

Implications of Recent Epidemiological Studies for Compensation of Veterans Exposed to Plutonium

Jan Beyea¹

Abstract—The objective of this paper is to compare post-2007 epidemiological results for plutonium workers to risk predicted by the software program NIOSH-IREP (IREP for short), which is used to determine the lowest dose for a US veteran to obtain cancer compensation. IREP output and methodology were used to predict excess relative risk per Gy (ERR Gy⁻¹) for lung cancer at the 99th credibility percentile, which is used for compensation decisions. Also estimated were relative biological effectiveness factors (RBE) predicted for workers using IREP methodology. IREP predictions were compared to results for Mayak and Sellafield plutonium workers, separately and pooled. Indications that IREP might underpredict 99th-percentile lung cancer plutonium risk came from (1) comparison of worker RBEs and (2) from comparison of Sellafield results separately. When Sellafield and Mayak data were pooled, ERR Gy⁻¹ comparisons at the 99th percentile roughly matched epidemiological data with regression dose range restricted to < 0.05 Gy, the most relevant region to veterans, but overpredicted for the full dose range. When four plausible distributions for lung cancer risk, including both new and old data, were combined using illustrative weighting factors, compensation cutoff dose for lung cancer matched current IREP values unless regression results below 0.05 were chosen for Sellafield, producing a two-fold reduction. A 1997 claim of a dose threshold in lung cancer dose response was not confirmed in later literature. The benefit of the doubt is given to claimants when the science is unclear. The challenge for NIOSH-IREP custodians is dealing with the Sellafield results, which might best match US claimants. *Health Phys.* 123(2):133–153; 2022

Key words: epidemiology; plutonium; relative biological effectiveness; risk estimates

INTRODUCTION

EPIDEMIOLOGICAL DATA play a key role in cancer compensation decisions for US veterans exposed to ionizing radiation as part of their military service (Kocher et al. 2008). The data serve as inputs to the software program NIOSH IREP, which is used to assess risk to a veteran or worker in the nuclear weapons industry given an individual's dose history (Otchin 2007; Kocher et al. 2008). NIOSH-IREP, herein called IREP for short, stands for "National Institute of Occupational Safety and Health Interactive RadioEpidemiological Program." Excess relative risks (ERR) predicted by IREP do not appear in the program's public output but are translated into another risk metric that is directly used for compensation, specifically an "assigned share" of an individual's cancer causation probability. Assigned share is calculated as *ERR* divided by $I + ERR$. It is an imputed probability of causation based on a simple biologic model, not a number identifiable from epidemiological data alone (Robins and Greenland 1989; Beyea and Greenland 1999; Greenland 1999). Assigned share is a useful policy construct, but the science lies in estimating ERRs and is the focus here. Unlike the situation in civil legal proceedings, compensation criteria in IREP are based on the upper 99th-credibility percentile and give additional benefits of the doubt to veterans and worker claimants, which makes questioning the uncertainties in assigned share a minor issue.

The program relies on epidemiological and other data analyzed prior to 2007. At the time of IREP's development, there were limited direct epidemiological data for plutonium exposures, so risks were determined primarily from data for the Japanese atomic bomb survivors (the low-LET reference source) using an uncertain quality factor, labeled as a relative effectiveness factor (REF), of which there are two in IREP for (high-LET) alpha particles (see Appendix). Multiplication by the appropriate REF converts an absorbed alpha particle dose into a biologically effective dose to be entered into

¹Senior Scientist, Emeritus, Consulting in the Public Interest, 53 Clinton Street, Lambertville, NJ 08530, jbeyea@cipi.com.

The authors declare no conflicts of interest. Over the last 20 y, the author has prepared pro bono dose and risk analyses for four US veterans with a claim before the US Veterans Administration.

For correspondence contact the author at the above address, or email at jbeyea@cipi.com

(Manuscript accepted 4 March 2022)

Supplemental digital content is available in the HTML and PDF versions of this article on the journal's website www.health-physics.com. 0017-9078/22/0

Copyright © 2022 The Author(s). Published by Wolters Kluwer Health, Inc. on behalf of the Health Physics Society. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

DOI: 10.1097/HP.0000000000001580

regression equations for risk determined from a low-LET source of radiation with photon energies above 250 KeV.

IREP documentation gives distributions for uncertain adjustment factors that can be used to convert an REF into a relative biological effectiveness factor (RBE) with confidence limits, as is done in this paper. The RBE, in turn, can be compared with results from plutonium worker studies. In these studies, the external radiation that served as the reference radiation was protracted and composed of both high and low energy photons, which was not the case in the radiation exposure experienced by atomic bomb survivors. In what follows, a “worker RBE” is one that was determined using risk coefficients for the external reference radiation experienced by workers. Use of an REF was not necessary for comparing plutonium ERR Gy⁻¹ because those could be obtained directly from IREP output by inverting the equation for assigned share.

When risk is assessed in IREP, dose-response linearity for exposure to alpha radiation is generally assumed for solid cancers, with the exception of exposure to radon (Kocher et al. 2008). Thus, there is no dose-and-dose-rate effectiveness factor (DDREF) used for alpha particle exposure above and beyond a small 20% average effect included to account for chronic exposures being slightly more deleterious than acute exposures (Kocher et al. 2005). However, dose linearity is not assumed for low-LET radiation, and uncertain distributions for a DDREF are used in IREP to account for non-linear possibilities for such low-LET exposures, as well as differences in dose rate.

Today, a number of plutonium worker studies are available and have been reviewed in ICRP Publication 150 (Tirmarche et al. 2021). The major post-2007 studies (for Mayak and Sellafield) are listed in Table 1. These studies separate the contributions of internal plutonium exposures from external radiation exposures using multivariate regression analysis. The external radiation was protracted, so any RBEs estimated in these studies cannot be used to infer plutonium risk from atomic bomb external reference radiation without adjustment (Tirmarche et al. 2021). An RBE is a quantity that is obtained experimentally by comparing the doses required to achieve a specific level of biological result (UNSCEAR 2012). Worker exposures fall into the class of unplanned experiments.

Two of the worker studies (Gilbert et al. 2013; Gillies et al. 2017a) presented worker RBEs for mortality complete with uncertainty ranges. These RBE values can be compared to corresponding IREP-based predictions, which are based primarily on animal and cell data and must be adjusted to account for the low-energy photon component that was present in the external radiation experienced by workers. Even with adjustment, the comparison would be somewhat inexact because the predictions of RBEs for workers derived from IREP methodology, like all IREP output, are for incidence and not mortality.

In addition, values of excess relative risk per Gy of plutonium exposure (ERR Gy⁻¹) for lung cancer diagnosed at age 60 are available from these two studies, along with a third (Labutina et al. 2013). The values and their uncertainty ranges can be compared to illustrative ERR Gy⁻¹ values derived from transformations of IREP program output. The IREP-based illustrative predictions sometimes require pooling of cohort data (Appendix). For instance, pooling was necessary for the IREP results by IREP-defined smoking status to match, to the extent possible, the smoking-status mix of epidemiological cohorts, which vary from study to study and do not generally match the mixes defined in IREP.

The worker plutonium data for lung cancer has some limited potential to check for nonlinearity in dose response by examining results that ratchet down the upper boundary for regressions and less satisfactorily by graphical analysis at low doses. Additional plutonium worker studies not listed in Table 1 are discussed separately. These are (1) a study of a European cohort (Grellier et al. 2017), (2) a cohort at the Los Alamos National Laboratory (LANL) (Boice et al. 2022), and (3) a cohort at Rocky Flats (Brown et al. 2004).

The subject of this paper is the implications of this new post-IREP plutonium risk data for claimant compensation decisions. For veterans, the cohorts whose claims might potentially be affected would include those service members who helped clean up plutonium contamination at Palomares in Spain in 1996 (Beyea and von Hippel 2019) and near Thule Air Force base in Greenland in 1968 (USAF 1970), as well as claims from service members who worked in, or helped clean up, plutonium contaminated areas on Johnston Atoll after 30 June 1963 (Rademacher 2016). Claims for workers in the nuclear weapons industry might potentially be affected also, specifically claims filed by those who worked in US plutonium-separation and nuclear weapon R&D and production sites at Rocky Flats, Hanford, Lawrence Livermore, Savannah River, and Los Alamos.

NIOSH-IREP and claimant benefits

IREP implements recommendations of a 2003 report of the NCI-CDC Working Group to Revise the 1985 NIH RadioEpidemiological Tables (US DHHS 2003) established at the request of the Department of Veterans Affairs (DVA). A claimant’s radiation “assigned share” of cancer causation probability determines compensation. As previously stated, assigned share is defined as excess relative risk divided by 1 + excess relative risk. It must reach 50% in IREP to justify an award, but the 50% is based on the 99th “credibility” percentile (US DHHS 2003; Kocher et al. 2008). Thus, the estimated likelihood of the assigned share actually reaching 50% can be as low as 1 in 100. This “benefit of the doubt” approach is intended to make it very unlikely that a rejected veteran was treated unfairly by the IREP methodology. Compensation is an all or nothing decision. The use of the

Table 1. Dose-response findings on lung, liver and bone, as reported in Mayak studies of workers exposed to plutonium.^a

| Study | Cancer site | Dose-response findings and comments | Dosimetry version | Study type ^b | Smoking adjustment | Quadratic term | Threshold analysis |
|-------------------------|-------------|--|-------------------|-------------------------|--------------------|----------------|--------------------|
| Tokarskaya et al. 1997 | Lung | Threshold at 0.8 Gy, quadratic or linear + quadratic | Pre-1997 | Case control | Yes | P = 0.0001 | Yes |
| Tokarskaya et al. 2002 | Lung | Stated that not adjusting for smoking could overestimate risks, possibly obscure threshold | 2000 | Case control | Yes | No | No |
| Tokarskaya et al. 2006) | Liver | Increased odds ratio for 2–16 Gy relative to 0–2 Gy category | 2000 | Case control | Yes | No | No |
| Labutina et al. 2013 | Lung | Linear fit to full dose range | 2008 | Cohort | Yes | P = 0.25 | No |
| | Liver | Linear-quadratic up to 6 Gy LQ effect “diminished” when restricted internal dose <2 Gy” | 2008 | Cohort | Yes | P = 0.001 | No |
| | Bone | No definite dose response Only 4 cases with Pu exposure | 2008 | Cohort | Yes | No | No |
| Gilbert et al. 2013 | Lung | Linear fit for full dose range and if restricted to <0.2 Gy Not mentioned by AF report. Lung cancer mortality | 2008 | Cohort | Yes | P = 0.5 | No |
| Gillies et al. 2017a | Lung | Linear fit for full dose range Dismissed in an Air Force (AF) report because of no smoking adjustment (Rademacher 2020) | 2013 | Pooled cohort | No | P > 0.5 | No |
| Stram et al. 2021 | Lung | Linear fit for full dose range, adjusted for dose error Adjusted for SES status. Published after AF report | 2013, 2016 | Cohort | Yes | No | No |

^aOther earlier Mayak papers were not listed here, because they have been superseded by the publications listed.

^bCohort = cohort study following workers over time.

word “credibility” rather than “probability” in the phrase, “99th-credibility percentage,” is a convention from Bayesian statistics indicating that uncertainties have been estimated based in part on prior knowledge and expert judgement.

All cancers in IREP are deemed to be radiogenic. Excess risk for 33 cancer types can be calculated as a function of estimated organ dose and converted into assigned share using the IREP program’s online calculator. Chronic lymphocytic lymphoma was added in a 2013 update (<https://irep.oraucoc.org/update.asp>). The focus in this paper is on solid cancer, primarily lung cancer. For blood cancers, like lymphoma, the target organs for dose calculations are not always obvious and must be determined before IREP can be used (NIOSH 2005). There is post-2007 epidemiological data of possible relevance to compensation for lymphoma,

which is noted here (Richardson et al. 2009; Hunter and Haylock 2022) but not discussed.

Specific inputs to IREP include a claimant’s age at exposure, age at diagnosis, smoking status, and dose history, which is entered year by year. An individual’s radiation dose history comes from the government agency that has historical dosimeter measurements and other records for a claimant or has experts/contractors who provide doses estimated when dosimeter readings are incomplete (CFR 2021b and c). Alternatively, veteran claimants may be able to provide their own dose estimates for consideration, provided their designated expert is deemed qualified by the agency. If a range of doses is provided for a veteran, the highest value in the range is entered into IREP, without an uncertainty range (Otchin 2007). For a nuclear weapons

worker, uncertainty in dose can be entered as a distribution into IREP. Doses inputted into IREP for alpha particles are entered as equivalent dose in Sv, i.e., the dose in Gy weighted by a factor of 20. Questions about assigning dose estimates in the face of limited dosimetry are also important, e.g., in the case of veterans of plutonium cleanups (Beyea and von Hippel 2019), but are not dealt with here.

