

**Mischaracterization of the epidemiological data on cancer from plutonium inhalation in a report prepared in July 2020 for the US Veterans Administration relating to compensation claims for veterans of the Palomares, Spain cleanup in 1966**

Jan Beyea,<sup>1</sup> Draft, 21 September 2020, corrected 23 September.

1. Overview:

Epidemiological data plays a key role in compensation decisions for US veterans exposed to ionizing radiation as part of their military service (Kocher et al. 2008). My coauthor, Frank N. von Hippel and I have made use of epidemiological data from studies of Russian Plutonium workers in a paper published in *Health Physics* concerning the US Veterans Administration's denial of claims of veterans of the Air Force plutonium cleanup after the 1966 crash of a nuclear-armed US B-52 in Spain (Beyea and von Hippel 2019). This paper has been discussed as part of submissions to the Court of Appeals for Veterans Claims (DVA 2020). See also: (YLC Online).

In July of 2020, the Air Force submitted to the Court of Appeals (DVA 2020) a report by Steven E. Rademacher (hereinafter, "SER") that responds, in part, to our 2019 paper. Subsequently, von Hippel responded to some of the issues raised in the SER report (Von Hippel 2020), finding that newly released information in the report actually strengthens the arguments in the *Health Physics* paper. In this research note,<sup>2</sup> I focus on the analysis of Mayak and Sellafield epidemiology studies contained in the SER report.

Of most note, the SER report claims evidence of a threshold below which plutonium doses do not cause cancer. In fact, this "evidence" is based primarily on an early (1997) analysis of Russian data with limited statistics that has been superseded by numerous subsequent analyses by an international consortium with access to more data and better dose estimates. The authors of these later papers have made no threshold claim. The Air Force report also includes its own *de Novo* analysis to the effect that a dose threshold lies unnoticed in some of these later papers, but this analysis is not statistically based. Moreover, the visual method relied upon can be turned around to suggest greater risks from Plutonium than now assumed in making compensation decisions.

**2. The SER Report**

One of the important sources of epidemiological data for assessing risks from inhalation of plutonium comes from the workers at the Mayak plutonium production site in Russia's Urals (Table 1). SER discusses purported findings from studies of Mayak workers, concluding that the Mayak studies "provide further conservative influence on compensation decisions." To reach this conclusion, which I find incorrect, SER does the following:

- 1) He emphasizes the early 1997 work that reported a nonlinear and threshold-like dose response.

---

<sup>1</sup> Senior Scientist, Emeritus, Consulting in the Public Interest, 53 Clinton Street, Lambertville, NJ, 08530, [www.cipi.com](http://www.cipi.com).

<sup>2</sup> This research note has been prepared for eventual journal submission. No attorney, nor law student, has seen or discussed this document prior to its release on September 21, 2020.

- 2) He relies on his own (largely unexplained) interpretation of later work (Table 1) to claim that the later studies agree with the 1997 work, when such a conclusion does not match what the authors actually wrote.
- 3) Thus, he does not fully and correctly describe the relevant statements in these later reports.
- 4) He omits the parts of these later studies concerning “relative biological effectiveness” factors (RBE) that are fully supportive of the cancer risk coefficients used in compensation decisions through the IREP<sup>3</sup> methodology (Table 2).
- 5) He focuses, without statistical analysis, on low data points at the low end of the lung cancer dose response curve in some studies while not mentioning the cases when the anomaly goes the other way and might suggest supralinearity. Ironically, applying this visual inspection to pooled Mayak and Sellafield data, an attorney for a veteran claimant could argue that the IREP program understates the uncertainty in the RBE, leading to lower compensation dose cutoffs.

As a result of these limitations in analysis, SER’s claim of over-conservatism is improperly sourced and ends up being misleading. SER omits mention of the agreement with the radiobiological effectiveness factor for lung cancer (RBE), which is what the compensation computer program, IREP, uses to assess plutonium risks (Table 2). The RBE relates Plutonium cancer risk (ERR/Gy) to the better-known ERR/Gy from gamma rays and x-rays. To estimate the risk for an exposed veteran, the veteran’s plutonium dose, measured in Gy, is multiplied by the RBE value, which converts the dose to Sieverts (Sv). The cancer risk is then taken from the IREP gamma ray risk formula for that Sv value. Uncertainties, which play a major role in compensation assessment, are then estimated, including the uncertainty in the RBE. This two-step approach is the heart of the decision process for determining compensation from Plutonium exposure.

