Effect of Alternative Percentile Definition on PC Values

White Paper

National Institute for Occupational Safety and Health, NIOSH

October 2021

Daniel O. Stancescu
Division of Compensation Analysis and Support, NIOSH

A. Iulian Apostoaei, Brian A. Thomas
Oak Ridge Center for Risk Analysis, Inc. (ORRISK, Inc.)

Reviewed by Timothy D. Taulbee
Division of Compensation Analysis and Support, NIOSH

ABSTRACT

The 99th percentile of the **p**robability of cancer **c**ausation (PC) for workers exposed to ionizing radiation is used to adjudicate claims for compensation filed under the Energy Employees Occupational Illness Compensation Program Act (EEOICPA). Such percentile is obtained from a probability distribution function (pdf) described by 2,000 or more samples produced using Monte-Carlo iterations to account for uncertainties in radiation dose and risk parameters representing the exposure situation. A previous report evaluated ten methods for estimating sample percentiles and determined that the method used by default in the SAS software has the smallest relative bias and smallest relative root mean squared error at all the sample sizes, when estimating the 99th percentile, while the method currently used in NIOSH-IREP tends to underestimate the 99th percentile especially for sample sizes lower than 20,000 values.

In this report, an analysis was conducted to investigate the effect on PC values of using the SAS percentile method, instead of the method implemented in the NIOSH-IREP software, for real claims. Significant differences were observed between the 99th percentiles of PC obtained using the alternative SAS method and those obtained using the NIOSH-IREP method, with the alternative SAS method always producing higher values. The analysis was carried out using several categories of previously analyzed cases selected from NOCTS. For 45 cases with 99th PC values between 44.00% and 44.99%, switching from the NIOSH-IREP method to the alternative method produced an average increase in the 99th PC values of 0.27 (maximum 1.46) at 2,000 Monte-Carlo iterations, an average increase of 0.06 (maximum 0.27) at 10,000 iterations, and an average increase of 0.02, (maximum 0.17) at 20,000 iterations. For 72 tested cases with 99th PC values between 45.00% and 51.99%, switching from the NIOSH-IREP method to the alternative method generated an average increase in the 99th PC values of 0.04 (maximum 0.11) at 10,000 iterations, an average increase of 0.02 (maximum 0.05) at 20,000 iterations, and an average increase of 0.02 (maximum 0.04) at 30,000 iterations.

To minimize the underestimation effect of the NIOSH-IREP percentile method and to generally improve the precision of the 99th PC values, either the alternative SAS percentile should be implemented in future NIOSH-IREP versions, or, preferably, the number of Monte-Carlo iterations used should be increased. Our analysis indicates that the number of iterations used in IREP should be increased at a minimum of 10,000, but preferably to 20,000 or more.

The number of iterations used in IREP Enterprise Edition (IREP-EE) should be respectively increased to a minimum of 20,000, but preferably to 30,000 or more. For cases with 99th PC values between 49.50% and 49.99% evaluated in IREP-EE, a much larger number of iterations is needed to reduce variability and increase the accuracy of the 99th PC value. This can be achieved not only by increasing the number of iterations per run but also by increasing the number of runs in IREP-EE from 30 to 100 or more runs.

TABLE OF CONTENTS

1.0 INTRODUCTION

This document is a follow-up to the "Comparison of Several Percentile Definitions" white paper (Stancescu et al., 2019) and is looking at the effect on Probability of Causation (PC) values of an alternative percentile definition, instead of the default percentile definition implemented in the NIOSH-IREP software. The document includes an analysis of the PC variability at different number of iterations, and an analysis of methods to select cases to be run on the IREP Enterprise Edition based on the PC values obtained in IREP.

2.0 DEFAULT PERCENTILE VERSUS ALTERNATIVE PERCENTILE

The "Comparison of Several Percentile Definitions" document summarized and compared several statistical methods to estimate the 99th percentile given multiple samples from a probability distribution and identified which of these methods is used in the NIOSH-IREP software, for the purpose of computing the 99th percentile of PC (99th PC values). The default percentile definition in Analytica, which is currently used in the NIOSH-IREP software (Type 7), underestimates the theoretical 99th percentile in all the scenarios tested (10 scenarios involving one distribution, and 20 scenarios involving the sum or product of two distributions), especially for sample sizes lower than 20,000 values. Based on the simulation scenarios tested, it was determined that the percentile definition used as default in the SAS software (Type 2) is the method with the smallest relative bias and smallest relative root mean squared error in estimating the theoretical 99th percentile at all sample sizes. In this report, the Type 7 percentile definition currently used in NIOSH-IREP will be denoted as the Default percentile definition, and the Type 2 percentile definition used in the SAS software will be denoted as the Alternative percentile definition. The purpose of this document is to compare the effect of using the Alternative percentile instead of the Default percentile on the 99th PC values reported by NIOSH-IREP.

3.0 SELECTION OF TEST CLAIMS

Several sets of claims were selected from the NOCTS database to test the effect of using the Alternative percentile instead of the Default percentile, on the 99th PC values. The cases were chosen using a query dated December 17, 2019.

Five sets of cases were selected based on the following ranges of the 99th PC values:

- Set 1: single cancer cases with 99th PC values between 44.00% and 44.99%
- Set 2: single cancer cases with 99th PC values between 45.00% and 48.99%
- Set 3: single cancer cases with 99th PC values between 49.00% and 49.99%
- Set 4: multiple cancer cases with Combined 99th PC values between 49.00% and 49.99%
- Set 5: single cancer cases with 99th PC values between 50.00% and 51.99%.