Regulations for veterans (CFR 2021b) require that, when there is an uncertain range of doses for a claimant, the highest level is to be chosen. However, IREP itself is intended to provide unbiased estimates of excess relative risks (ERR) and assigned shares of causation of cancers, given the input dose history, whenever possible. "All the effort in modeling that is incorporated in IREP is directed at estimating ERR for a specific cancer in an exposed individual and its uncertainty" (Kocher et al. 2008). When the science is reasonably clear, the only benefit given to the claimant in the programming is the use of the 99th-percentile credibility level. Alternatively, when a parameter value needed for an IREP calculation is so uncertain that no clear scientific basis exists to make a choice for its distribution, assumptions favorable to the claimant are made. For instance, should it not be possible to choose one risk model incorporated into IREP over another, and should there be conditions of exposure that are considered plausible, the model or assumption most favorable to the claimant is chosen (Kocher et al. 2008; USC 2011).

These are all benefits given for service to the country, an example of the "veteran's canon," which goes back to US Supreme Court decisions as early as 1943: "The Soldiers' and Sailors' Civil Relief Act is always to be liberally construed to protect those who have been obliged to drop their own affairs to take up the burdens of the nation" (Harper 2019). Another motivation for giving benefit of the doubt to claimants may have been to assure future soldiers that, should they be exposed to radiation, they could expect favorable treatment in risk assumptions and dosimetry analysis.

The veteran's canon does not apply to civilian workers in the nuclear weapons industry. Instead, they are covered under the Energy Employees Occupational Illness Compensation Program Act (EEOICPA), which adopted the claimant favorable approach used in IREP's predecessor, the 1985 Radioepidemiological Tables (US DHHS 2003). The use of NIOSH-IREP to replace the Tables for use in compensation under EEOICPA was formalized in regulations (CFR 2022). There is a history of favorable treatment for worker compensation claims in general, but the degree varies by state (Spieler 2016).

Is there a plutonium dose threshold in lung cancer dose response?

A recent Air Force report (DVA 2020; Rademacher 2020), written in defense of the Air Force's treatment of vet-

erans of the Palomares, Spain, plutonium cleanup in 1966, claims evidence of a dose threshold for lung cancer in the dose response for Mayak plutonium workers below 0.2 Gy.² If correct, such a finding could have a major impact on IREP, possibly challenging the long-standing assumption of dose linearity for alpha radiation. However, no uncertainty range was assigned, either in risk or threshold dose position. The claim is therefore operationally unusable because a veteran's compensability is determined by the upper tail of the risk distribution at the veteran's assigned dose. If there should be a true threshold in cancer dose response at dose, d , the key uncertainty of interest for compensation is not the uncertainty around the position of the dose threshold but the uncertainty around the dose response risk at d and below. Neither uncertainty is covered by the Air Force report nor in the one paper from 1997 (Tokarskaya et al. 1997) that actually claimed a threshold (using out-of-date dosimetry and a small number of datapoints for regression). To get usable uncertainty information, there would likely need to be a fit to the original data with a dose response threshold function, followed by extraction of prediction errors calculated from the fit as a function of dose. This could only be done by the original authors unless the data were made publicly available.

Without an uncertainty range, an uncertain claim of a dose threshold in lung cancer dose response cannot be used to dismiss as irrelevant criticism of dosimetry made by the US Court of Appeals for Veterans Claims in its order for a reassessment of Palomares veterans' dosimetry (CAVC 2020). Nevertheless, it is possible that the dose threshold claim might play a role in the reassessment ordered by the Court, should the order survive the Veteran Administration's appeal to the Federal Circuit court (PACER 2021). Therefore, a search was made for any claims of dose threshold in the published post-IREP epidemiological plutonium studies.

Particular attention has been given here to the lowest dose region for which plutonium epidemiological data are available, which is well below the 0.2-Gy region that is suspect to the Air Force. The focus here is on the dose region below 0.05 Gy, which is a range for which ERR Gy⁻¹ has been determined from regression of Mayak and Sellafield data. If there is no threshold in dose below 0.05 Gy, there will not be one above. Even lower than 0.05 Gy and of most direct relevance to claimants are plutonium lung doses below 0.02 Gy (0.4 Sv equivalent dose). This is the level judged compensable at the 99th-credibility percentile by the Air Force for lung cancer in veterans exposed during the Palomares cleanup, as listed in Table 1 of Beyea and von Hippel (2019). It is the Sellafield data, which the Air Force report did not discuss, that dominates the number of excess cancers at plutonium doses below 0.02 Gy.

²Note that, unless specifically labeled as weighted values, all doses in Gy are unweighted, i.e., no factor of 20 or any other quality factor or RBE value used.

In subsequent sections, after Materials and Methods, four sets of results are presented:

1. Statements about linearity and dose threshold in cancer dose response from the assessed studies are listed, as are relative risk values for all studies discussed at doses below 0.2 Gy;
2. Linear fits to lung cancer ERR for the dose region 0–0.05 Gy are plotted, along with individual datapoints, to allow a (subjective) visual assessment of a hypothetical dose threshold in lung cancer dose response;
3. RBE values predicted for workers generated using IREP methodology are compared to those study values that were accompanied by published confidence limits; and
4. Illustrative IREP-based predictions for lung cancer ERR Gy⁻¹ are compared to ERR Gy⁻¹ values standardized to age 60 that were given in three studies (Gilbert et al. 2013; Labutina et al. 2013; Gillies et al. 2017a).

What these results might mean for future IREP methodology is then discussed.

MATERIALS AND METHODS

Prediction of RBEs for workers using IREP methodology

IREP methodology provides the steps for translating its relative effectiveness factors into predictions of RBEs (Appendix) for different reference radiations, including predictions of RBEs for use with protracted reference exposures at any dose and for any energy of ionizing photons. In the case of photon energy, IREP accounts for different biological efficiencies in three ranges, which are <30 KeV, between 30 and 250 KeV, and > 250 KeV. Although the external radiation received by plutonium workers contained gammas and x rays with photon energies <250 KeV, the percentage distribution does not appear in the published literature. External radiation came from fission products in the reactor, radiochemical plant, and plutonium production plant, including contamination from process leaks. The energy range for the protracted external radiation delivered in the Mayak study included considerable flux of gamma rays between 100 and 300 KeV for many worker situations (Vasilenko et al. 2007). As for Sellafield workers, although 90% of the doses were estimated to come from photon energies in the range of 100 KeV to 3 MeV (Thierry-Chef et al. 2007, 2015), Sellafield workers were exposed to some ionizing photons with energies < 250 KeV.

Therefore, in the absence of a known distribution for photon energies, three separate estimates of RBEs for workers have been made for IREP's three different photon energy ranges, all for low dose/low dose rate photons. If the published RBEs for workers are above the highest of the three IREP-derived predictions, the published RBEs would also be above any weighted sum of the three estimates.

The RBEs estimated in the worker studies are relative to the protracted external radiation to which workers were exposed and not to prompt exposures to which atomic bomb survivors were exposed. However, these RBE predictions can be compared to RBEs estimated from fits to epidemiological data for Mayak and Sellafield workers.

Predicting lung cancer ERR Gy⁻¹ using IREP output

Unlike predicting worker RBEs, no knowledge of the inner workings of IREP is necessary to obtain ERR Gy⁻¹. All that is needed is to convert IREP output for assigned share into excess risks, as is done in this section. The conversion is possible because assigned share (AS) and excess relative risk (ERR) have a one-to-one correspondence.

The output from the online calculator gives assigned share at various percentage credibility levels: 1, 5, 50, 95, and 99. The 99%-credibility limit, favorable to claimants, is the one used in adjudicating claims for compensation. Values for 2.5 and 97.5 were obtained through simulation of a lognormal distribution. A lognormal distribution was found to be consistent with the output percentage values provided (Appendix). Assigned share (AS), expressed as a fraction, is related to excess relative risk (ERR) by the equation:

$$AS = \frac{ERR}{1 + ERR}. \quad (1)$$

AS is a “probability of causation” assuming a simple, “independent of background” biological model (Beyea and Greenland 1999). AS is often labeled as PC/AS. The inverse of equation 1 gives the formula for converting AS output from IREP to the underlying ERR:

$$ERR = \frac{AS}{1 - AS}. \quad (2)$$

If ERR values at various credibility levels as determined by eqn (1) are plotted as a function of input dose, slopes of the curves can be determined to give numerical values for ERR Gy⁻¹. These in turn can be compared to results from epidemiological studies of plutonium workers, although some adjustment of an IREP-based result is necessary in principle to account for transfer of risk values, e.g., between worker populations in the US, the UK, and the Russian Federation (Appendix). For this purpose, the generic methodology used in IREP to transfer risk from Japanese atomic bomb survivors to US populations was adopted, which depends on the baseline mortality rates for that cancer in each country/area considered.

Linear fits to the ERR data for lung cancer to obtain the slope values (ERR Gy⁻¹) were made using the “lm” function in the R-statistical programming language (R_Core_Team 2020). For the regression dose ranges considered here, all fits were found to be linear with no intercept. They passed through the origin with a deviation that was within computational error.

The ERR Gy⁻¹ values were obtained for 3 smoking histories: Never, Former, and Current (with 20–39 cigarettes per day as a midrange choice).

IREP allows inputs for claimant's age at diagnosis. To most easily match IREP-based ERR Gy⁻¹ predictions for lung cancer to published epidemiological values, analysis was restricted to one attained age. Age 60 was chosen because ERR Gy⁻¹ results for lung cancer risks standardized to age 60 are often presented in the plutonium worker literature (Gilbert et al. 2013; Labutina et al. 2013; Gillies et al. 2017a). Each worker had a range of years when exposed to plutonium, but information sufficient to specify the distributions was not available. Analysis indicates, however, that the IREP lung cancer results are similar over a broad range of exposure ages (Table 2). Cumulative doses received at ages 30–50 can therefore be assigned to one exposure age to generate an approximate result. An age at exposure of 35 was chosen for this purpose. The resulting ERR Gy⁻¹ comparisons cannot be precise, but they should be useful for identifying broad trends in IREP-based risk predictions relative to epidemiologically obtained values.

The inverse of ERR Gy⁻¹ at the 99th-credibility percentile is called the compensation dose in this paper. Multiplying the compensation dose by ERR Gy⁻¹ gives an ERR of unity. At unity, the ERR is half the relative risk of 1 + ERR, which means the assigned share has reached 50% at 99th-credibility percentile, the minimum value for compensation under IREP. Compensation dose is strictly defined for the claimant condi-

tions generating the ERRs, which here include exposure at age 35 and lung cancer diagnosis at age 60.

Combining four plausible ERR Gy⁻¹ distributions

In some IREP analyses, there were models, such as absolute vs. multiplicative risk model, that needed to be combined because there was scientific and/or theoretical support for an in-between model. The IREP program handles this with an uncertain distribution for merging. In this paper, for a simplified example, there were four ERR Gy⁻¹ distributions, E_i, that were merged. A simple weighting scheme was used. Let s₁, s₂, s₃ be random variables uniformly distributed between 0 and 1. Then the merged distribution is written as

$$E_{\text{merged}} = s_1 * (s_2 * E_1 + (1-s_2) * E_2) + (1-s_1) * (s_3 * E_3 + (1-s_3) * E_4). \quad (3)$$

This is equivalent to treating E₁ and E₂ equally likely, as well as treating E₃ and E₄ equally likely, while also treating the two-paired combinations as equally likely.

Review of relevant plutonium epidemiology

Epidemiological studies of cancer in Mayak and pooled Mayak and Sellafield workers published since 2007 were examined (Table 1). Also included in the review were additional studies of plutonium lung cancer mortality in European worker populations (Grellier et al. 2017), 60% of whose cancer cases were Sellafield plutonium workers, and two studies of US plutonium workers, one published in 2004 (Rocky Flats) (Brown et al. 2004) and one in 2022 (Los Alamos National Laboratories) (Boice et al. 2022). In addition, consideration was given to a recent detailed assessment of the major plutonium worker studies, ICRP Report 150, which included point-value estimates of worker RBEs derivable from Mayak and Sellafield studies, although without corresponding estimates of confidence limits (Tirmarche et al. 2021).

A 2011 study of bone sarcomas in atomic bomb survivors was a non-plutonium study assessed for this paper (Samartzis et al. 2013). Older epidemiological studies that were available during the development of IREP, and therefore implicitly included in the IREP uncertainty analysis, were not generally considered.

RESULTS

Reported findings about plutonium dose response

Findings about plutonium dose response reported by the bulk of the modern plutonium studies are summarized in Table 1 (Tokarskaya et al. 1997, 2002, 2006; Gilbert et al. 2013; Labutina et al. 2013; Gillies et al. 2017a, 2017b; Stram et al. 2021). Also listed are early studies by Tokarskaya et al., which the Air Force report emphasizes. Not listed in Table 1 are studies carried out on workers at Los Alamos and Rocky Flats; they are discussed separately.

