M. *et al.* Multisite de novo mutations in human offspring after paternal  
490 exposure to ionizing radiation. *Sci. Rep.* **8**, 1–5 (2018).
- 491 23. Degraeve, E., Meeusen, B., Grivegnée, A.-R., Boniol, M. & Autier, P. Causes of death  
492 among Belgian professional military radar operators: A 37-year retrospective cohort  
493 study. *Int. J. Cancer* **124**, 945–951 (2009).
- 494 24. Wong, W. S. *et al.* New observations on maternal age effect on germline de novo  
495 mutations. *Nat. Commun.* **7**, 10486 (2016).
- 496 25. Knijnenburg, T. A. *et al.* Genomic and molecular characterization of preterm birth. *Proc.*  
497 *Natl. Acad. Sci.* **116**, 5819–5827 (2019).
- 498 26. Lin, M. F. *et al.* GLnexus: joint variant calling for large cohort sequencing. *BioRxiv*  
499 343970 (2018).
- 500 27. Kong, A. *et al.* Rate of de novo mutations and the importance of father’s age to disease  
501 risk. *Nature* **488**, 471–475 (2012).
- 502 28. Sasani, T. A. *et al.* Large, three-generation CEPH families reveal post-zygotic mosaicism  
503 and variability in germline mutation accumulation. *bioRxiv* 552117 (2019).

- 504 29. Besenbacher, S. *et al.* Multi-nucleotide de novo Mutations in Humans. *PLOS Genet.* **12**,  
505 e1006315 (2016).
- 506 30. Belyeu, J. R., Sasani, T. A., Pedersen, B. S. & Quinlan, A. R. Unfazed: parent-of-origin  
507 detection for large and small de novo variants. *Bioinformatics* **37**, 4860–4861 (2021).
- 508 31. Yamada, M. *et al.* Congenital malformations and perinatal deaths among the children of  
509 atomic bomb survivors: a reappraisal. *Am. J. Epidemiol.* **190**, 2323–2333 (2021).
- 510 32. Zlobina, A. *et al.* Impact of Environmental Radiation on the Incidence of Cancer and  
511 Birth Defects in Regions with High Natural Radioactivity. *Int. J. Environ. Res. Public.*  
512 *Health* **19**, 8643 (2022).
- 513 33. Wood, K. A. & Goriely, A. The impact of paternal age on new mutations and disease in  
514 the next generation. *Fertil. Steril.* **118**, 1001–1012 (2022).
- 515 34. Wang, S., Meyer, D. H. & Schumacher, B. Inheritance of paternal DNA damage by  
516 histone-mediated repair restriction. *Nature* **613**, 365–374 (2023).
- 517 35. Semenenko, V. & Stewart, R. A fast Monte Carlo algorithm to simulate the spectrum of  
518 DNA damages formed by ionizing radiation. *Radiat. Res.* **161**, 451–457 (2004).

519  
520  
521