3.1. Single Cancer Cases with 99th PC Values between 44.00% and 44.99%

Of the 283 single cancer cases with 99th PC values between 44.00% and 44.99%, 45 cases were randomly selected for testing purposes, such that at least one case was selected for the 29 different IREP Cancer Models available in this range.

Each of the 45 cases selected had the 99th PC value between 44.00% and 44.99%, based on the IREP version in use at the time the case was processed. Since there were multiple IREP versions used for the selected cases, each of the 45 cases was rerun on current IREP version (v.5.8.2 version). Rerunning the cases on the IREP v.5.8.2 version would result in different PC values than the PC values on record for the cases that were run on IREP versions before the IREP v.5.8.2 version was introduced. As a result of this, some of the PC values for these 45 cases obtained on IREP v.5.8.2 version might not fall in the 44.00% to 44.99% range. All the cases with the 99th PC values between 44.00% and 44.99% were originally run at 2,000 iterations and random seed 99, and the same number of iterations and random seed was used when each case was rerun on IREP v.5.8.2 version. In addition to each case being rerun at 2,000 iterations, it was also run at 10,000 and 20,000 iterations on IREP v.5.8.2. For each case and each number of iterations used, the 99th PC value computed using the Default percentile and the 99th PC value computed using the Alternative percentile were recorded.

Table 1 shows summary statistics for the differences between the Alternative percentile and the Default percentile for the 99th PC values for the 45 cases selected, computed for each number of iterations. The Alternative percentile is greater than the Default percentile for all cases at 2,000 and 10,000 iterations, and greater or equal at 20,000 iterations. The difference between Alternative percentile and the Default percentile has an average increase of 0.27, with a maximum increase of 1.46 at 2,000 iterations, an average increase of 0.06, with a maximum increase of 0.27 at 10,000 iterations, and an average increase of 0.02, with a maximum increase of 0.17 at 20,000 iterations.

Table 1. 99th PC difference (Alternative percentile - Default percentile) for cases with PC values between 44.00% and 44.99%.

Number of Iterations	N	Minimum	25th %ile	Median	Mean	75th %ile	95th %ile	Maximum
2,000	45	0.02	0.11	0.17	0.27	0.35	0.69	1.46
10,000	45	0.01	0.03	0.05	0.06	0.08	0.16	0.27
20,000	45	0.00	0.01	0.01	0.02	0.03	0.05	0.17

Most of the cases had an increase from the Default percentile to the Alternative percentile of less than 1 at 2,000 iterations, except two cases which had increases of 1.06 and 1.46; however, the two cases with the largest increases between the two percentile definitions at 2,000 iterations are not the same cases who had the largest increases at 10,000 and 20,000 iterations, so is not exactly

clear what is causing these extreme differences between the two percentile definitions at 2,000 iterations.

Figure 1 shows the 99th PC values computed using both the Default percentile and the Alternative percentile for all 45 cases at 2,000 iterations. Since all cases were rerun on IREP v.5.8.2, just a fraction of the recomputed PC values will fall in the 44.00% to 44.99% range, with most of them being outside this range.