Table 2. ERR Gy⁻¹ for lung cancer incidence at age 60 predicted using output from NIOSH-IREP for a claimant who is a former smoker exposed to high-LET radiation at 5 earlier ages, all for various credibility percentiles.^a

| Credibility percentile | Age exposed | | | | |
|------------------------|----------------------|------|------|------|-------|
| | 20 | 25 | 35 | 50 | 55 |
| Male | ERR Gy ⁻¹ | | | | |
| 2.5 ^b | 0.58 | 0.46 | 0.40 | 0.38 | 0.003 |
| 5 | 0.82 | 0.64 | 0.58 | 0.54 | 0.005 |
| 50 | 5.0 | 3.8 | 3.8 | 3.6 | 0.13 |
| 95 | 30 | 23 | 22 | 22 | 1.80 |
| 97.5 ^b | 43 | 32 | 31 | 29 | 2.95 |
| 99 | 69 | 50 | 40 | 38 | 5.59 |
| Female | ERR Gy ⁻¹ | | | | |
| 2.5 ^b | 1.4 | 1.4 | 1.4 | 1.3 | 0.015 |
| 5 | 2.0 | 2.0 | 2.0 | 1.8 | 0.025 |
| 50 | 12 | 12 | 12 | 11 | 0.40 |
| 95 | 65 | 65 | 65 | 61 | 5.3 |
| 97.5 ^b | 90 | 90 | 90 | 84 | 8.7 |
| 99 | 121 | 121 | 121 | 117 | 16 |

^aAssigned share outputs from IREP converted to ERR for a range of doses (not shown) that were fitted to a linear function. Unweighted Gy. IREP output for chronic alpha exposure.

^bERR Gy⁻¹ values in this row were extrapolated assuming lognormal distribution.

In only one case in the post-2007 studies was a quadratic term retained in modeling for the plutonium dose response. That was for liver cancer, although the nonlinearity “diminished” when the internal doses were restricted to less than 2 Gy (Labutina et al. 2013). As for lung cancer, a quadratic term was suggestive in the European cohort study of Grellier et al. (2017), with a P value of 0.07. Note that none of the modern plutonium studies were able to successfully fit bone cancer data to a dose-response curve, which is not surprising given the paucity of cases.

Quotations from the modern Mayak studies about dose effects among the plutonium workers are presented in Supplemental Digital Content Table S-1, <http://links.lww.com/HP/A217>. No support was expressed by modern authors for a dose threshold in cancer dose response. As will be discussed, a number of the studies do, however, provide considerable information that allows examination of lung cancer risks at the very low end of the dose range, either by assessing model fits to ERR Gy⁻¹ restricted to the low dose range or by (subjective) visual assessment of clusters of datapoints whose risks seem high or low.

Published lung cancer RBE values for workers compared to predictions based on IREP documentation

A number of estimates of RBEs derived from data on plutonium workers have been published without confidence limits, as listed and discussed in Supplemental Digital Content, Text S-1, <http://links.lww.com/HP/A217>, but central values alone are not useful for compensation purposes. Thus, attention here is focused on the two published studies that included confidence limits along with central values (Gilbert et al. 2013; Gillies et al. 2017a). The comparison is indirect because various non-statistical (not random) uncertainties (Amrhein et al. 2019) in the study data were not included in confidence limits, a step that appears to have been an important part of the IREP development process when the original datasets were incorporated. Such assessments are best made by a review committee, so none were attempted for this paper.

An RBE value of 45 (21–240, 95% CI) was given for Mayak workers in the 2013 study by Gilbert et al. Assuming the underlying distribution was lognormal, the imputed 99th percentile would be 330 (Table 3). An alternate result was 33 (14, 98, 95% CI), with an imputed 99th-percentile value of 120 (Table 3). The alternate result was for a conditional model in which the modifying effects of sex and attained age were assumed the same for internal and external dose. Note that smoking was controlled for to the extent possible in this study (Gilbert et al. 2013).

In a 2017 study of pooled Mayak and Sellafield workers (Gillies et al. 2017a and b), an RBE value complete with confidence limits was given, but only for mortality and only for the slow, not the fast, case of plutonium nitrate solubility. The

Table 3. Worker radiobiological effectiveness factors (RBEs) for lung cancer *mortality* with confidence limits (95%) as reported in plutonium worker studies compared with lung cancer *incidence* predictions based on IREP methodology for 3 different protracted photon energy mixes.^a 99th-percentile values estimated.^b

| Source | RBE from published studies or IREP-derived | Details |
|---|---|-----------|
| Mayak (Gilbert et al., 2013) ^c | 45 (21, 240) ^d 99 th = 330 | Mortality |
| Mayak + Sellafield (Gillies et al., 2017a) ^c | 21 (9, 178) 99 th = 270 | Mortality |
| IREP-derived RBE for photon energies >250 KeV ^f | 18 (3.4, 100) 99 th = 140 | Incidence |
| IREP-derived RBE for photon energies >30 KeV and < 250 KeV ^g | 10 (1.4, 70) 99 th = 99 | Incidence |
| IREP-derived RBE for photon energies <30 KeV ^h | 7.8 (1.0, 54) 99 th = 78 | Incidence |

^aWorker external radiation was protracted, with a mix of photon energies. The study values were the only identified values complete with confidence limits. They cannot be used to infer plutonium risk from A-bomb external reference radiation without adjustment. Note the different endpoints: mortality for worker studies, incidence for IREP-based predictions.

^bAssuming a lognormal distribution for the RBE.

^cAdjusted for smoking.

^d33 (14, 98), 99th = 120, for a conditional model in which the modifying effects of sex and attained age were assumed the same for internal and external dose.

^ePooled study, not adjusted for smoking. This result applies only to cancer mortality and the case of slow solubility of Pu nitrate. Results were not given for fast solubility coefficient.

^fThis is the IREP alpha REF_L result, defined for photons with energies > 250 KeV. As a starting point for subsequent Table entries, the upper 95% confidence limit and the central value were used to generate 30,000 replications from a lognormal distribution.

^gTo obtain the statistics for this case, each replication of the alpha REF_L lognormal distribution (footnote f) was divided by a draw from a photon REF_L with a mixed distribution (Kocher 2005): 75% weight assigned to a lognormal distribution with 95% confidence interval between 1.0 and 5.0 (geometric mean of 2.2). 25% weight assigned to the value 1.0

^hTo obtain the statistics for this case, the distribution given in footnote g was further divided by draws from a triangular distribution with a minimum of 1.0, a maximum of 1.6, and a mode of 1.3 (Kocher 2005).

result was 21 (9–178, 95% CI), with an imputed 99th-percentile value of 270. There was no control for smoking. Because only 1 of 4 possible RBE results were given in the paper, there is a potential for selection bias in using the one set of numbers available, which must be borne in mind. Gillies et al. (2017a) made calculations for two solubility numbers for plutonium nitrate because, unlike absorption of plutonium oxide, there was no solid data on which to rely.

Putting aside the alternate conditional calculation made by Gilbert et al. (2013) for the moment, the other two RBE confidence limits for lung cancer presented in the epidemiological studies are higher by at least a factor of 1.8 than the values that would be predicted by IREP-based models for all three external photon energy ranges that span the mix of external radiation to which workers were exposed (Table 3). Therefore, although the percentage to be assigned to each external photon energy mix is not known, it doesn't matter because the epidemiology-derived RBEs are higher than any

combination of the three IREP-based versions that were calculated for different photon energy mixes. Thus, according to this illustrative comparison of worker RBEs, IREP makes it harder for a claimant to receive lung cancer compensation than would be justified based on the epidemiological values for RBE.

As for the alternate RBE estimated by Gilbert et al. (2013), its upper 95% confidence limit of 98 matches the IREP-based estimate for photons with energies >250 KeV and exceeds it for photon energies <250 KeV. It was not included in comparisons with IREP-based RBEs. Had it been, the combined scaling factor would have been 1.6, not 1.8. The alternate RBE was a second choice of the authors, with an unquantified ranking, making its scientific use unclear. Therefore, it was excluded from consideration, consistent with the veterans' canon.

Illustrative predictions of ERR Gy⁻¹ based on IREP output

Table 4 gives the ERR Gy⁻¹ predictions and confidence intervals that were obtained from the IREP output by transforming assigned share (AS) into ERR as a function of input dose, followed by linear regression as a function of input dose. Results for three smoking histories are given in Table 4 for males and females. Results for 90% CI come directly from regressions to the IREP ERR output. The 95% CI, however, were imputed using lognormal simulation as described in the Appendix. Table 4 shows that predicted ERR Gy⁻¹ declines with increased smoking and that the predicted ERR Gy⁻¹ is higher for females. IREP makes predictions about individual risk, dependent on claimant characteristics, such as smoking status. As such, Table 4 gives predicted conditional probabilities that may need to be statistically

Table 4. ERR Gy⁻¹ statistics for lung cancer incidence at age 60 predicted using NIOSH-IREP output for claimant exposed to high-LET radiation at age 35 for 3 smoking histories.

| Statistic identifier (C-dose ^a = compensation dose) | Smoking status at age 60 | | |
|--|--------------------------|----------------|---------------------------|
| | Never | Former | Current (20-39 cigs/d) |
| Male | | | |
| 90% CI | 6.9 (0.92, 47) | 3.8 (0.58, 22) | 2.7 (0.41, 17) |
| 95% CI ^b | 6.9 (0.63, 67) | 3.8 (0.40, 31) | 2.7 (0.29, 24) |
| 99 th percentile | 91 | 40 | 34 |
| C-dose ^a (Gy) | 0.011 | 0.025 | 0.029 |
| Female | | | |
| 90% CI | 20 (3.1, 130) | 12 (2.0, 65) | 8.8 (1.3, 52) |
| 95% CI ^b | 20 (2.2, 185) | 12 (1.4, 90) | 8.8 (0.94, 72) |
| 99 th percentile | 200 | 120 | 99 |
| C-dose ^a (Gy) | 0.0036 | 0.0083 | 0.01 |

^aCompensation dose (C-dose) is the inverse of the 99th-percentile ERR Gy⁻¹, corresponding to a 50% assigned share at 99th-credibility limit for a claimant exposed at age 35 and diagnosed with lung cancer at age 60.

^bValues in this row were extrapolated assuming a lognormal distribution.

combined to compare to epidemiological results for different subgroups (Appendix).

Also included in Table 4 are results for the 99th percentile and the "compensation dose." Compensation cutoff dose (C-dose in the Table) is the inverse of the 99th-percentile ERR Gy⁻¹, corresponding to a 50% assigned share at 99th-credibility percentile for a claimant exposed at age 35 and diagnosed with lung cancer at age 60. Compensation doses are presented as unweighted Gy, not Sv, although the absorbed dose would need to be converted to Sv (by multiplying by a factor of 20) before being entered into the IREP online calculator.

To compare illustrative predictions based on IREP output to published ERR Gy⁻¹, it is necessary to account for smoking histories. The three smoking history choices covered in Table 4, available in IREP output, do not always exactly match the smoking categories that have generally been used in studies of plutonium workers, which in most cases did not distinguish nonsmokers by Never and Former categories. In other words, the nonsmoking category in most of the epidemiological studies considered here is made up of an aggregate of former smokers and never smokers. The IREP-based ERR Gy⁻¹ predictions for different cohorts shown in Table 4 were aggregated separately for males and females using weighted pooling (Appendix) based on the smoking prevalence assumptions listed in Supplemental Digital Content Table S-2, <http://links.lww.com/HP/A217>.

Table 5 gives the ERR Gy⁻¹ comparisons for the study by Gillies et al. (2017a and b), whose confidence intervals were 90%-values, not 95%-values (Gillies et al. 2017a). Study results are only included in the Table for the case of fast plutonium nitrate absorption, which was generated by Gillies et al. (2017a) using parameters based on results for volunteers from the UK's Public Health England. Subsequent to publication, the UK absorption parameters were adopted for Mayak workers (Vostrotin et al. 2018).

As for the ERR Gy⁻¹ comparisons with the lung cancer studies by Labutina et al. (2013) and Gilbert et al. (2013), they are shown in Table 6 (95% CI). Note that the results from Gilbert et al. (2013) are for cancer mortality, not incidence, so only the comparisons to the results of Gillies et al. (2017a) and Labutina et al. are strictly equivalent as to end points. However, the ERR Gy⁻¹ values tend to be the same for mortality and incidence across the range of studies of plutonium workers considered here (Supplemental Digital Content Text S-2, <http://links.lww.com/HP/A217>).

In general, the ERR Gy⁻¹ illustrative age-60 predictions given in Tables 5 and 6 using IREP output have central estimates that are lower than the central values found in the epidemiological studies (Tables 5 and 6). However, it is the upper percentiles of the ERR Gy⁻¹ distributions that determine compensation in IREP. There the results are mixed. The regressions to Sellafield data alone (Table 5) have (imputed) 99th percentiles that are 2 to 3 times higher than the

Table 5. ERR Gy⁻¹ statistics for pooled lung cancer incidence of monitored workers at age 60 from Gillies et al. (2017a and b), compared with illustrative IREP-based incidence predictions made at age 60 for alpha particle exposure to US workers at age 35.^{a,b} 90% CI.

| Cohort | Statistic identifier (C-dose ^c) = compensation dose) | IREP-based illustrative prediction for US workers | Gillies et al. 2017a and b | |
|--------------------------------|---|--|--|--|
| | | | Dose range restrictions | |
| | | | None (all doses) Fast Pu nitrate ^d | Dose < 0.05 Gy Fast Pu nitrate ^d |
| Sellafield males | | | | |
| | ERR Gy ⁻¹ (90 th CI) | 3.8 (0.67, 23) ^c | 26 (-1.9, 65) | 59 (19, 110) |
| | 99 th percentile | 48 ^f | > 81 ^f | > 130 ^f |
| | C-dose ^b (Gy) | 0.021 | < 0.012 | < 0.0076 |
| Mayak males | | | | |
| | ERR Gy ⁻¹ (90 th CI) | 2.9 (0.46, 18) ^g | 7.3 (5.2, 10.2) | -0.08 (<0, 14) |
| | 99 th percentile | 38 ^f | > 11 ^f | > 20 ^f |
| | C-dose ^c (Gy) | 0.026 | < 0.091 | < 0.05 |
| Mayak females | | | | |
| | ERR Gy ⁻¹ (90 th CI) | 20 (2.1,180) ^h | 24 (12, 49) | NA |
| | 99 th percentile | 280 ^f | > 66 ^f | |
| | C-dose ^c (Gy) | 0.0038 | < 0.015 | |
| Sellafield males + Mayak males | | | | |
| | ERR Gy ⁻¹ (90 th CI) | 3.4 (0.55, 21) ⁱ | 7.5 (5.3, 10.4) | 8.4 (<0, 24) |
| | 99 th percentile | 44 ^f | > 12 ^f | > 30 ^f |
| | C-dose ^c (Gy) | 0.023 | < 0.083 | < 0.033 |

^aAlthough IREP-based predicted values for ERR Gy⁻¹ distributions were obtained for but one exposure age, the values are similar for a wide range of exposure ages (Table 2). Study results for male workers with monitored doses come from Tables 5 and S10 in Gillies et al. (2017a and b). Mayak females, age 60, from Table 3. CI for IREP-based predictions = credibility interval. The < and > signs indicate that the raw study values are limits because they have not been adjusted to account for non-statistical (nonrandom) uncertainties.