As indicated in Table 2, the RBE determined in the various Mayak studies varies from 10 to 45. The corresponding range in IREP is 3.4 to 100 at 95% credibility values, with a central value of 18. The Mayak range falls in the middle of the IREP range, which indicates good agreement.

As for nonlinearities in the dose response curves reported by the modern studies, the findings are summarized in Table 1. Also, listed is the earlier report in 1997 by Tokarskaya et al. of a nonlinear and threshold-like dose response. This is the report whose reported findings SER emphasizes.

When judging differences over time in the historical studies, it should be noted that the modern studies were carried out using the dosimetry developed by an International Consortium to support and extend the work of Russian analysts (Preston et al. 2017). The modern studies also have more follow-up cases, which gives them more statistical power.

---

<sup>3</sup> Interactive Radioepidemiological Program, 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..

Table 1. Dose-response findings on lung, liver and bone, as reported in the Mayak studies cited, and two not cited, by SER <sup>a)</sup>									
Study	Cancer site	Findings	Dosimetry Version	Study Type	Adjust for smoking	Quadratic term	Threshold analysis	RBE given	Comment
(Tokarskaya et al. 1997)	Lung	Threshold at 0.8 Gy, Quadratic or Linear + Quadratic fit	Pre-1997	Case Control	Yes	P = 0.0001	Yes		
(Tokarskaya et al. 2002)	Lung	Not adjusting for smoking could overestimate risks; possibly obscure threshold	2000	Case Control	Yes	No	No		
(Tokarskaya et al. 2006)	<b>Liver</b>	Increased odds ratio for 2-16 Gy category relative to 0-2 Gy category	2000	Case Control	Yes	No	“ “		Not mentioned by SER. Limited data
(Labutina et al. 2013)	Lung	Linear fit to full dose range	2008	Co-hort	Yes	P = 0.25	“ “		
	<b>Liver</b>	Linear fit for dose range below 2 Gy	“ “	Co-hort	Yes		“ “		Linear-Quadratic for full dose range (6 Gy)
	<b>Bone</b>	No definite dose response (only 4 cases with Pu exposure)	“ “	Co-hort	Yes		“ “		
(Gilbert et al. 2013)	Lung (mortality)	Linear fit for full dose range	“ “	Co-hort	Yes	P = 0.5	“ “	Yes	Not mentioned by SER.
(Gillies et al. 2017)	Lung	Linear fit for full dose range	2013	Co-hort	No	P > 0.5	“ “	Yes	Dismissed by SER because no smoking adjustment.
a) There are other Mayak papers not listed here, because they have been superseded by the publications listed here.									

Table 2. Radiobiological effectiveness factor (RBE) for lung cancer estimated from Mayak worker data compared to value used in IREP program for compensation decisions <sup>a)</sup>					
	Relative to risks from protracted Mayak external gamma dose		Relative to A-bomb risks from prompt external gamma		Dosimetry version
	Incidence	mortality	Incidence	Mortality	
(Gilbert et al. 2013)		45 (21-240) <sup>b,c)</sup>		20 (exposed at ages 15-60)	2008
(Labutina et al. 2013)		Not reported. Value would be 35, if take ratio of ERR/Gy's from Table 4 in paper	Not reported. Likely similar to Gilbert et al. <sup>d)</sup>		2008
(Preston et al. 2017)				28 (exposed at ages 20-39)	2008
(Gillies et al. 2017)		10-25 (Pooled estimate. Not smoking adjusted) <sup>e)</sup>			2013 for Mayak component)
IREP	18 (3.4 – 100) <sup>b,f)</sup>		18 (3.4 – 100)] <sup>b)/</sup> DDREF <sup>g)</sup>		