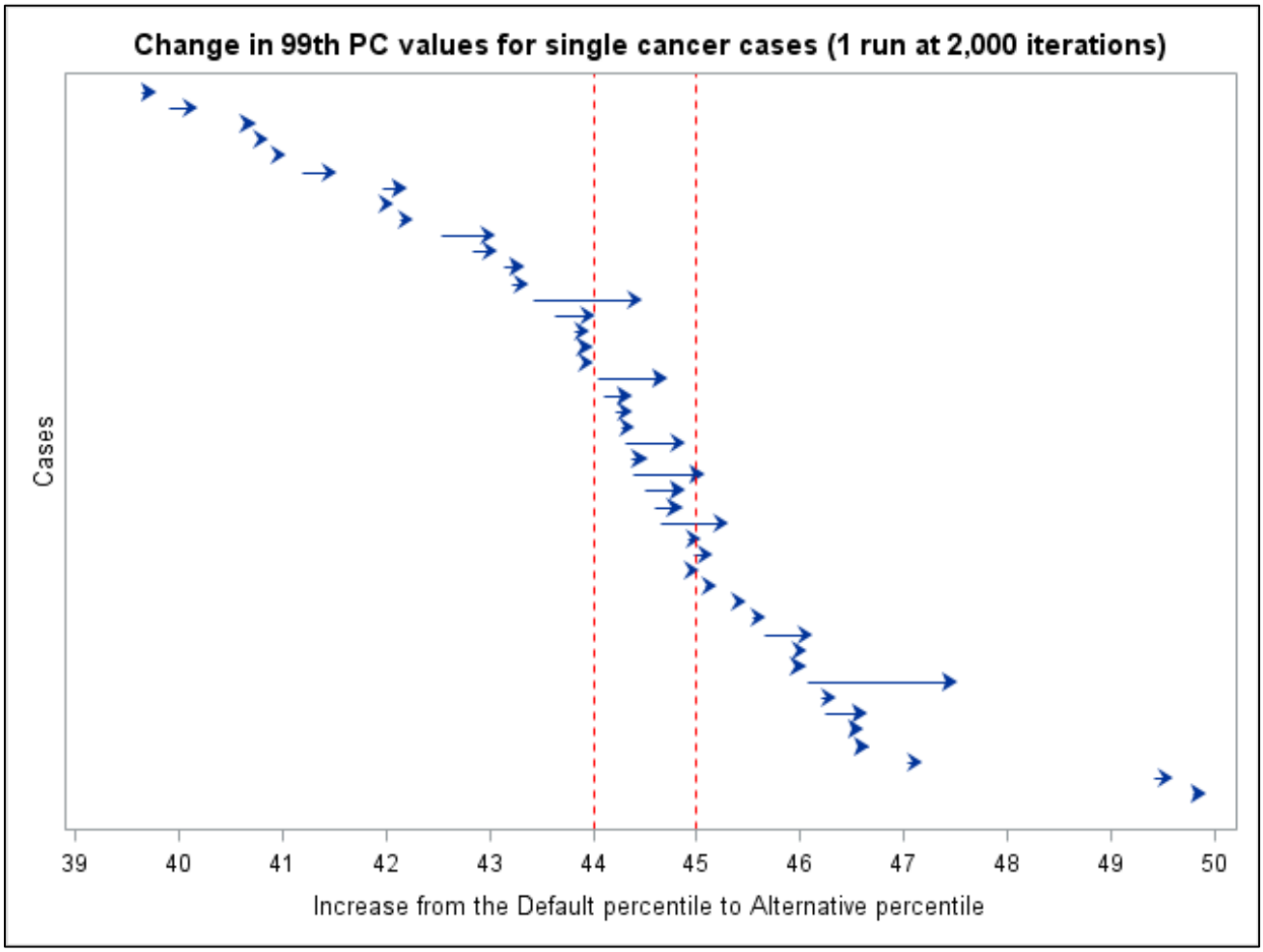

Figure 1. The increase in the 99th PC values from using the Default percentile versus using the Alternative percentile at 2,000 iterations, for the 45 cases with PC values between 44.00% and 44.99%.

Figure 2 shows the magnitude of the increase in the 99th PC value from using the Default percentile versus using the Alternative percentile at 2,000 iterations for all 45 cases, sorted from the lowest increase to the highest increase. The plot also shows the IREP Cancer Model for each of the 45 cases.

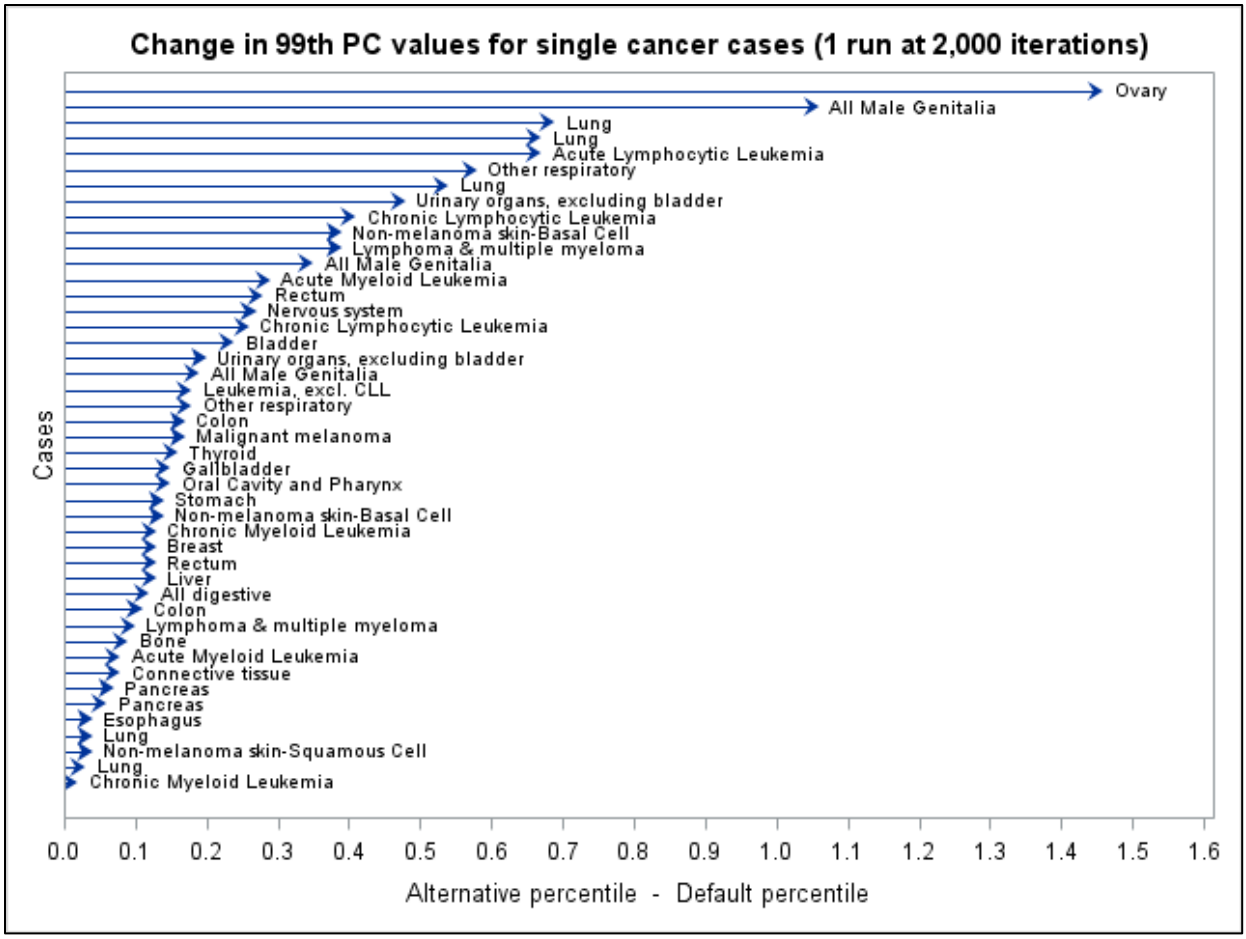

Figure 2. The increase in the 99th PC value from using the Default percentile versus using the Alternative percentile at 2,000 iterations, for the 45 cases with PC values between 44.00% and 44.99%.