^bAccounting for risk transfer from study cohorts to a US population using IREP methodology would have led to a reduction of approximately +20%, -20%, and -9% in the 99th-percentile values for Mayak males, Mayak females, and Sellafield males, respectively, using 2005 lung cancer mortality rates (Supplemental Digital Content Tables S-3 and S-8, <http://links.lww.com/HP/A217>).

^cCompensation dose (C-dose) is the inverse of the 99th-percentile ERR Gy⁻¹.

^dAssumed nitrate absorption coefficient used in plutonium dose estimation.

^e4% never smokers, 96% former smokers for Sellafield Males. If the 96% is assigned to current smokers instead, the results drop to 2.8 (0.43, 18), 99th-percentile = 40.

^fIREP-based 99th values were based on logarithmic difference between upper CI and the central value, divided by 1.65 for 90% CI, which was used for the logarithmic standard deviation. For 99th values in the study column, a normal distribution was assumed. The result is not necessarily the same as the 99th percentile that would be found in actual study data.

^g76% current smokers with remaining 24% assumed to be never smokers for Mayak males. If assume 24% are former smokers, the IREP results increases to 3.3 (0.49, 24), 99th percentile = 56.

^h4% current smokers; 96% nonsmokers.

ⁱUsing the alternate smoking assumptions indicated in footnotes d and f, the IREP results are 3.0 (0.46, 21), 99th percentile = 48.

99th percentiles for the illustrative IREP-based ERR Gy⁻¹ values, suggesting that IREP might be underpredicting risk. On the other hand, for the Mayak worker data alone, the imputed 99th-percentile values are 2 to 5 times lower than the illustrative IREP-based values, suggesting that IREP might give too much benefit of the doubt to claimants. The 99th percentiles for the pooled Mayak and Sellafield data are also lower than the corresponding IREP-based values, although only by 30% when regressions are limited to < 0.05 Gy. Note that the study values for ERR Gy⁻¹ have not been adjusted for transfer of risk from a Russian or UK population with different lung cancer mortality rates, but this is not an adjustment that would change the overall picture. For instance, transfer might increase the imputed 99th-percentile values for Mayak males by about 20% (Supplemental Digital Content

Table S-3, <http://links.lww.com/HP/A217>), bringing them closer to, but still below, the IREP-based values. Transfer calculations were not possible for the pooled UK and Russian Federation cohorts using IREP methodology because of the mixed countries of origin.

Tempering all of these comparisons is the fact that the IREP-based calculations are only illustrative, as well as the fact that 99th-percentile values for the epidemiological studies do not account for study limitations. Ideally, before comparisons with IREP-based predictions were finalized, study values would be adjusted to quantitatively account for study limitations that go beyond the statistical measures of random error that come out of regression analyses (Amrhein et al. 2019) and possibly to account for unexplained variations across study results. Accounting for additional uncertainties

Table 6. ERR Gy⁻¹ statistics for lung cancer incidence and mortality at age 60 from 2 Mayak studies (2013) compared with illustrative IREP-based incidence predictions for age 60 made for exposure of US worker exposed to high-LET radiation at age 35. ^{a,b} 95% CI.

| Study and cohort | Statistic identifier (C-dose ^c) = compensation dose | IREP-based illustrative incidence | |
|---------------------------------|--|--|-------------------------------|
| | | prediction (for a US worker at age 60) | Mayak worker, full dose range |
| Labutina et al. 2013, incidence | | | |
| Mayak Non-smokers | | | |
| | ERR Gy ⁻¹ (95% CI) | 13 (1.0, 140) ^d | 31 (21.9, 40.8) |
| | 99 th percentile | 220 | > 43 ^e |
| | C-dose (Gy) | 0.0046 | < 0.023 |
| Mayak Smokers | | | |
| | ERR Gy ⁻¹ (95% CI) | 2.8 (0.31, 17) | 8.0 (5.2, 11.7) |
| | 99 th percentile | 25 ^e | > 12.4 ^e |
| | C-dose (Gy) | 0.040 | < 0.080 |
| Gilbert et al. 2013, mortality | | | |
| Mayak Males | | | |
| | ERR Gy ⁻¹ (95% CI) | 2.9 (0.32, 25) ^{b,f} | 7.4 (5.0, 11) |
| | 99 th percentile | 38 ^{b,e} | > 12 ^e |
| | C-dose (Gy) | 0.026 | < 0.083 |
| Mayak Females | | | |
| | ERR Gy ⁻¹ (95% CI) | 20 (2.1, 180) ^{b,g} | 24 (11, 56) |
| | 99 th percentile | 280 ^{b,e} | > 62 ^e |
| | C-dose (Gy) | 0.0038 | < 0.015 |

^aThe 2 epidemiologic study results come from Table 3 in each of their articles. Although IREP-based predicted values for ERR Gy⁻¹ distributions were obtained for but one exposure age, the values are similar for a wide range of exposure ages (Table 2). The < and > signs indicate that the raw study values are limits because they have not been adjusted to account for non-statistical (nonrandom) uncertainties.

^bAccounting for risk transfer from study cohorts to a US population using IREP methodology would have led to a reduction of approximately +20% and -20% in the 99th-percentile values for Mayak males and females, respectively, using 2005 lung cancer mortality rates (Supplemental Digital Content Tables S-3 and S-8. <http://links.lww.com/HP/A217>).

^cCompensation dose (C-dose) is the inverse of the 99th-percentile ERR Gy⁻¹.

^dAssumes all male nonsmokers are former smokers and all female nonsmokers are never smokers.

^e2.3 standard deviations above the central value. Based on an assumed normal distribution for regression coefficients, the standard deviation was obtained by dividing the difference between upper CI and the central value by 1.96. The 99th-percentile values here are not necessarily the same as the 99th percentile in the actual study data.

^f76% current smokers with remaining 24% assumed to be never smokers. If assume the 24% are former smokers, the results change to 3.3 (0.49, 24), 99th percentile = 56.

^g4% smokers; 96% never smokers.

can only increase the 99th-ERR Gy⁻¹ study percentiles and only decrease their inverses (compensation doses). For this reason, entries in Tables 5 and 6 have “<” and “>” symbols to indicate that the listed values are upper or lower limits, respectively.

Merging discordant information, the simplified example

There are four models for ERR Gy⁻¹ that are plausible given the new and old data, assuming the illustrative IREP-based predictions are reasonably correct and assuming more precise assessments verified them:

1. First, there is the current IREP approach that was used to predict ERR at a dose, d ;
2. The second ERR approach is based on doubling the upper confidence limit to match the higher values in the assessed RBEs from Gilbert et al. (2013) and Gillies et al. (2017a) This might be done in IREP by doubling the 95th upper confidence limit for the appropriate uncertain relative effectiveness factor (REF);

3. The third model that might play a role in compensation calculations is based on the direct ERR Gy⁻¹ incidence values found in the Gillies et al. study (2017a and b) for the Sellafield cohort, either for the full regression dose range or the dose range < 0.05 Gy; and
4. The fourth approach is to take the all-dose incidence results of Gillies et al. (2017a and b) for the Mayak cohort. This last term leads to a less favorable compensation dose for a claimant than is calculated by the current IREP program.

Reliance on the models in Gillies et al. (2017a and b) is consistent with their use in the ICRP Report 150 to compute lifetime risk from plutonium intake (Tirmarche et al. 2021), although here the separate Sellafield low dose regression (below 0.05 Gy) is also used as an alternate because it represents the model most favorable to a US veteran cohort. The two direct ERR terms taken from study results would need in principle to be adjusted for non-random uncertainty factors (Amrhein et al. 2019) before being incorporated into IREP

data, but this has not been attempted for the simplified example. Not considered for this simplified example is accounting for risk transfer from study ERR Gy⁻¹ to values appropriate for a US population, but this is a modest effect using the IREP transfer methodology with 2005 lung cancer mortality rates: approximately +20%, -20%, and -9% for Mayak males, Mayak females, and Sellafield males, respectively (Supplemental Digital Content Table S-3, <http://links.lww.com/HP/A217>).

The numerical weighting distributions needed to combine the four EER distributions would likely be best determined by consensus of a review committee. For the simplified example, a combination of uniform distribution discussed in the Methods section was used. No adjustment in weighting was made for the possible selection bias in the RBE-based values. In any case, given the uniform weighting, the lung cancer compensation dose was cut in half if the Sellafield term with regression doses was limited to below 0.05 Gy. If, on the other hand, the Sellafield term with no dose restriction on regression was used in the combining equation, the corresponding compensation dose remained roughly unchanged from the current IREP value.

A valid alternate approach in IREP to merging these four distributions would be to take only the term that gave the most favorable outcome for a claimant, which would be the Sellafield fit, which produces compensation cutoff doses 2-3 times lower than under the current IREP program, as shown in Table 5, depending on which regression dose range is considered.

Other lung cancer risk data for plutonium doses below 0.05 Gy

There is additional epidemiological information on ERR Gy⁻¹ values at low doses, which, due to the absence of results for a single worker age, could not be compared quantitatively to the IREP-based predictions. Nevertheless, the results are informative. For instance, the European lung cancer mortality case/control study, which did control for smoking, has (approximate) relative-risk values at internal plutonium doses below 0.05 Gy, as shown in Supplemental Digital Content Table S-4, <http://links.lww.com/HP/A217>, that are consistent with the Sellafield results given in Gillies et al. (2017a and b), which did not consider smoking (Gillies et al. 2017a and b). The Sellafield data dominates the European plutonium cohort, so the agreement provides additional support for the idea that controlling for smoking does not affect results for internal plutonium exposure. Although the European cohort study did control for smoking, the authors did recognize that the smoking data was limited (Grellier et al. 2017).

An earlier case/control study of lung cancers following plutonium exposures at Rocky Flats also found little effect of smoking on worker plutonium risks (Brown et al. 2004). Finally, a 2021 study of workers at Los Alamos National Laboratory (LANL) receiving low plutonium doses had only 19 lung cancers among the 447 workers with plutonium lung

doses >5 weighted mGy (weighting factor of 20 assumed). With so few cases exposed to plutonium, the study had limited statistical ability to assess risk with precision (Boice et al. 2022). The hazard ratio for mortality of 1.15 (95% CI 0.66, 1.99) obtained for lung cancer per (weighted) 100 mGy of internal plutonium exposure was compatible with plutonium risks obtained by regression to Mayak data, according to the authors. Note that a hazard ratio is approximately equal to a relative risk when the outcome is rare (VanderWeele 2020), as is the case with lung cancer. The results of these three studies are discussed further in Supplemental Digital Content Text S-3, <http://links.lww.com/HP/A217>, along with an explanation of why the results from Grellier et al. (2017) that were given as excess odds ratios are approximately equal to excess relative risks.

Graphical presentation of lung cancer ERR below 0.05 Gy plutonium dose

If hypothetical variations in lung cancer ERR Gy⁻¹ as a function of plutonium dose not included in IREP were to be taken into account in compensation decisions based on epidemiological results, then the dose region most relevant to cohorts like the Palomares cleanup veterans, based on Air Force dose assignments to Palomares veterans (Beyea and von Hippel 2019), would be the region below 0.05 Gy, i.e., below an equivalent dose of 1 Sv. Current compensation cutoffs in IREP already allow for lung cancer compensation for Palomares veterans at doses of 0.015–0.025 Gy for males diagnosed at age 60 after exposure at a wide range of ages, as can be deduced from the ERR Gy⁻¹ values given in Table 2, by taking their reciprocals at 99th-percentage credibility.