a) IREP = Interactive Radioepidemiological Program (Kocher et al. 2008).  
b) 95% confidence or credibility intervals  
c) Based on full dataset  
d) The ERR/GY for males are similar for Pu in both the Labutina et al. and the Gilbert et al. studies. (7.1/Gy and 7.4/Gy, respectively.) The A-bomb estimates are also similar for lung cancer incidence and mortality for the relevant age ranges (Gilbert et al. 2013).  
e) Pooled Sellafield and Mayak data. Smoking data for UK Sellafield works was not usable. Because Preston et al. only cite the Gilbert et al. mortality study, not the incidence study by Labutina et al., their RBE has been assigned to the mortality column.  
f) Note that this value would be called a Radiation Effectiveness Factor i(REF) n IREP documentation (Kocher et al. 2008). The value quoted is the so-called “Low” REF (Kocher et al. 2008 at Table 3).  
g) DDREF = dose and dose rate effectiveness factor for comparison gamma risk, which is a variable with its own uncertainty distribution in IREP, with a mean value around 2. Multiplication by the DDREF converts, apparently, the so-called “Low” IREP REF to the underlying RBE (Kocher et al. 2008). The corresponding credibility intervals for this entry would be a combination of the uncertainty in the REF and the DDREF. Not considered here is the transfer of risks from the Japanese survivor population to a US populations, which would modify the entry in the Table further.



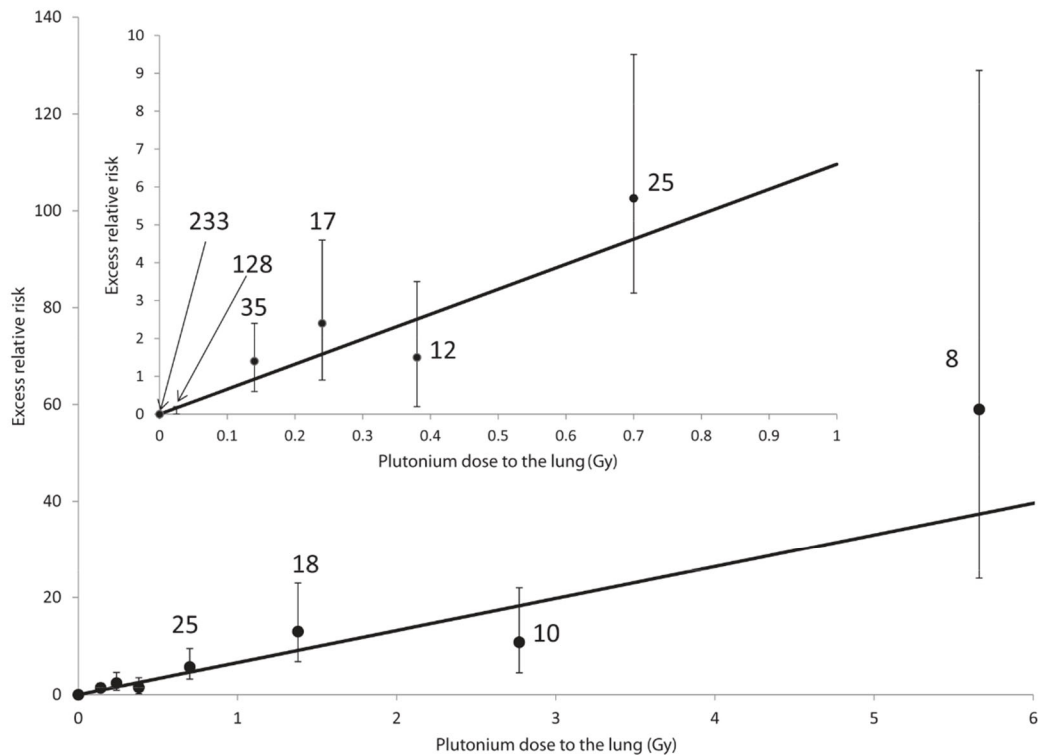
### 3. Statements from the modern Mayak studies

With one exception, nonlinearities were not found in the modern Mayak studies (Table 1). In this section, the actual statements in the papers are presented. For example, a 2013 lung cancer mortality study, not cited by SER, described its findings as linear:

“The dose-response relationship is well described by a linear function, and a linear-quadratic function did not significantly improve the fit ( $P > 0.5$ ). If the dose-response relationship was expressed as a power of plutonium dose (dplua), the estimated power ( $\alpha$ ) was 1.02 (95% CI: 0.84–1.23).