Figure 3 shows a needle plot with the increase in the 99th PC value from using the Default percentile to using the Alternative percentile at 2,000 and 10,000 iterations for all 45 cases. The plot also shows the reference lines corresponding to the average difference between the Alternative percentile and the Default percentile at 2,000 and 10,000 iterations. As it can be seen from this plot, for most of the cases the differences between the Alternative percentile and the Default percentile are much larger at 2,000 than at 10,000 iterations, with just a few exceptions for which the Default percentile and the Alternative percentile values are almost identical.

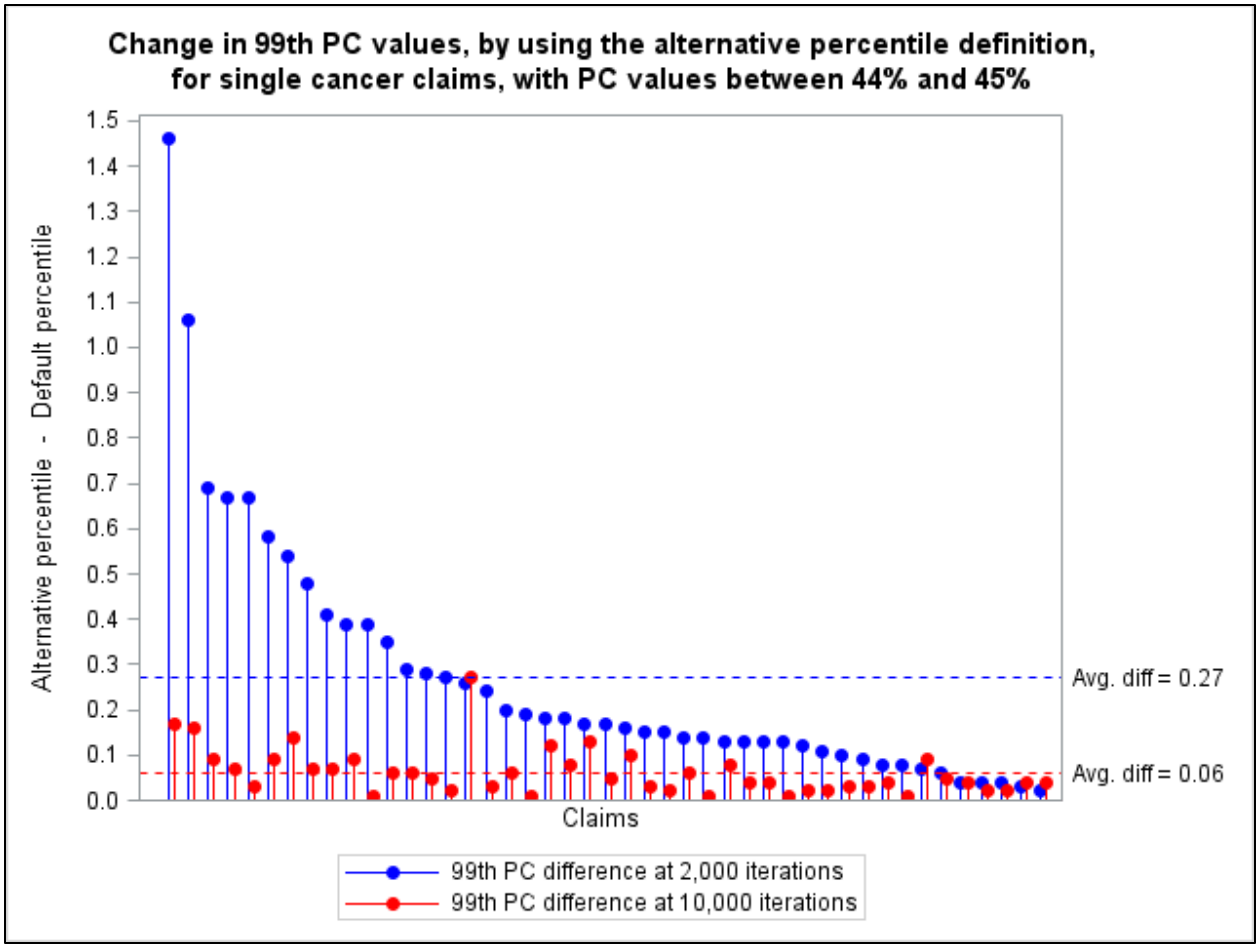

Figure 3. Increase in the 99th PC values from the Default percentile to the Alternative percentile at 2,000 and 10,000 iterations, for the 45 cases with PC values between 44.00% and 44.99%.

3.2. Single Cancer Cases with 99th PC values between 45.00% and 48.99%

Of the 134 single cancer cases with 99th PC values between 45.00% and 48.99%, 25 cases were randomly selected for testing purposes, such that at least one case was selected for the 24 different IREP Cancer Models available in this range.