Risks for lung cancer below 0.02 Gy plutonium dose can be extracted from data presented in the paper by Gillies et al. (2017a and b) for cancer incidence and are plotted in Fig. 1 and Supplemental Digital Content Fig. S-1, <http://links.lww.com/HP/A217>, with 90% CI shown. Recall that it is the credibility limits for ERR Gy⁻¹ that determine compensation cutoff doses in NIOSH-IREP, not the central values. The worker data in Fig. 1 was obtained by fitting risk to dose data determined assuming a “fast” solubility for the subset of workers exposed to plutonium nitrate, mainly those who worked in radiochemical plants, as opposed to production plants, where plutonium oxides were the dominant exposure, at least for those with high exposures. At Sellafield, the default exposure assumption in Gillies et al. (2017a) was that 90% of those workers were exposed to the nitrate form of plutonium. Supplemental Digital Content Fig. S-1, <http://links.lww.com/HP/A217>, shows the low-dose results assuming slow nitrate solubility, which look broadly similar to the results assuming fast nitrate solubility shown in Fig. 1. In any case, no conclusion of this paper would change if results for the slow solubility assumption were used for plutonium nitrate intakes.

The raw data used for the plotted Sellafield and Mayak datapoints can be found in Supplemental Digital Content

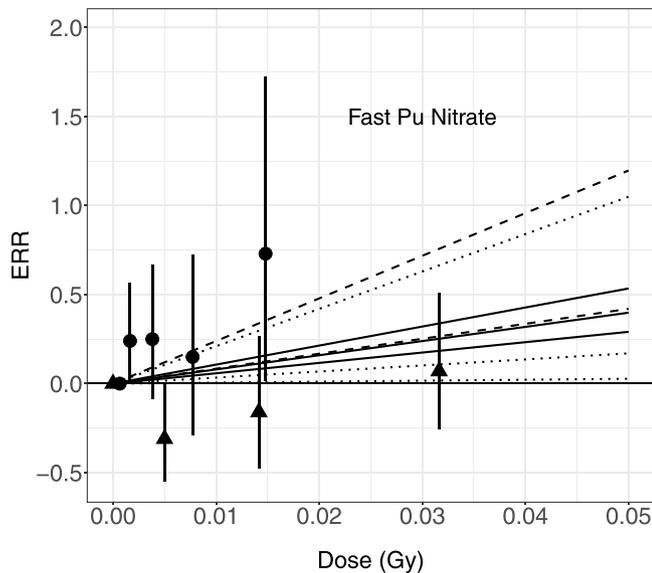


Fig. 1. Excess relative risks of lung cancer incidence in Mayak and Sellafield male workers for plutonium doses below 0.05 unweighted Gy as reported in Gillies et al. (2017a and b). 90% confidence limits on data points as reported in Table S10 of Gillies et al. (2017b). If available in study data, fitted linear lines have upper and lower 90% confidence limits around them. The linear fits were made to combined data for Mayak and Sellafield workers. Risk values shown were calculated in the Gilles et al. study assuming a fast absorption solubility ratio for plutonium nitrate. Circles = Sellafield data; Triangles = Mayak data; Solid lines = fit to all dose data; Dashed lines = fit to dose data below 0.05 Gy but without a lower 90% CI slope, given only as <0. Dotted lines are illustrative IREP-based predictions for male Mayak and Sellafield workers combined at age 60 exposed at age 35 (Table 5).

Table S-4, <http://links.lww.com/HP/A217>, expressed these as *relative risks*, and in Supplemental Digital Content Tables S-5 and S-6, <http://links.lww.com/HP/A217>, expressed as *excess relative risks*. Also plotted in Fig. 1 are linear fits to the pooled Sellafield and Mayak data, with 90% CI lines included. Gillies et al. (2017a and b) made these fits to ERR using (1) all of the pooled data (solid lines) and (2) the pooled data restricted to doses below 0.05 Gy (dashed lines). As expected, the upper-limit line rises faster for the 0.05 restriction.

In both Fig. 1 and Supplemental Digital Content Fig. S-1, <http://links.lww.com/HP/A217>, the Mayak excess relative risks are relatively low, with all central values falling below the lower confidence limit on the linear fit to all of the data. Although confidence limits on individual datapoints are useful for single comparisons, they are not so useful when assessing clusters without a combined statistical analysis. The smooth curves we may draw in our minds or sketch on top of a plot do not have statistically determined confidence bounds. The visual impression given to some analysts by the cluster of low-dose Mayak datapoints is that the risk has suddenly dropped well below a linear fit, suggesting threshold behavior, if limitations on the low-dose dosimetry are discounted (Gillies et al. 2017a). On the other hand, the Sellafield datapoint cluster has central values that tend to fall above the linear fit to the combined data, suggesting

supralinearity if limitations are discounted (Gillies et al. 2017a). When looking at the full set of datapoints, the idea of a dose threshold in lung cancer dose response put forth in the Air Force report does not seem plausible if visual assessment is to be the standard.

What can be said statistically? All four Sellafield datapoints lie above the fitted lines in Fig. 1, whereas all three Mayak datapoints lie below the fitted lines. Assuming it is equally likely for a point to fall above or below the linear-fit line, the chances of such a disparity are 1 in 64 using simple coin-tossing math, as was checked by a simulation in R statistical language using the function “rbinom.” In this case, a visual assessment of a *difference* in datapoint clusters would be supported by statistical analysis. However, a statistical difference between the cohorts does not necessarily mean that there is a real difference in dose-response; the difference may reflect data problems at low doses as suggested by Gillies et al. (2017a). Note that in this analysis, a disparity would be counted when a first point in the sequence was either above or below the linear-fit line, so the chances are 1 in 64, not 1 in 128.

Looking at the Mayak and Sellafield cohorts separately using the same simple coin-tossing approach provides only a weak signal for nonlinear dose response. Despite what a fit in the mind’s eye might suggest, the chances that three datapoints would randomly all lie above or below the line is 1 in 4. For 4 datapoints, the chances are 1 in 8.

The third set of (dotted) straight lines in Fig. 1 shows the illustrative incidence predictions based on IREP output, with 90% confidence limits. The IREP-based, upper-limit line matches the upper-limit line for the fits to pooled data with doses restricted to < 0.05 Gy, suggesting consistency with the epidemiological data for this dose range. On the other hand, the IREP-based, upper-limit line is higher than the upper-confidence-line for the fit without restriction on maximum dose, suggesting that IREP is overly favorable to claimants if the full-dose-range comparison is used. Not shown in the graphs are the linear fits to Sellafield data alone, because the upper limits rise so fast as to dwarf the separate datapoints and the linear lines including the IREP-based illustrations. Nevertheless, the unshown Sellafield lines would suggest that IREP is insufficiently favorable to claimants.

DISCUSSION

Dose threshold in lung cancer dose response

Overall, the five datasets with plutonium-associated risks estimated in the dose regions most relevant to veterans (Mayak, Sellafield, European cohort, LANL, Rocky Flats) are not definitive about ruling in or ruling out a dose threshold in lung cancer dose response, because fits to dose thresholds in dose response were not made by study authors and because visual assessment of datapoint clusters with large

confidence intervals on individual points cannot justify such an inference. However, four of the five show positive central effects, which suggests that it would be hard to argue that a meta-analysis of all the data would overturn the default hypothesis of no dose threshold unless one weights studies strongly to disfavor some or all of the four. Such strong weighting processes may take place in contested legal proceedings. For instance, a critic of the Sellafield data might point to Gillies et al. (2017a) stating that the Sellafield excess was “caused by an absence of lung cancer events at doses in excess of 50 mGy” (Gillies et al. 2017a). As a result, there would not have been events at high doses that could have stabilized a linear fit. Someone on the other side might point to the views of Gillies et al. (2017a) on the absence of precise low dose data in the Mayak worker cohort (MWC) due to a higher limit of detection for plutonium in bioassays: “This pattern could perhaps be interpreted as evidence of a threshold effect, but it is more likely related to the higher limit of detection in the early MWC, which results in an inability to detect variation in risk below a certain dose level” (Gillies et al. 2017a).

As for the 1997 study highlighted by the Air Force claiming a threshold in lung cancer dose response at 0.8 Gy (16 Sv) (Tokarskaya et al. 1997), there was no formal fit with confidence limits. The reported finding was based on the shape of ad hoc curves, usually straight lines drawn between six data points. Subsequent studies have found excess lung cancers well below such a high threshold, as shown in Supplemental Digital Content Table S-4, <http://links.lww.com/HP/A217>, which compiles all of the relative risks below 0.2-Gy plutonium dose given in the lung cancer studies listed in Table 1. Furthermore, when judging differences over time in the historical studies, it should be noted that the modern Mayak studies were carried out using the dosimetry developed by an international consortium of analysts to support and extend the work of Russian analysts (Napier 2017; Preston et al. 2017; Vostrotin et al. 2018). The modern Mayak studies also have more follow-up cases, which gives them more statistical power, although the statistical power is still weak for detecting risks other than to the lung due to the small number of excess cases.

As quoted in Supplemental Digital Content text S-4, <http://links.lww.com/HP/A217>, the Air Force report also implies that the findings of a 2013 study (Labutina et al. 2013) are similar to the 1997 study by Tokarskaya et al. However, the Labutina et al. study did not report a dose threshold in lung cancer dose response, publishing a Table, here reproduced as Supplemental Digital Content Table S-7, <http://links.lww.com/HP/A217>, that the Air force report interpreted as showing such a threshold. All that Table actually showed was that the two lowest relative risks (Supplemental Digital Content Table S-7, <http://links.lww.com/HP/A217>), which had dose categories below 0.2 Gy, did not exclude

the null at 95% confidence, an indication that the results could be compatible with a dose threshold but also compatible with many other possibilities that excluded a dose threshold. Visual inspection of datapoint clusters in a Table or graph is not a statistical assessment with confidence limits that can be used for compensation analysis. Moreover, as stated earlier, when visual inspection is expanded to include more than the relative risks reported in Labutina et al. (2013), such as the full set of studies listed in Supplemental Digital Content Table S-4, <http://links.lww.com/HP/A217>, any visual hint of threshold behavior disappears with the presence of the high Sellafield relative risk results compensating for the low risk Mayak datapoints (Fig. 1).

The Air Force report was so concerned about smoking effects on risk that it discounted an important analysis of plutonium worker data (Gillies et al. 2017a) and thus the datapoints in Fig. 1, because the study did not control for smoking. Yet, this analysis of the combined Mayak and Sellafield cohorts had the lowest dose categories and was deemed sufficient to be used to estimate RBEs in ICRP report 150 (Tirmarche et al. 2021). The Air Force concern was based on a 2002 study by Tokarskaya et al. (2002) that was carried out before the modern Mayak dosimetry analysis. Tokarskaya et al. had argued that, based on their results, failure to control for smoking could lead to overestimates of risk and obscure a threshold. However, other studies not mentioned in the Air Force report did control for smoking, one for Mayak workers (Gilbert et al. 2013) and one for a European cohort primarily made up of Sellafield workers (Grellier et al. 2017). The fits to these datasets were consistent with the findings of Gillies et al. (2017a and b) (Supplemental Digital Content Table S-4, <http://links.lww.com/HP/A217>). Thus, a dose threshold in lung cancer dose response that is hidden by smoking, a possibility claimed in the early Mayak study by Tokarskaya et al., seems unlikely. The data do not provide a basis for changing veteran compensation protocols or excusing any lack of defensible dosimetry analysis for Palomares veterans as found by a Veterans Appeals Court (CAVC 2020). An expanded discussion of smoking issues is presented in Supplemental Digital Content Text S-5, <http://links.lww.com/HP/A217>.

Unlike lung cancer, for bone cancer there is statistical evidence from a 2011 atomic bomb survivor study, that a dose threshold in dose response for bone cancer is compatible with the data at 95% confidence, given that the 95% lower confidence limit excluded zero dose. The central value found was 0.85 Gy (95% confidence interval, 0.12 to 1.85 Gy), with a linear dose-response above this threshold. This would imply a low central value for plutonium dose threshold in Gy, after dividing by an appropriate RBE and after disregarding the small neutron component contained in the 0.85 Gy estimate (Preston et al. 2004). On the other hand, given uncertainties in RBEs, a small

central value prediction for the corresponding plutonium dose would have a large uncertainty.

Not identifiable from the study was the behavior of 99th percentiles in ERR at and below the threshold dose value, which would govern compensation decisions at low doses. The number of bone cancer cases in the atomic bomb study was quite small, with only seven above the reference level of 12, sparsely spread out over the regression dose range. Confirmation and more details might be needed before modification of the bone cancer IREP risks, and uncertainty ranges should be made less friendly to claimants. Furthermore, bone sarcomas are rare, accounting for only 0.2% of all malignancies diagnosed in the United States (Franchi 2012) and so are unlikely to play a significant role in compensation, certainly compared to lung cancer. For this reason, focus here has been on lung cancer.

Uncertainties in study results not included in comparison with IREP-based predictions

In judging differences between the IREP-based illustrative predictions and the ERR Gy⁻¹ epidemiological confidence limits, it needs to be borne in mind, as mentioned earlier, that the full IREP approach requires adding uncertainties to study numbers that have not been accounted for in the raw ERR Gy⁻¹ results presented in epidemiological studies. Study results can differ due to different researcher choices in data analysis and statistical methods (“researcher degrees of freedom,” “auxiliary hypotheses”) that may not be accounted for in confidence limits (Forstmeier et al. 2017; Amrhein et al. 2019). Such private choices have been put forward to explain in part the so-called “crisis of unreplicable research” (Amrhein et al. 2019), which appears to be a concern in a great number of fields of science (Baker 2016).