“Highly statistically significant dose-response relationships were found even when analyses were restricted to plutonium doses less than 0.2 Gy [200 mGy]. Estimates of the ERR per Gy were similar regardless of the dose range restriction.” [Conversion of 0.2 Gy to 200 mGy added.]

Figure 1 reprinted from that study shows the linear fits to lung cancer graphically.



**FIG. 1.** Excess relative risk (with 95% CI) of lung cancer and number of lung cancer deaths by categories of plutonium dose to the lung. Shown for males at age 60. Estimated linear function also shown.

A 2013 cancer incidence study, cited by SER as showing a threshold like dose response curve with linearity only above 0.2 Gy, described the results differently. No caveats were included about the results not holding below 0.2 Gy: In fact, Labutina found a linear relation using all the data, including the counts at and below 0.2 Gy (200 mGy).

“There was clear evidence for the linear association between internal plutonium dose and the risk of lung cancer.”

“The lung cancer risks were best described by a linear dose response relationship for internal dose to the lung. Adding a quadratic term to the model did not significantly improve the fit ( $p = 0.25$ ).”

It is true that, at the lowest categories, the point values for relative risk are low, albeit with large error bars. Fluctuations are inevitable in study counts and can overwhelm the small predicted excesses at the lowest dose categories. There are statistical ways to assess the importance of fluctuations, such as regression analysis. When regression slopes and their errors no longer exclude the null, as happened in the Gilbert study for the dose range 0 to 0.1 Gy, little can be said about the meaning of such data points.

Visual inspection of low dose cancer counts appears to be the method used in the SER report to assess the study by Labutina et al. A major 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 fluctuating data would be judged to indicate nonlinearity? I have worked on calculating the expected frequency of false positives for nonlinear dose response that occur using visual methods of inference and, along with colleague, George R. Hoffman, will be publishing sample results; however, they are not available for this draft.

Visual inference of nonlinearity is inherently subjective and vulnerable to confirmation bias. Once policy is based on visual inferences from data points with large fluctuations, the connection to science can be lost. Moreover, critics can easily find examples, where the fluctuations go the other way. I do this below in Section 4 by showing graphs from a 2017 study that pooled Mayak worker data with UK Sellafield data. A major goal of that study was to explore how risks might translate from Mayak workers to other groups. The authors reported finding linearity of dose response.

“There was clear evidence of a linear association between cumulative internal plutonium lung dose and risk of both lung cancer mortality and incidence in the pooled cohort.”

SER dismisses this study, because the data was not adjusted for smoking. It was not possible to adjust for smoking because no usable smoking data was available for Sellafield workers. The authors acknowledged this limitation, but argued that adjusting for smoking in Mayak studies did not produce a major change in risk coefficients.

Having not cited the lung cancer mortality study, and dismissing the pooled study, SER was left with only the Labutina study and his novel characterization of it as showing a threshold. Only then could he claim, as he did, that the modern cancer studies supported the 1997 study that reported a threshold.

In only one case in the modern studies was a quadratic term needed. That was for liver cancer, and then only for the full dose range, which went out to doses over 6 Gy (120 Sv). The nonlinearity in the liver cancer dose response fit “disappeared” when “the internal doses were restricted to less than 2 Gy.” Doses above 2 Gy are certainly not of any relevance to claims that Palomares veterans might have made. For instance, 0.02 Gy (0.4 Sv) was the dose recommended by the Air Force J for lung

compensation decisions for non-High 26 Palomares cleanup veterans (Beyea and von Hippel 2019 at Table 1)

For the range below 2 Gy, the authors reported a dose response that was linear. The reported slope is reasonably consistent, I checked, with risks determined from gamma ray exposures in the A-bomb epidemiological data as well as in the Mayak gamma epidemiological data, when the standard RBE of 20 is used. Note that none of the modern studies were able to fit bone cancer to a dose response curve, which is not surprising given the paucity of cases.