Each of the 25 cases selected had either the 99th PC value obtained using the IREP version or the Avg. 99th PC value obtained using the IREP Enterprise Edition (IREP EE) version between 45.00% and 48.99%, based on the IREP or IREP EE version in use at the time the case was processed. Since there were multiple IREP versions used for all the selected cases, each of the 25 cases was rerun on IREP EE v.5.8.2 version. Rerunning the cases on the IREP EE v.5.8.2 version would result in different PC values than the PC values on record for the cases that were run on

IREP versions before the IREP EE v.5.8.2 version was introduced. As a result of this, some of the PC values obtained on IREP EE v.5.8.2 version might not fall in the 45.00% to 48.99% range. Most of the cases with the 99th PC values between 45.00% and 48.99% were originally run at 10,000 iterations and using a set of 30 different random seeds, and the same set of 30 random seeds were used when each case was rerun on IREP EE v.5.8.2 version. In addition to each case being rerun at 10,000 iterations, it was also run at 20,000 and 30,000 iterations on IREP EE v.5.8.2. For each case and each number of iterations used, the Avg. 99th PC value computed using the Default percentile and the Avg. 99th PC value computed using the Alternative percentile were recorded.

Table 2 shows summary statistics for the differences between the Alternative and Default percentiles for the Avg. 99th PC values for the 25 cases, computed for each number of iterations. The Alternative percentile is greater than the Default percentile for all cases and each number of iterations. The difference between Alternative percentile and the Default percentile has an average increase of 0.04, with a minimum increase of 0.02, and maximum increase of 0.06 at 10,000 iterations, an average increase of 0.02, with a maximum increase of 0.03 at 20,000 iterations, and an average increase of 0.016, with a maximum increase of 0.03 at 30,000 iterations.

Table 2. Avg. 99th PC difference (Alternative percentile - Default percentile) for cases with PC values between 45.00% and 48.99%.

Number of Iterations	N	Minimum	25th %ile	Median	Mean	75th %ile	Maximum
10,000	25	0.02	0.03	0.04	0.04	0.05	0.06
20,000	25	0.01	0.02	0.02	0.02	0.02	0.03
30,000	25	0.01	0.01	0.02	0.016	0.02	0.03

Figure 4 shows the magnitude of the increase in the Avg. 99th PC value from using the Default percentile versus using the Alternative percentile based on 30 runs at 10,000 iterations for all 25 cases, sorted from the lowest increase to the highest increase. The plot also shows the IREP Cancer Model for each of the 25 cases.

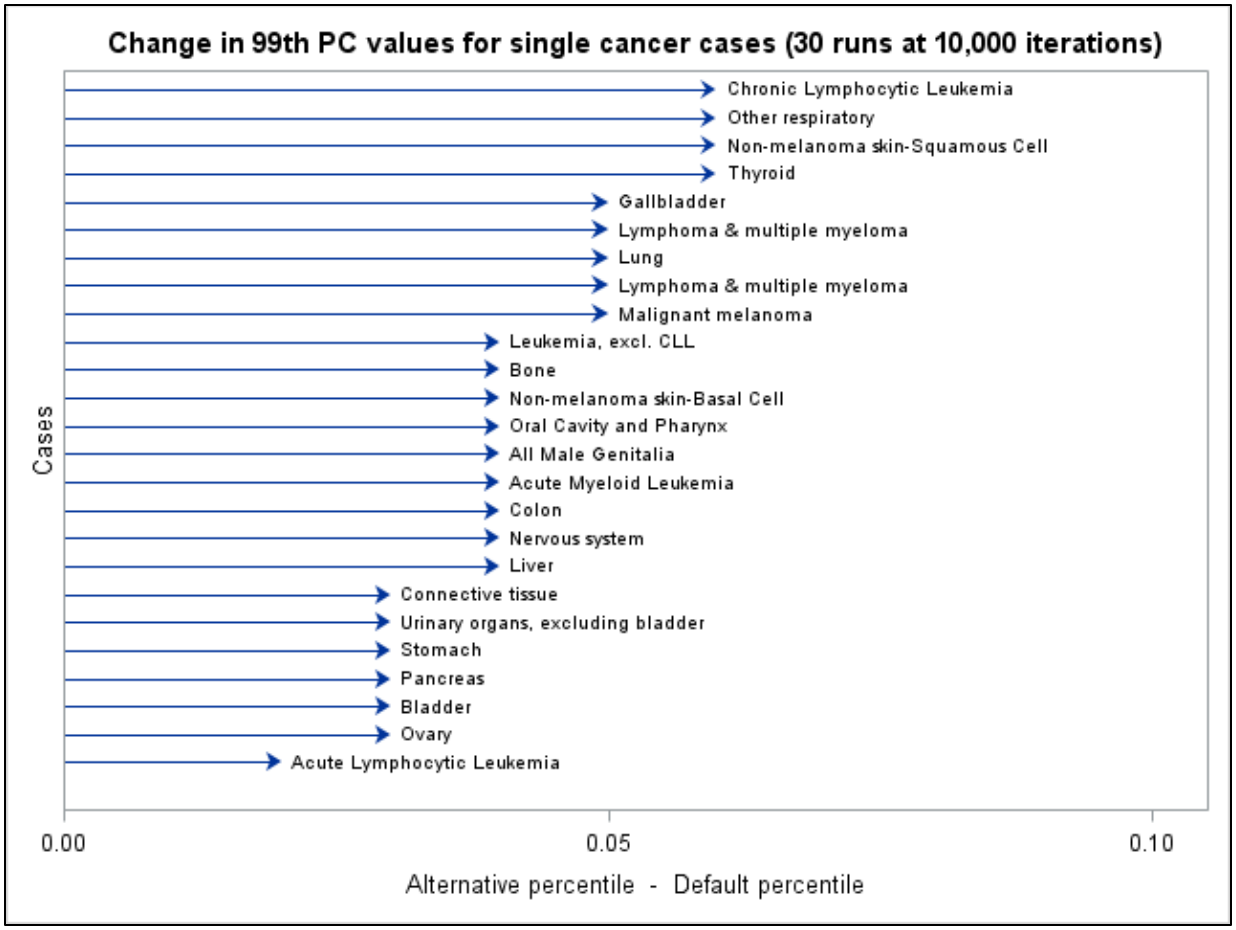

Figure 4. The increase in the Avg. 99th PC value from using the Default percentile versus using the Alternative percentile, for the 25 cases with 99th PC values between 45.00% and 48.99%.