With the exception of one study that considered dose uncertainty (Stram et al. 2021) (Supplemental Digital Content Text S-6, <http://links.lww.com/HP/A217>), the epidemiological confidence limits include only statistical estimates of uncertainties due to randomness in data, whereas there are a number of additional specific uncertainties in worker studies that have been discussed in the literature (Gillies et al. 2017a; Tirmarche et al. 2021). These are listed in Table 7. All of these possibilities could be most problematic for the low dose categories but less problematic for the higher dose categories with their greater percentages of excess cases.

The disparity in risk between the Sellafield and Mayak low-dose data clusters (Fig. 1 and Supplemental Digital Content Fig. S-1, <http://links.lww.com/HP/A217>), which is unlikely to be due to chance, is testimony that uncaptured uncertainties are large. The existence of uncaptured uncertainties should not be surprising, because historical dose reconstruction of protracted exposures can be a difficult undertaking and may require the making of many assumptions.

One of the uncounted uncertainties listed in Table 7 is the effect of missing data for both plutonium dose and smoking

Table 7. Uncertainties in worker study results discussed in the literature that can in principle introduce variance not always accounted for in confidence limits.

| Uncertainty type | Uncertainty |
|-----------------------------------|---|
| Dosimetry | Uncertainty in assessment of dose modeling parameters, such as those governing deposition and absorption in respiratory tract. ^a Challenges in assessing organ-/tissue-specific doses: “uncertainties associated with internal dose assessments based on bioassay data can be quite large” (Tirmarche et al. 2021). Low limits of detection, which is particularly an issue for Mayak results in the low-dose region. |
| Different experimental conditions | Generic issues of selection bias, ascertainment bias, loss to follow-up, misclassification, radiation exposure and risk |
| Modeling | Question as to the extent that Mayak results can be extrapolated to low doses and other populations. Treatment of joint occurrence of internal and external exposure data. ^{b)} |
| Missing data | Large percentage of missing bioassay data, particularly for Mayak workers, in full or in part, which in some studies was handled by a simple occupation-based surrogate, without any multiple imputation. Data for smoking was also missing for a high percentage of individuals in full or in part |

^aNote that the paper by Stram et al. does account for plutonium dosimetry uncertainties, although it does not account for the uncertainty associated with the simple surrogate used for missing bioassay data.

^bGilbert et al. found no evidence for modification of the external dose response by sex, age and smoking, but uncertainty ranges were not presented (Gilbert et al. 2013).

^ctreatment of multiplicative vs absolute risk model in smoking interaction.

status. Here is an area where techniques exist to help quantify the uncaptured uncertainty, although cooperation of those with access to the raw data are needed. For instance, a number of the studies (Labutina et al. 2013; Gillies et al. 2017a; Stram et al. 2021) used a 6-category surrogate index for unmonitored plutonium workers derived from occupational history, which does not account for individual uncertainty and is a methodology that can perform badly (Greenland 1995). Doses were not imputed (say, using a study-wide covariate matrix) (van Buuren and Groothuis-Oudshoorn 2011). Had they been, multiple imputation could have been used to quantify variance due to missing dose and smoking data.

The study by Gilbert et al. generally did not use the surrogates for assessment of ERR Gy⁻¹ as opposed to RBE estimates and did not find much difference in ERR Gy⁻¹ when surrogates were included (Gilbert et al. 2013). However, excluding unmonitored workers and unmonitored periods of work, as done by Gilbert et al., introduces its own problems, leading to a large amount of missing data that could affect the true variance of results (Greenland 1995; White and Carlin 2010). Unmonitored status may not have occurred at random, and true dose within surrogates may have unusual distributions; assignment to surrogate categories has unknown misclassification potential.

Ideally, in the face of the many unresolved uncertainties in lung cancer risks to workers from plutonium exposure, there would be sufficient cohorts studied in different populations to allow a direct assessment of between-study variance at low doses, which could help account for uncaptured uncertainties affecting low doses. However, as discussed in this paper, the number of cohorts in different populations with clear results is limited at the present time to Mayak, Sellafield, and the European cohort (the European cohort includes some study subjects other than Mayak and Sellafield workers). With so few results available for low doses, reliance on the judgment of a committee of experts, including some of the researchers involved in the plutonium worker studies, may be the best way at the current time to quantify any increase in worker study confidence limits to account for uncounted variance that might be relevant to compensation decisions.

Implications for regulatory policy

Although there is no need to consider changes relative to threshold for lung cancer dose response, there is a challenge posed by the conflicting signals in the comparisons of the IREP-based predictions and the epidemiological results. There are four options suggested by the data that are considered in this paper.

Option 1 - Make no changes to IREP. Multiple comparisons of IREP predictions with data are bound to produce discrepancies, because “Observed effect sizes can easily differ across settings” (Amrhein et al. 2019). The fact that some indications from the comparisons are more favorable to claimants than the IREP program would decide today, and some are less favorable, suggests that on balance, IREP got it right, which is a testimony to careful analysis. Differences appear to be in the noise of the data and predictions. Still, regulatory policy has to consider a wide range of issues, including the wider purpose of the policy, as well as veterans’ and public perception. Getting it right on balance and within the noise may not be sufficient.

Option 2 - Make IREP uncertainty bounds for lung cancer dose-dependent based on the pooled Sellafield and Mayak ERR Gy⁻¹ results.

These results (Gillies et al. 2017a) were the data of choice for analyzing RBEs in ICRP Publication 150 (Tirmarche et al. 2021) and so deserve consideration as the basis for any IREP adjustment. Table 5 shows that the IREP-based, illustrative predictions for the 99th percentile appear to be too high compared to the pooled results, with the extent depending on the dose range chosen for regression. The difference is a modest 30%, when doses are restricted to < 0.05 for regression analysis, and that difference would likely be overcome should consideration of study limitations be taken into account as previously discussed. However, for the full dose range, more than a 3-fold overprediction in 99th percentile occurred (Table 5). To incorporate the difference, a

modification could be made to either IREP’s lung cancer ERRs or the underlying parameter distribution REF_L so that confidence limits were retained at doses below 0.05 Gy but narrowed above, with a transition region in between. This would be equivalent to establishing a new dose dependent DDREF distribution for plutonium alpha particle exposures, allowing for possible deviation from dose linearity, upward or downward, but not necessarily changing dose linearity on average. Such a novel change might better reflect the epidemiological plutonium results but would likely be moot for many lung cancer claimants, because the current lung cancer compensation cutoff dose is already in the low dose region for claimants like the Palomares veterans. However, were the change made to REF_L, which is the starting point for all plutonium cancer risks in IREP, the change would not necessarily be moot for solid cancers other than lung that have higher lower limits for cutoff doses for compensation. For those cancer types, compensation doses might be increased.

Option 3 - Combine varying and discordant lung cancer ERR Gy⁻¹ distributions.

As in the simplified example, where there were four plausible models that have some support given the new and old data, assuming the illustrative predictions are reasonably correct. In the simplified example with its neutral weighting factors, the lung cancer compensation doses were unchanged, when the full regression dose range was used with the Sellafield data of Gillies et al. (2017a). The compensation doses were cut in half (i.e., made more favorable to claimants) when the regression dose range was kept to low doses (i.e., those below 0.05 Gy). Thus, it is possible that an attempt to merge the new information with the old would lead to no changes in IREP, or possibly a reduction in compensation dose by half. There are additional datasets that might be added to the merger, including the results for the European cohort, LANL, and Rocky Flats, as well as any post-2007 animal and cellular data that covers the same dose range, but their inclusion was not considered for this paper.

There is scientific advantage to combining datasets or averaging over them, which in effect averages over limitations in the different studies, allowing confidence limits to be based on the results from the combined analysis. This is consistent with the IREP approach of being scientific whenever possible and giving the benefit of the doubt, in general, through use of the 99% credibility criterion. However, deciding on the weighting factors and their distributions would be challenging. For instance, there is the possibility of selection bias in the limited number of RBE study values available that have accompanying confidence limits, suggesting some discounting of their use should be made.

Option 4 - Adjust IREP plutonium models to match the Sellafield lung cancer results.

This could be done in IREP by modifying the confidence limits for the IREP parameter, REF_L, or by adding an ad hoc adjustment to calculated lung cancer ERR. IREP

does not merge data when there is little basis for selecting or weighting one alternative over the other. In such a case, the most claimant-favorable alternative is chosen (Kocher et al. 2008), which here might be the 99th-percentile ERR Gy⁻¹ estimated for the Sellafield data alone. Even though it is scientifically weak to rely on a single study population (Amrhein et al. 2019), the policy argument in favor is that the Sellafield workers and their dose ranges are probably a better match for, and more relevant to, US claimants. Also, the monitoring at Sellafield was relatively strong, with approximately 500,000 urine sample results available over a long period of time for over 12,000 plutonium-monitored workers (Tirmarche et al. 2021).

Until risks determined in the studies of the Mayak and Sellafield contingents are adjusted to account for uncounted study limitations and fully transformed to apply to a US population, the policy argument for taking the Sellafield results is strong. Consider the high relative risk confidence limits on the datapoints for the Sellafield cohort at low doses (Fig. 1 & Supplemental Digital Content Fig. S-1, <http://links.lww.com/HP/A217>) and the high 99th-percentile limits for the Sellafield ERR Gy⁻¹ estimated in Table 5. Specifically, the Sellafield 99th percentile was 1.7 times higher than the 99th percentile for the IREP-based illustrative prediction when considering the full regression dose range and 2.7 times higher considering the low regression dose range < 0.05 Gy. Should not such findings for a relevant cohort, if verified by subsequent analysts, be sufficient to change 99th lung cancer percentiles in IREP? In any case, if only a factor of 2 or 3 is involved in compensation cutoff doses, giving the extra benefit of the doubt to claimants might not even be an issue. Waiting for the next update of the Sellafield and Mayak studies might be preferable from the scientific perspective. With more data, both the Sellafield and Mayak confidence limits should tighten, but veterans exposed to plutonium in the 1960's, like the Palomares cohort (Beyea and von Hippel 2019) and the Thule cohort (USAF 1970), have limited time to wait. Also, ironically, more data can lead to more stratification of analysis with a greater number of statistical tests being reported, thereby introducing a potential multiple comparisons problem and making it likely that a subset of the expanded number will continue to show large uncertainties in 99th percentiles relevant to some claimants.

For Options 2-4, there is another decision to be made: namely, deciding how to adjust 99th percentiles for cancers other than lung for which there is no equivalent plutonium data. The easiest solution would be to keep changes in REF_L that might be made for lung cancer, which would leave relative variations in plutonium risk by cancer endpoint to be determined by the proportionate risks found in the atomic bomb survivors.

Procedures for updating NIOSH-IREP

There are administrative rules for updating NIOSH-IREP (CFR 2021a). Recommendations may come internally

from NIOSH or from the Advisory Board on Radiation and Worker Health, the Department of Labor, independent reviews by scientific organizations, or public comment. If NIOSH then decides to proceed with substantive changes to NIOSH-IREP, the Advisory Board on Radiation and Worker Health reviews them, with NIOSH thereafter addressing any recommendations of the review before proceeding with the change. In the final stages, proposed changes must appear in the Federal Register allowing for public comment and response by the agency before actual implementation of a change in IREP. Until such time as adjustments might be made in the IREP program in light of the new epidemiological data, claims assessors might still take the new data into account, as could independent experts writing in support of individual claimants.

NIOSH and advisory committees would want to be highly confident in the science before making it more difficult for a veteran claimant. What should not be done is to try to infer nonlinearity visually, which is inherently subjective and vulnerable to confirmation bias. The generic problem with trying to infer nonlinearity using visual interpretation of multipoint data is that the likelihood of finding false positives is unknown, so confidence limits on a result cannot be correctly set. For instance, how many combinations of clustered datapoints would be judged visually to indicate nonlinearity? Once policy is based on visual inferences from data with large fluctuations in central values, the connection to science can be lost. Moreover, a visual basis is vulnerable to critics of the policy finding counterexamples in the literature, where the fluctuations from nonlinearity visually appear to contradict the original claim, as was the case with the Air Force claim of dose nonlinearity in the Mayak data.

Strengths and limitations

Strengths of this study include consideration of all of the large, modern epidemiological plutonium studies, not a selected subset as was the choice made in the Air Force report. In addition, a variety of IREP-based statistics have been compared with epidemiological data. In particular, illustrative IREP-based predictions for ERR Gy⁻¹ have been developed, translating IREP output assigned shares into quantities that can be roughly compared to the epidemiological data, albeit with caveats. A simplified example was used to show how disparate lines of evidence in this case might be combined for compensation purposes using a method that is found in IREP.

There are a number of limitations of this paper to consider. Only one representative exposure year was entered into IREP rather than a range of years spread out over employment, although for most years there was little difference. Incomplete transfer of lung cancer risk to a US population from risks determined in the Russian Federation and the UK was an additional limitation. There was a possible selection bias in the available RBE estimates that met the criterion of having accompanying confidence limits.