#### 4. Threshold

The most startling claim by SER is that a plutonium dose threshold was found for lung cancer in the later studies. No such mention is found in the later papers; instead, SER appears to have reached his conclusion from inspection of data in a 2013 Table, particularly the lowest entries in the Table. It is not possible to be sure, because the claim is stated indirectly. First, he states that Labutina et al. found a linear relationship above 200 mGy for lung cancer and then implies that this was a similar finding to the 1997 paper's finding of a dose response with "a non-linear threshold character." In fact, Labutina found a linear relation using all the data, including the counts at and below 200 mGy.

Similarly, his claim of a threshold being found for bone cancer appears to be based on his inspection of data in another 2013 Table. This is fine, but readers should be made aware of the distinction, including in the executive summary, between a reported finding and a personal call. And, then, it seems to me that the rationale should be made explicit, so that a technical reader doesn't have to guess at the logic. Also, all readers should be made aware of data in Tables or Figures that suggests a supralinear response, something that SER did not do.

The region where SER apparently finds a threshold-like dose response is from 0 to 0.2 Gy, which corresponds in effective dose to 0 to 4 Sv. This is a range that extends far beyond the dose range at which compensation decisions would be contested.

If SER is correct, then the RBE must drop precipitously at 0.2 Gy, because at the same Sievert (4 Sv) the A-bomb response is strong. A threshold in the RBE is a novel claim that I have previously not heard expressed. I do not think a novel claim such as this can be used to suggest that there is more conservatism in the compensation decisions than is already built into the system through the use of the 99<sup>th</sup> credibility risk and use of the maximum dose in a range. The underlying logic is that, at the compensation cutoff point, it should be very unlikely that the veteran's exposure had contributed much to the claimed illness. This is a very different standard than proving a tort claim in a US Court of law. There is certainly much to criticize in the compensation analysis, and I have done so in the past, for instance, in collaboration with biostatistician, Sander Greenland, with regards to the improper labeling of the output of the IREP program as a probability (of causation) (Beyea and Greenland 1999). Because probability of causation requires the assumption of an underlying biologic model, the outcome of IREP is best called an assigned share. Nevertheless, the IREP program as used in the NIOSH online calculator provides in principle a framework that merges science and policy in a workable manner that removes



the need for toxic tort litigation in federal and state courts, by giving much benefit of the doubt to the veteran.

#### **5. Sellafield data**

A dose range more relevant to the Palomares claimants than the Mayak dose range can be found in the pooled study of UK Sellafield and Mayak workers. The plutonium doses to the Sellafield cohort were about an order of magnitude lower than the doses to Mayak workers. The Sellafield ERR/Gy values range from  $\sim 1.5$  to 3 times the ERR/Gy of the Mayak workers in the Gillies et al. study, depending on the solubility case used in assessing Sellafield Plutonium doses. Hence, the Gillies et al. study provides graphs with datapoints in a range to 0.07 Gy, as shown in the bottom panes of Fig. 2.