Figure 5 shows a needle plot with the increase in the Avg. 99th PC value from using the Default percentile to using the Alternative percentile at 2,000, 10,000, and 20,000 iterations for all 25 cases. The Avg. 99th PC value at 2,000 iterations was computed and is shown here just for the purpose of comparing it with the results obtained for the first set of cases (cases with 99th PC values between 44.00% and 44.99%), when only one run at 2,000 iterations was used. The plot also shows the reference lines corresponding to the average difference between the Alternative percentile and the Default percentile at 2,000, 10,000, and 20,000 iterations.

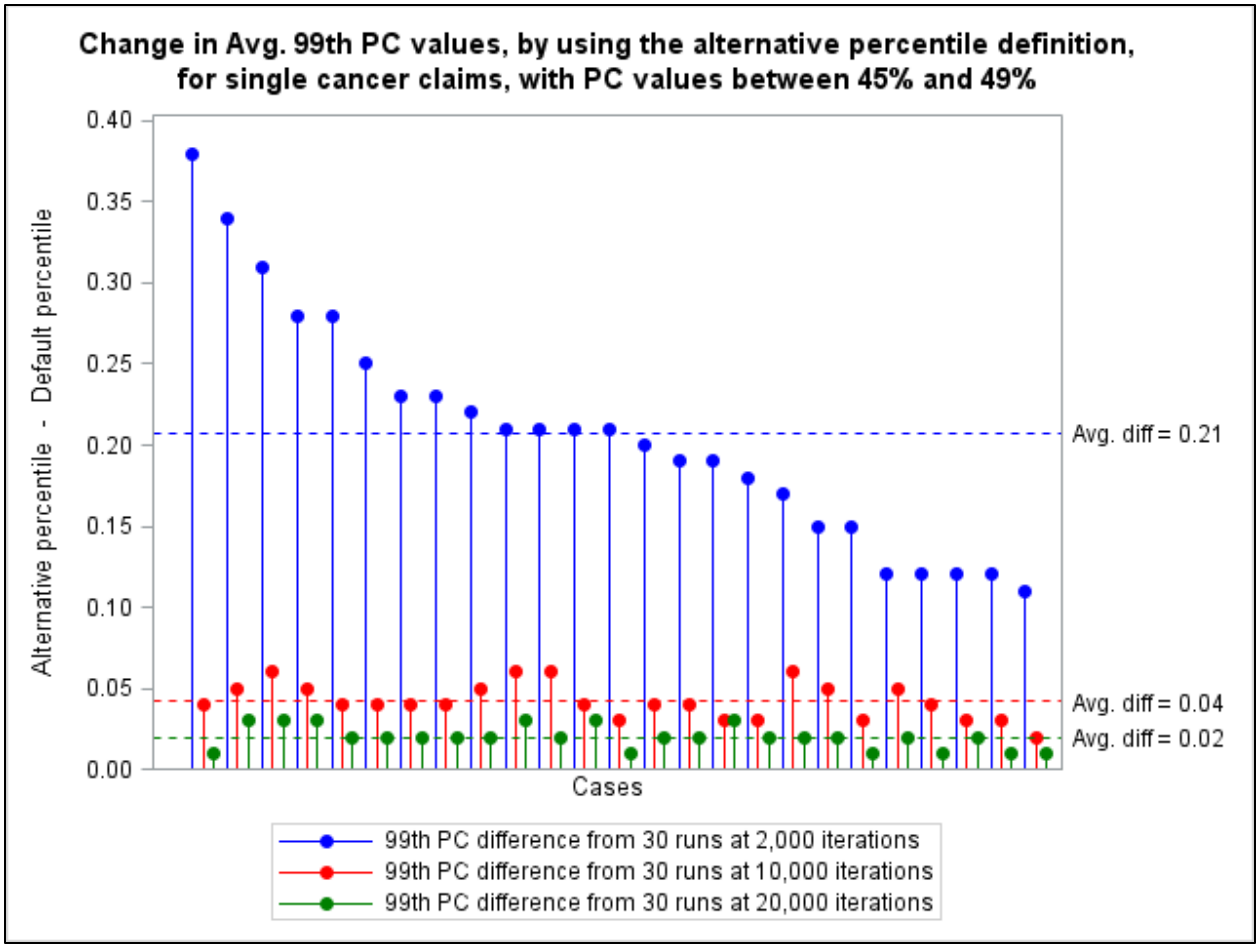

Figure 5. Increase in the Avg. 99th PC values from the Default percentile to the Alternative percentile at 2,000, 10,000, and 20,000 iterations, for the 25 cases with 99th PC values between 45.00% and 48.99%.

3.3. Single Cancer Cases with 99th PC values between 49.00% and 49.99%

All 22 single cancer cases with 99th PC values between 49.00% and 49.99% were selected for testing purposes, but there are only 12 IREP Cancer Models corresponding to the 22 cases in this range.