As discussed in Table 7 and accompanying text, there are also limitations in the original studies that may have led to underestimates in confidence limits. These limitations carry over to comparisons made here to the study results. Finally, there is an inherent limitation in IREP itself: assigned shares are imputed probabilities, not true probabilities, and thus have modeling uncertainties implicitly attached to them. Different models for cancer can give different probabilities of causation for the same dose and assigned share (Beyea and Greenland 1999). Furthermore, IREP does not add uncertainty to account for individual susceptibility. These latter two limitations, however, would have to contribute a very large uncertainty to make an important difference in the 99th percentiles.

CONCLUSION

A review of the post-IREP epidemiological plutonium data and comparison with IREP-based predictions indicates that:

1. A 1997 claim of a dose threshold in Mayak lung cancer dose response was not borne out by subsequent studies. As for the visually anomalous central datapoint values in the low end of the Mayak dose response distribution (Fig. 1), it is not valid statistical inference to focus on visual hints of nonlinear effects. Nor would it be valid to make inferences based on visual assessment of the Sellafield cluster. Clusters of Sellafield and Mayak datapoints at low doses relevant to military personnel like the Palomares veterans appear on different sides of the linear-fit lines, as shown in Fig. 1, with a 1-in-64 chance of occurring randomly, suggesting possible systematic errors, which would be best handled in a compensation context by pooling datasets or taking the one most favorable to the claimant, if there was not a clear scientific basis for the pooling;
2. A 2011 study of atomic bomb survivors found evidence for a dose threshold in *bone cancer* dose response (following low-LET exposure), specifically a finding that the confidence interval did not include the null at 95% confidence, although the number of cases was small. What this would mean for a 99th-percentile ERR around and below the threshold point, and hence what it would mean for compensation decisions at low doses, is neither identifiable from study results nor easily simulated;
3. An illustrative IREP-based prediction of RBE for Mayak conditions suggests that IREP might underpredict 99th percentile risk compared to worker epidemiological data, thereby producing a compensation cutoff dose that is insufficiently favorable to claimants;
4. Mayak values for lung cancer ERR Gy⁻¹ suggest that IREP might overpredict 99th-percentile risk based on illustrative calculations, thereby producing compensation cutoff doses that are too favorable to claimants. Sellafield values for lung cancer ERR Gy⁻¹ suggest the opposite:

- IREP might underpredict 99th-percentile risk, especially if regressions are limited to doses below 0.05 Gy;
5. A simplified example combining four lines of evidence, new and old, suggests that IREP-based predictions for lung cancer are consistent with the results for the combined dataset when the regressions for the Sellafield evidence component include the full dose range, but suggest that the appropriate compensation cutoff should be made more favorable to US claimants by a factor of 2 if the Sellafield data component is only regressed for doses < 0.05 Gy; and
6. Both the epidemiological results and the IREP-based, illustrative calculations have uncalculated uncertainties that make the comparisons tentative.

Although tentative, these mixed results present a challenge for NIOSH and its review committees. It is possible to imagine a number of outcomes for changing ERR Gy⁻¹ values for lung cancer or changing the IREP radiation effectiveness factors (REFs) that are used to scale plutonium risks from risks determined in the studies of atomic bomb survivors. Four possibilities were discussed, including (1) no change, and (2) widening the uncertainty bands for REFs at plutonium doses below 0.05 Gy and tightening them above 0.05 Gy. Another possibility was (3) combining ERR Gy⁻¹ distributions as in the simplified example presented here, which would be difficult to do reliably at this point due to the many judgments necessary to choose weighting parameters. Yet another possibility considered was (4) using the Sellafield data to adjust REFs or ERR, given the likelihood that Sellafield workers better match US claimants. Making such adjustments would be the most favorable option for claimants, which is the appropriate choice for compensation purposes when the alternates are scientifically unclear.

In its reports and any future reassessments, the Air Force should avoid trying to infer nonlinearity visually, which is inherently subjective and vulnerable to confirmation bias. The Air Force reports should take into account the full set of Mayak studies as well as Sellafield and European cohort data. It is not appropriate to discount the Sellafield data because of the absence of control for smoking in the study by Gillies et al. (2017a), particularly because controlling for smoking has not been shown to make any major difference for plutonium risks in the many studies discussed that have controlled for it. The only change reported by smoking status in any modern Mayak study was in the lung cancer ERR Gy⁻¹ for external radiation, not internal radiation (Stram et al. 2021), a change that, if used, might decrease RBEs from the one of two plutonium worker studies reporting RBE with confidence limits that did not control for smoking (Gillies et al. 2017a).

Although IREP has limitations (Beyea and Greenland 1999; CRS 2021), it remains a valid compensation tool, informed by consistent scientific analysis. IREP provides a

framework that merges science and policy in a workable, transparent manner that removes any need for toxic tort litigation in federal and state courts by giving benefit of the scientific doubt to the veteran. It may be argued that the hint of supralinearity and/or high lung cancer ERR Gy⁻¹ found in the Sellafield data has introduced new scientific uncertainty at the lowest doses, but there is no guarantee that the IREP 99th-percentile values would change depending on how the new uncertainty were to be combined with all of the other evidence.

Prior to any decision by NIOSH, veterans still have a possible way to get the Sellafield data at least considered by the Veterans Administration if they can find an expert deemed qualified to support their case. Experts acting on behalf of veterans can always submit supportive affidavits. They could argue perhaps that the ERR Gy⁻¹ distributions implicit in IREP should be widened at low doses, given the lung cancer results from Sellafield and the European Cohort, especially in a compensation program that tries to give the benefit of the doubt to the claimant when the science is not clear. Furthermore, the Veterans Administration could, depending on the outcome of its Palomares appeal, recommend such a position itself.

Acknowledgments—I thank Frank N. von Hippel for comments on an earlier version of the manuscript. I thank F. Owen Hoffman and David Kocher for reviewing descriptions of REF_L as presented in an earlier version of the manuscript.

REFERENCES

- Amrhein V, Trafimow D, Greenland S. Inferential statistics as descriptive statistics: there is no replication crisis if we don't expect replication. *The American Statistician* 73:262–270; 2019.
- Baker M. 1,500 scientists lift the lid on reproducibility. *Nature* 533:452–454; 2016.
- Beyea J, Greenland S. The importance of specifying the underlying biologic model in estimating the probability of causation. *Health Phys* 76:269–74; 1999.
- Beyea J, von Hippel FN. History of dose, risk, and compensation assessments for US veterans of the 1966 plutonium cleanup in Palomares, Spain. *Health Phys* 117:625–636; 2019. Available at https://journals.lww.com/health-physics/Fulltext/2019/12000/History_of_Dose,_Risk,_and_Compensation.13.aspx. Accessed 19 August 2021.
- Boice JD, Cohen SS, Mumma MT, Golden AP, Howard SC, Girardi DJ, Ellis ED, Bellamy MB, Dauer LT, Samuels C, Eckerman KF, Leggett RW. Mortality among workers at the Los Alamos National Laboratory, 1943–2017. *Internat J Radiat Biol* 98:722–749; 2022.
- Brown SC, Schonbeck MF, McClure D, Baron AE, Navidi WC, Byers T, Rutenber AJ. Lung cancer and internal lung doses among plutonium workers at the Rocky Flats Plant: a case-control study. *Am J Epidemiol* 160:163–72; 2004.
- CAVC. Skaar v. Wilkie, 121720 USCAVC. 17–2574; 2020.
- CFR. 42 CFR § 81.12. Procedure to update NIOSH-IREP. 42 CFR § 8112; 2021a.
- CFR. Claims based on exposure to ionizing radiation. 38 CFR § 3.311; 2021b.
- CFR. Methods for conducting dose reconstruction under the Energy Employees Occupational Illness Compensation Program Act of 2000. 42 CFR Part 82; 2021c.
- CFR. Guidelines for determining probability of causation under the Energy Employees Occupational Illness Compensation Program Act of 2000. 42 CFR part 81; 2022.
- CRS. Update: the radiation exposure compensation act (RECA): compensation related to exposure to radiation from atomic weapons testing and uranium mining. Washington, DC: Congressional Research Service; 2021.
- DVA. Notice of case development. In the United States Court of Appeals for Veterans Claims Victor B. Skaar, Appellant, v. Vet. App. No. 17-2574, Robert L. Wilke, Secretary of Veterans Affairs, Appellee; 2020.
- Forstmeier W, Wagenmakers E-J, Parker TH. Detecting and avoiding likely false-positive findings—a practical guide. *Biol Rev* 92:1941–1968; 2017.
- Franchi A. Epidemiology and classification of bone tumors. *Clin Cases Miner Bone Metab* 9:92–95; 2012.
- Gilbert ES, Sokolnikov ME, Preston DL, Schonfeld SJ, Schadilov AE, Vasilenko EK, Koshurnikova NA. Lung cancer risks from plutonium: an updated analysis of data from the Mayak worker cohort. *Radiat Res* 179:332–342; 2013.
- Gillies M, Kuznetsova I, Sokolnikov M, Haylock R, O'Hagan J, Tsareva Y, Labutina E. Lung cancer risk from plutonium: a pooled analysis of the Mayak and Sellafield worker cohorts. *Radiat Res* 188:645–660; 2017a.
- Gillies M, Kuznetsova I, Sokolnikov M, Haylock R, O'Hagan J, Tsareva Y, Labutina E. Supplement: lung cancer risk from plutonium: a pooled analysis of the Mayak and Sellafield worker cohorts. *Radiat Res* 188:645–660; 2017b.
- Greenland S. Dose-response and trend analysis in epidemiology: alternatives to categorical analysis. *Epidemiol* 6:356–65; 1995.
- Greenland S. Relation of probability of causation to relative risk and doubling dose: a methodologic error that has become a social problem. *AJPH* 89:1166–1169; 1999.
- Grellier J, Atkinson W, Bérard P, Bingham D, Birchall A, Blanchardon E, Bull R, Guseva Canu I, Challeton-de Vathaire C, Cockerill R, Do MT, Engels H, Figuerola J, Foster A, Holmstock L, Hurtgen C, Laurier D, Puncher M, Riddell AE, Samson E, Thierry-Chef I, Tirmarche M, Vrijheid M, Cardis E. Risk of lung cancer mortality in nuclear workers from internal exposure to alpha particle-emitting radionuclides. *Epidemiol* 28: 675–684; 2017.
- Harper CJ. Give veterans the benefit of the doubt: Chevron, Auer, and the veteran's canon. *Harv JL Pub Poly* 42:931; 2019.
- Hunter N, Haylock R. Radiation risks of lymphoma and multiple myeloma incidence in the updated NRRW-3 cohort in the UK: 1955–2011. *J Radiol Prot* 42:011517; 2022.
- Islami F, Torre LA, Jemal A. Global trends of lung cancer mortality and smoking prevalence. *Transl Lung Cancer Res* 4: 327–338; 2015.
- Kocher DC, Apostoaei AI, Henshaw RW, Hoffman FO, Schubauer-Berigan MK, Stancescu DO, Thomas BA, Trabalka JR, Gilbert ES, Land CE. Interactive Radioepidemiological Program (IREP): a web-based tool for estimating probability of causation/assigned share of radiogenic cancers. *Health Phys* 95:119–147; 2008.
- Kocher DC, Apostoaei AI, Hoffman FO. Radiation effectiveness factors for use in calculating probability of causation of radiogenic cancers. *Health Phys* 89:3–32; 2005.
- Labutina EV, Kuznetsova IS, Hunter N, Harrison J, Koshurnikova NA. Radiation risk of malignant neoplasms in organs of main deposition for plutonium in the cohort of Mayak workers with regard to histological types. *Health Phys* 105: 165–176; 2013.