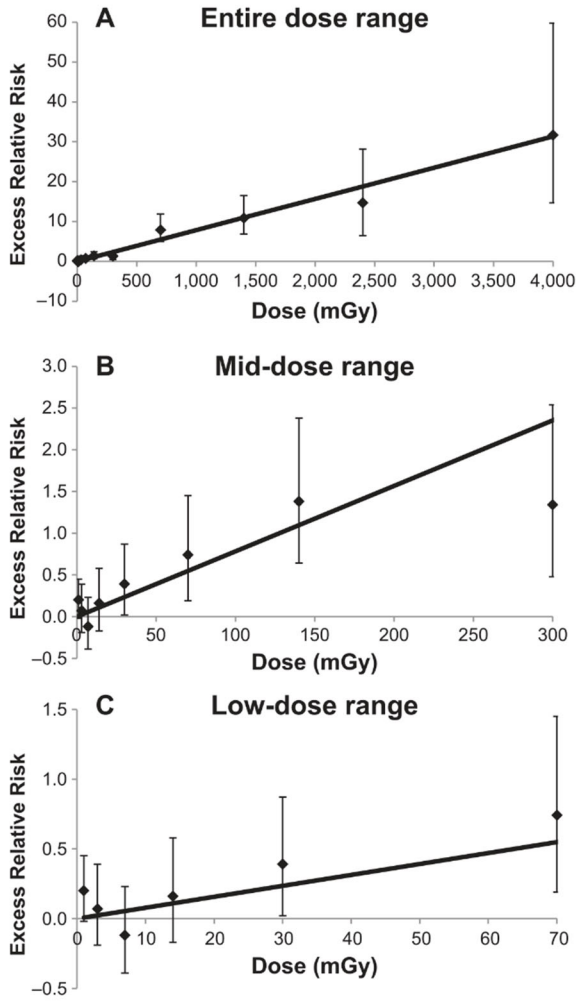


FIG. 3. Pooled ERR estimates by plutonium dose category (using the *ss\_fast* solubility assumption), linear ERR/Gy estimate reference line superimposed (estimated using the entire dose range for all three panels). Panels A–C: Entire, mid- and low-dose range, respectively.

Fig. 2. From Gillies, 2017

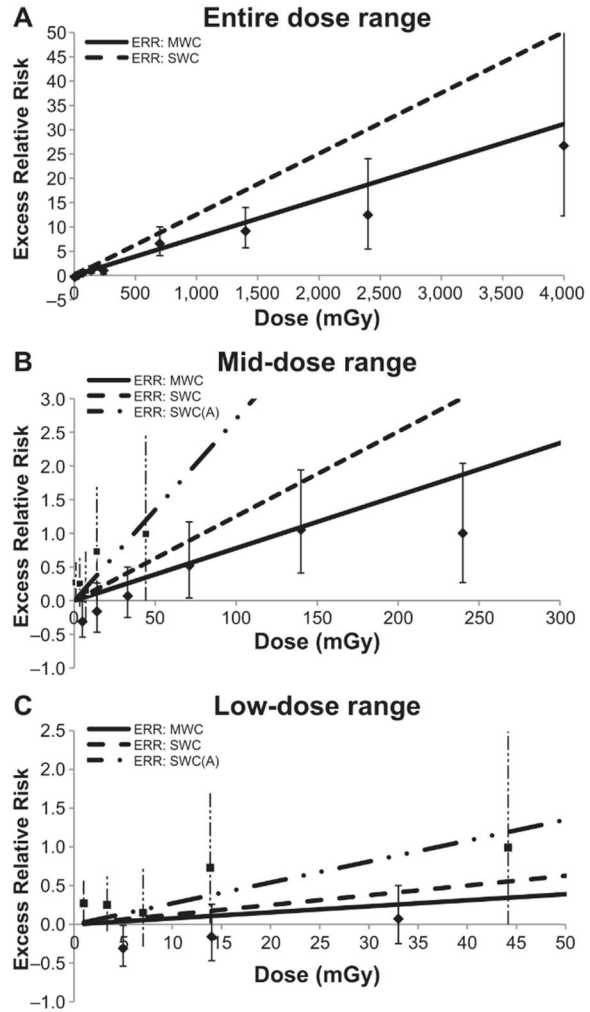


FIG. 4. Cohort-specific ERR estimates by plutonium dose category (using the *ss\_fast* solubility assumption). SWC and MWC indicated by squares and diamonds, respectively. The linear ERR/Gy estimate reference line is superimposed (estimated using the entire dose range for all three panels). Panels A–C: Entire, mid- and low-dose range, respectively. The line labeled ERR-SWC(A) represents the linear ERR/Gy estimate for the SWC when the model includes a parameter for age effect modification, i.e., the line represents risk at attained age 60.