Each of the 22 cases had an Avg. 99th PC value obtained using the IREP Enterprise Edition version between 49.00% and 49.99%, based on the IREP EE version in use at the time the case was processed. Since there were multiple IREP EE versions used for all the selected cases, each of the 22 cases was rerun on IREP EE v.5.8.2 version. Rerunning the cases on the IREP EE v.5.8.2 version would result in different PC values than the PC values on record for the cases that

were run on IREP EE versions before the IREP EE v.5.8.2 version was introduced. As a result of this, some of the PC values obtained on IREP v.5.8.2 version might not fall in the 49.00% to 49.99% range. All the cases with the Avg. 99th PC value between 49.00% and 49.99% were originally run at 10,000 iterations and using a set of 30 different random seeds, and the same set of 30 random seeds were used when each case was rerun on IREP EE v.5.8.2 version. In addition to each case being rerun at 10,000 iterations, it was also run at 20,000 and 30,000 iterations on IREP EE v.5.8.2. For each case and each number of iterations used, the Avg. 99th PC value computed using the Default percentile and the Avg. 99th PC value computed using the Alternative percentile were recorded.

Table 3 shows summary statistics for the differences between the Alternative percentile and the Default percentile for the Avg. 99th PC values for the 22 cases, computed for each number of iterations. The Alternative percentile is greater than the Default percentile for all cases and each number of iterations. The difference between Alternative percentile and the Default percentile has an average increase of 0.04, with a minimum increase of 0.02, and maximum increase of 0.07 at 10,000 iterations, an average increase of 0.02, with a maximum increase of 0.04 at 20,000 iterations, and an average increase of 0.015, with a maximum increase of 0.02 at 30,000 iterations. These statistics shown in Table 3 for the cases with 99th PC values between 49.00% and 49.99% are very similar to the statistics shown in Table 2 for the cases with 99th PC values between 45.00% and 48.99%.

Table 3. Avg. 99th PC difference (Alternative percentile - Default percentile) for single cancer cases with PC values between 49.00% and 49.99%.

Number of Iterations	N	Minimum	25th %ile	Median	Mean	75th %ile	Maximum
10,000	22	0.02	0.03	0.04	0.04	0.05	0.07
20,000	22	0.01	0.02	0.02	0.02	0.03	0.04
30,000	22	0.01	0.01	0.015	0.015	0.02	0.02

Figure 6 shows the Avg. 99th PC values computed using both the Default percentile and the Alternative percentile for all 22 cases, based on 30 runs at 10,000 iterations. As it can be seen from the plot, three of the PC values recomputed on the IREP EE v.5.8.2 fall outside the 49.00% to 49.99% range.