- Napier BA. The Mayak worker dosimetry system (MWDS-2013): an introduction to the documentation. *Radiat Protect Dosim* 176: 6–9; 2017.
- NIOSH. Summary of NIOSH's re-examination of lymphoma target organ selection. The National Institute for Occupational Safety and Health [online]. 2005. Available at <https://www.cdc.gov/niosh/ocas/pdfs/misc/lymphoma/summary.pdf>. Accessed 11 February 2022.
- Otchin NS. The use of the Interactive Radioepidemiological Program (IREP) by the Department of Veterans Affairs [online]. 2007. Available at <https://healthdocbox.com/Cancer/118984667-Comments-for-the-meeting-of-the-veterans-advisory-board-on-dose-reconstruction-vbdr-march-7-2007-las-vegas-nv.html>. Accessed 11 February 2022.
- PACER. Skaar v. McDonough, United States Court of Appeals for the Federal Circuit, Docket # 21-1812; 2021.
- Preston DL, Pierce DA, Shimizu Y, Cullings HM, Fujita S, Funamoto S, Kodama K. Effect of recent changes in atomic bomb survivor dosimetry on cancer mortality risk estimates. *Radiat Res* 162:377–389; 2004.
- Preston DL, Sokolnikov ME, Krestinina LY, Stram DO. Estimates of radiation effects on cancer risks in the Mayak worker, Techa River and atomic bomb survivor studies. *Radiat Protect Dosim* 173:26–31; 2017.
- R_Core_Team. R: a language and environment for statistical computing [online]. 2020. Available at <http://www.R-project.org>. Accessed 4 November 2020.
- Rademacher SE. Plutonium exposures to personnel assigned to Johnston Atoll. Kirtland Air Force Base, Texas, Air Force Safety Center; 2016.
- Rademacher SE. Topical issues for assessment of dose to Palomares accident recovery workers. United States Air Force Safety Center [online]. 2020. Available at <https://www.safety.af.mil/Portals/71/documents/Weapons/Palomares%20Dose%20Assessment%20Rpt%207%20Jul%202020.pdf?ver=2020-08-05-113936-200>. Accessed 19 August 2021.
- Richardson DB, Sugiyama H, Wing S, Sakata R, Grant E, Shimizu Y, Nishi N, Geyer S, Soda M, Suyama A, Kasagi F, Kodama K. Positive associations between ionizing radiation and lymphoma mortality among men. *Am J Epidemiol* 169:969–976; 2009.
- Robins J, Greenland S. The probability of causation under a stochastic model for individual risk. *Biometrics* 45:1126–1138; 1989.
- Samartzis D, Nishi N, Cologne J, Funamoto S, Hayashi M, Kodama K, Miles EF, Suyama A, Soda M, Kasagi F. Ionizing radiation exposure and the development of soft-tissue sarcomas in atomic-bomb survivors. *JBJS* 95:222–229; 2013.
- Spieler EA. (Re) Assessing the grand bargain: compensation for work injuries in the United States, 1900–2017. *Rutgers UL Rev* 69:891; 2016.
- Stram DO, Sokolnikov M, Napier BA, Vostrotin VV, Efimov A, Preston DL. Lung cancer in the Mayak workers cohort: risk estimation and uncertainty analysis. *Radiat Res* 195:334–346; 2021.
- Thierry-Chef I, Marshall M, Fix JJ, Bermann F, Gilbert ES, Hacker C, Heinmiller B, Murray W, Pearce MS, Utterback D, Bernar K, Deboodt P, Eklöf M, Griciene B, Holan K, Hyvonen H, Kerekes A, Lee MC, Moser M, Pernicka F, Cardis E. The 15-country collaborative study of cancer risk among radiation workers in the nuclear industry: study of errors in dosimetry. *Radiat Res* 167:380–395; 2007.
- Thierry-Chef I, Richardson DB, Daniels RD, Gillies M, Hamra GB, Haylock R, Kesminiene A, Laurier D, Leuraud K, Moissonnier M, O'Hagan J, Schubauer-Berigan MK, Cardis E. Dose estimation for a study of nuclear workers in France, the United Kingdom and the United States of America: methods for the International Nuclear Workers Study (INWORKS). *Radiat Res* 183:632–642; 2015.
- Tirmarche M, Apostoaei I, Blanchardon E, Ellis ED, Gilbert E, Harrison JD, Laurier D, Marsh JW, Sokolnikov M, Wakeford R, Zhivin S. ICRP Publication 150: cancer risks from plutonium and uranium exposure. *Ann ICRP* 50:1–143; 2021.
- Tokarskaya ZB, Okladnikova ND, Belyaeva ZD, Drozhko EG. Multifactorial analysis of lung cancer dose-response relationships for workers at the Mayak nuclear enterprise. *Health Phys* 73:899–905; 1997.
- Tokarskaya ZB, Scott BR, Zhuntova GV, Okladnikova ND, Belyaeva ZD, Khokhryakov VF, Schollnberger H, Vasilenko EK. Interaction of radiation and smoking in lung cancer induction among workers at the Mayak nuclear enterprise. *Health Phys* 83:833–846; 2002.
- Tokarskaya ZB, Zhuntova GV, Scott BR, Khokhryakov VF, Belyaeva ZD, Vasilenko EK, Sychikov VA. Influence of alpha and gamma radiations and non-radiation risk factors on the incidence of malignant liver tumors among Mayak PA workers. *Health Phys* 91:296–310; 2006.
- UNSCEAR. Sources, effects and risks of ionizing radiation. Report to the General Assembly, with scientific annexes. New York, NY: United Nations Scientific Committee on the Effects of Atomic Radiation; 2012.
- USAF. Project Crested Ice. USAF Nuclear Safety J special edition (AFRP 122-1) 65 Pt 2 1:1–97; 1970. Available at https://archive.org/details/DTIC_AD0701493. Accessed 10 September 2021.
- US DHHS. Report of the NCI-CDC working group to revise the 1985 NIH radioepidemiological Tables. Washington, DC: US Department of Health and Human Services, US National Cancer Institute; 2003. Available at <https://www.cdc.gov/niosh/docket/archive/pdfs/NIOSH-209/0209-010103-land.pdf>. Accessed 19 August 2021.
- USC. Veterans' benefits, claimant responsibility; benefit of the doubt 38 United States Code § 5107. 2011.
- van Buuren S, Groothuis-Oudshoorn K. Multivariate imputation by chained equations in R. *J Stat Software* 45:1–67; 2011.
- VanderWeele TJ. Optimal approximate conversions of odds ratios and hazard ratios to risk ratios. *Biometrics* 76:746–752; 2020.
- Vasilenko E, Gorelov M, Smetanin M, Knyazev V, Teplyakov I, Alexandrova O, Miller S, Eckerman K, Fix J, Scherpelz R. Mayak worker study project 2.4. Volume II: dose assignment methodology used to calculate annual organ doses to Mayak workers from external radiation. *Health Phys* 93:1–100; 2007.
- Vostrotin VV, Napier BA, Zhdanov AV, Miller SC, Sokolova AB, Bull RK, Suslova KG, Efimov AV, Smith MA, Vvedensky VE. The Mayak worker dosimetry system (MWDS-2016): internal dosimetry results and comparison with MWDS-2013. *Radiat Protect Dosim* 184:201–210; 2018.
- White IR, Carlin JB. Bias and efficiency of multiple imputation compared with complete-case analysis for missing covariate values. *Stat Med* 29:2920–2931; 2010.

APPENDIX

Adjustments needed to IREP distributions to match published worker results

IREP documentation specifies how to build distributions that can be used to (1) generate IREP-based predictions for worker RBEs to compare with results given in studies of plutonium workers exposed to low dose or low dose-rate external, low-LET radiation, and (2) to transfer ERR Gy⁻¹ values estimated in the UK and the Russian Federation to values appropriate for the US population assumed in IREP output.

Predictions of risk for inhaled plutonium in IREP start with the distribution for the so-called low radiation effectiveness factor, REF_L, which is log-normally distributed with central value and 95% confidence limits equal to 18 (3.4, 100). REF_L was determined from alpha data, primarily animal and cell data, relative to low-LET radiation with photon energies above 250 KeV, delivered at low doses and/or low dose rates (Kocher et al. 2005). It and its uncertainty range are assumed to be independent of dose within the low dose and/or low dose-rate regions for which it is defined.

Adjustments to REF_L that produce an RBE distribution for workers that can be compared to results from plutonium worker studies

Had the plutonium worker external exposures all been to photons with energies above 250 KeV, REF_L would be a direct prediction of RBEs found in worker studies. However, this was not the case. IREP documentation provides modifying factors for REF_L, as described in Table A1, which are used for (1) photons with energies between 30 KeV and 250 KeV and (2) photons with energies below 30 KeV (Kocher et al. 2005). Because the distribution of photon energies at Mayak and Sellafield were not available, calculations were made for each of the photon energy ranges and compared separately to published worker RBE values (Table 3). To build final RBE distributions, each draw from the modifying distribution in Table A1 was multiplied by a draw from the REF_L distribution obtained from the “rlnorm” function in the R-statistical programming language (R_Core_Team

Table A1. Probability distributions of REFs for ionizing photons with different energy groupings as developed for IREP in Kocher et al. (2005, Table 11).

| Photon energy range | Distribution |
|---------------------|---|
| > 250 KeV | Single-valued at 1.0 (assumed reference radiation) |
| 30 – 250 KeV | Hybrid distribution with—75% weight to lognormal distribution with 95% confidence interval between 1.0 and 5.0; 25% weight to value 1.0; |
| < 30 KeV | Product of two distributions—(1) hybrid distribution for 30 – 250 keV photons; and (2) triangular distribution with minimum of 1.0, mode of 1.3, and maximum of 1.6 |

2020). The standard percentile values, including the 99th, were then extracted from the distributions.

Adjustments needed to transfer risks to US populations estimated for worker cohorts in the Russian Federation and the UK

Before comparing with IREP-based predictions, ERR Gy⁻¹ values estimated in studies of workers in the Russian Federation or the UK should be transformed to values appropriate for a US population. To estimate the size of the effect, risk transfers have been estimated here, when possible, using IREP methodology but only reported in the text and in Table footnotes. Risk transfer in IREP requires data on cancer mortality. Lung cancer mortality rates for males and females in the US, UK, and Russian Federation were taken from published data for the year 2005 (Islami et al. 2015). Transfer calculations were made separately for each sex. Transfers were not attempted by smoking status because no corresponding historical mortality data by smoking status were found for the Russian Federation.

There are two IREP approaches for lung cancer risk transfer, with the one producing the largest 99th-percentile EER Sv⁻¹ accepted (Kocher et al. 2008). The choice of year to use for lung cancer mortality rates is not specified in IREP documentation; thus, a range of years was investigated because the mortality data most relevant to a claimant would likely be from a year prior to the date of cancer diagnosis. In addition to 2005 for the main analysis, data from 1960 and 1995 were also considered, as shown in Supplemental Digital Content Tables S-3 and S-8, <http://links.lww.com/HP/A217>, with the following results. Using the 2005 lung cancer mortality data, the percentage changes in ERR Gy⁻¹ were approximately +20% for Mayak males and approximately -20% for Mayak females. For Sellafield male workers, the percentage changes were ~ -9%. As for the 1995 choice for lung cancer mortality rates, the percentage changes rose to +27% for Mayak males, with no change for Sellafield males. When 1960 mortality rates were used, there was a large percentage increase for Sellafield males (85%–89%) and a 0-to-12% decrease for Mayak males. Nevertheless, such percentage changes found for 1960 mortality data would only amplify the differences between the cohorts noted in this paper and would not change any conclusions.

Calculation of RBEs involve ratios of risk. Because transference of risk from one population to another will apply to both numerator and denominator, risk transfer effects will cancel for RBEs to first order.

Pooling of IREP predictions across cohort subcategories

IREP makes predictions about individual risk, conditional on such factors as sex and smoking status at time of diagnosis. To use these conditional risk estimates to simulate risks predicted for worker cohorts whose members had

varying risk factors, it was necessary to perform a weighted sum over IREP distributions (marginalize them). This required having prevalence numbers for the risk factors in the cohort, which were obtained from the literature, and full knowledge of the individual IREP distributions. IREP-based ERR Gy⁻¹ distributions, obtained from inverted IREP assigned-share output, were assumed to have a lognormal underlying distribution. The 95th- and 50%-credibility percentiles were used to obtain a log standard deviation for ERRs. This was accomplished by dividing the logarithmic difference by the associated z-score critical value of 1.65. (1.65 times a standard deviation added to the 50% value reaches the 95th-percentile value for a normal distribution.) Note that the 95th-percentile is the upper limit for a 90th percent confidence interval.

Using the 50% value with the derived log standard deviation, draws from the lognormal distribution were made using the “rlnorm” function in R-code. For pooling, the number of replications was chosen to be 10 times the number of subjects in a particular worker cohort, male or female, smokers or nonsmokers, with percentages described in Supplemental Digital Content Table S-2, <http://links.lww.com/HP/A217>. With these individual distributions in hand, now weighted by the risk category percentages, pooling of risks for cohort subcategories could be made and the resulting ERR Gy⁻¹ values for the standard credibility levels extracted from the pooled distributions.

IREP documentation discusses multiple uses of the lognormal distribution. To check that the lognormal approximation was a reasonable choice for lung cancer ERR Gy⁻¹ distributions, the 99th-percentile values were extracted from the simulated lognormal distributions as discussed above and compared with IREP-based 99th-percentile output values for males and females, each for three different smoking histories. The approximated values tended to be slightly higher than the IREP output, with an increase ranging from approximately 0% to 11%. Thus, the IREP distributions have

a slightly narrower upper tail than expected for a pure lognormal distribution, but the difference in the 99th-percentile predictions, which determines compensation decisions in IREP, is minor compared to the overall uncertainties.

Because results from epidemiological studies of plutonium exposure did not use the same smoking breakdown as did IREP, it was necessary to combine the IREP predictions to best match the worker categories used in the studies that presented results for lung cancer diagnosis at age 60 (Supplemental Digital Content Table S-2, <http://links.lww.com/HP/A217>). When smoking assignment data were not available, percentages were assigned to the IREP smoking category that would increase pooled results. The impact of the alternate choice is presented along with the main results in Table footnotes.

To impute the 99th percentile for published study data, a normal distribution was assumed rather than a lognormal distribution, given that the output was from a regression analysis and generally with tight confidence limits. If results were computed assuming an underlying lognormal distribution, the 99th-study percentiles were also similar, at most higher by 20%. The standard deviation estimate was obtained from the difference of a study's upper and 50% confidence limits divided by the appropriate statistical critical value, either 1.65 for 90% confidence intervals (CI) or 1.96 for 95% CI, depending on the study's choice for confidence limits. These values were used to define 10,000 replicates of a normal distribution using the R-code routine (rnorm) from which the 99th percentile was then extracted. These exploratory 99th-percentile values are placed alongside study confidence intervals in Table columns in the results section. The values are not necessarily the same as the 99th percentiles that would be found in actual study data, which might not be pure normal.