Those datapoints at  $\sim 0.001$  and  $0.002$  Gy in the bottom panes show these first two central points to be above the linear line using the Mayak fit. These two points with doses in Sieverts of 20 and 40 mGy fall below the region in the A-bomb dose response curve where the shape of the dose response curve is uncertain. A supralinear response in Sellafield workers at these low doses would not be inconsistent with what is clearly known from the study of A-bomb survivors and thus not necessarily requiring a shift in the RBE value.

If one is going to use visual assessment of dose response curves to justify compensation decisions (which I am not recommending without statistical analysis), then it should be balanced by mentioning this apparent supralinear response at the lowest doses measured. Now, the pooled study did not adjust

for smoking, which is a major limitation, although not necessarily one that would make a difference according to the authors, because, as previously stated, when smoking has been adjusted in the Mayak studies the risk coefficients did not change much. Nevertheless, the identification of a visual, supralinear excess at the very lowest doses, those relevant to Palomares compensation decisions, should give one pause. Either adjustment for smoking flips risk coefficients from very small to very large, or just maybe, all of these deviations from linearity are fluctuations and it is improper to draw conclusions about best estimates of dose response based on visual inspection of a few points at the lowest dose end of a plot. Although it is improper to draw conclusions about central estimates based on visual inspection, this method is often used out of necessity to assess uncertainty ranges. And uncertainty in the high-risk direction can affect the dose cutoff for a veteran's compensation. The uncertainties in the low-risk direction that SER sees would not affect compensation decisions, if the central value is maintained. Ironically, applying this visual inspection method to pooled Mayak and Sellafield data, an attorney for a veteran claimant could argue that the IREP program understates the uncertainty in the RBE, leading to lower compensation dose cutoffs.

## 5. Summary:

There is no statistically reliable evidence that would allow SER's conclusion about extra conservatism in IREP to stand. Visual assessment of data points is not a substitute for statistical analysis. It is subjective, subject to confirmation bias, and the rate of false positives from such an assessment is not known. The later Mayak studies fully support the use of RBE factors to ascertain plutonium risks for exposed veterans. Uncertainty in RBE's is incorporated in IREP. The higher-risk Sellafield results given in the pooled study, although unadjusted for smoking, indicate that assessment of this cohort should be pursued and watched for IREP implications. In the meantime, it would make sense, using statistical methods, to explore possible modifications to the RBE uncertainty distribution based on the Sellafield data in hand, which could lead to lower dose cutoffs for compensation.

## 6. Acknowledgements:

I thank Frank N. von Hippel for comments. I thank F. Owen Hoffman and David Kocher for reviewing Table 2 for compatibility with IREP.

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 (Available at: [https://journals.lww.com/health-physics/Fulltext/2019/12000/History\\_of\\_Dose,\\_Risk,\\_and\\_Compensation.13.aspx](https://journals.lww.com/health-physics/Fulltext/2019/12000/History_of_Dose,_Risk,_and_Compensation.13.aspx) Accessed 22 Sept. 2020). *Health Phys* 117: 625-636; 2019.

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. WILKIE, Secretary of Veterans Affairs, Appellee.; 2020.

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; 2017.
- 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.
- 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.
- 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 Prot Dosim* 173: 26-31; 2017.
- 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: 833-46; 2002.
- Tokarskaya ZB, Zhuntova GV, Scott BR, Khokhryakov VF, Belyaeva ZD, Vasilenko EK, Syrchikov 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.
- Von Hippel FN. Comments on the new Air Force report, Topical Issues for Assessment of Dose to Palomares Accident Recovery Workers (1966) by Steven E. Rademacher (SER), July 2020, Draft, 27 August (Available at: <https://sgs.princeton.edu/pdf/vonhippel-2020.pdf>. Accessed 21 Sept. 2020). Program on Science and Global Security, Princeton University; 2020.
- YLC. Palomares FOIA Litigation (Available at: <https://law.yale.edu/studying-law-yale/clinical-and-experiential-learning/our-clinics/veterans-legal-services-clinic/palomares-foia-litigation>; <https://law.yale.edu/yls-today/news/veteran-sickened-plutonium-brings-nationwide-class-action>. Accessed 7 June 2018.). New Haven: Yale Law Clinic; Online.