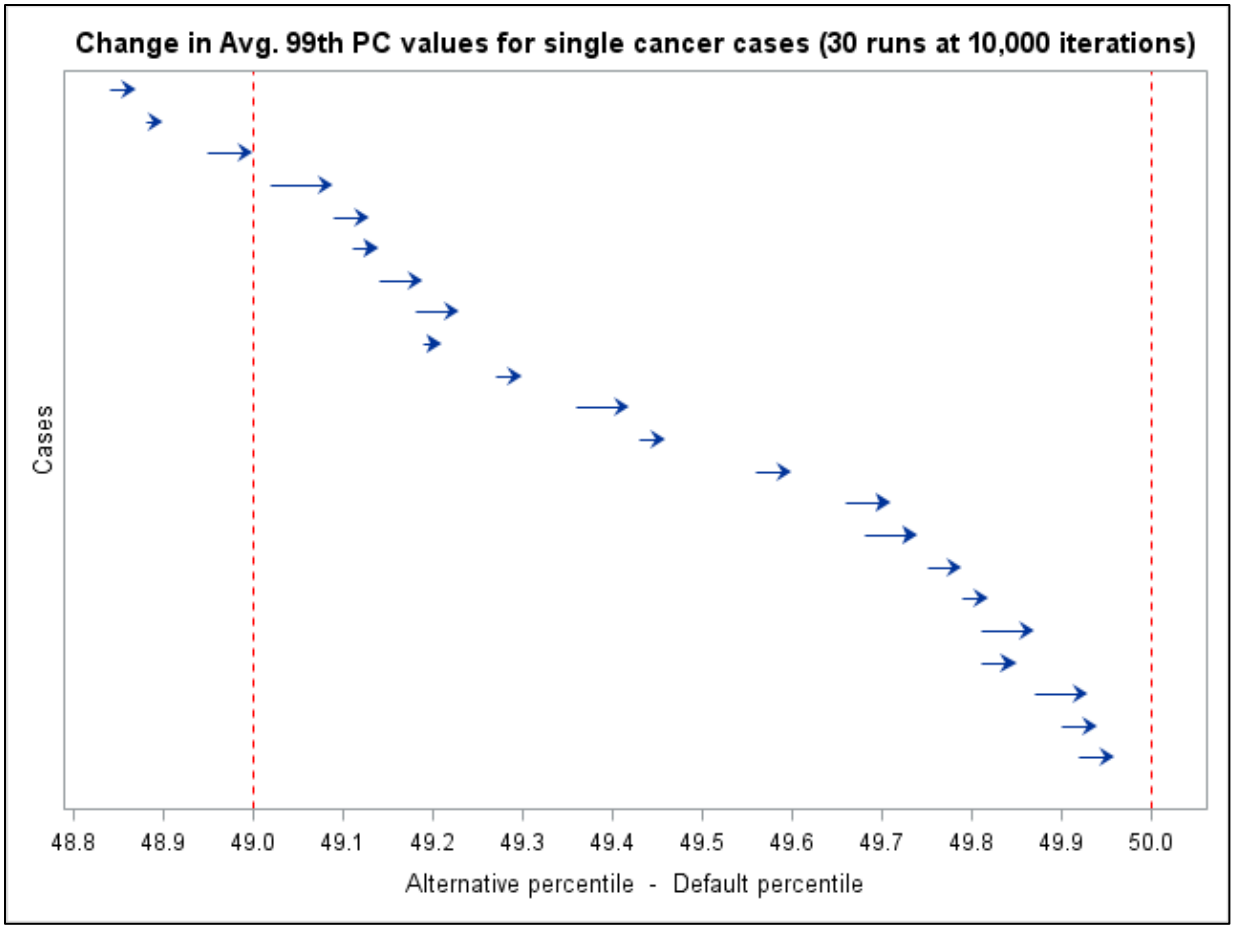

Figure 6. The increase in the Avg. 99th PC values from using the Default percentile versus using the Alternative percentile, for the 22 single cancer cases with PC values between 49.00% and 49.99%, based on 30 runs at 10,000 iterations.

Figure 7 shows the magnitude of the increase in the Avg. 99th PC value from using the Default percentile versus using the Alternative percentile for all 22 cases, sorted from the lowest increase to the highest increase. The plot also shows the IREP Cancer Model for each of the 22 cases.

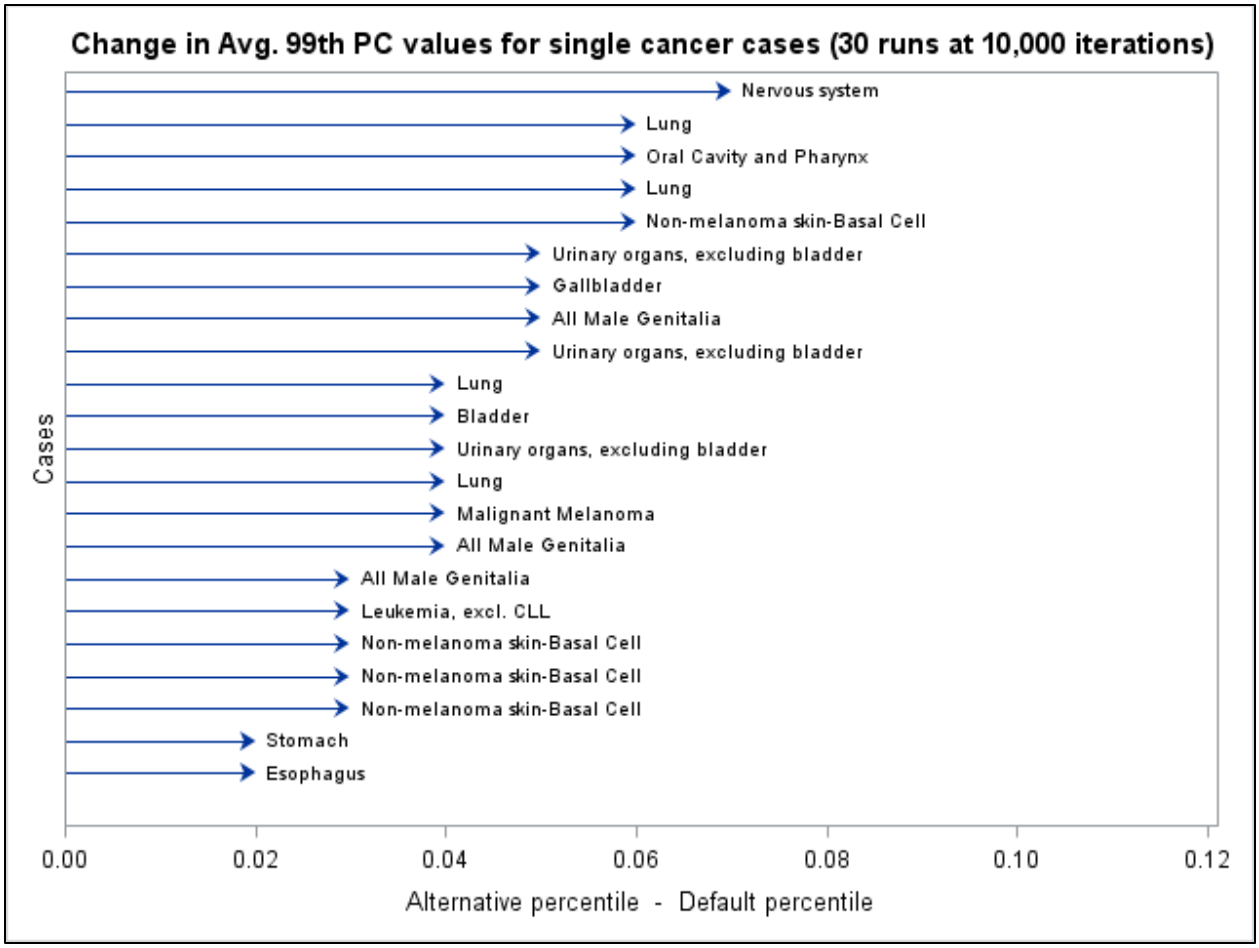

Figure 7. The increase in the Avg. 99th PC value from using the Default percentile versus using the Alternative percentile, for the 22 single cancer cases with Avg. 99th PC values between 49.00% and 49.99%.

Figure 8 shows the Avg. 99th PC values computed using both the Default percentile and the Alternative percentile for all 22 cases based on 30 runs at 20,000 iterations. The lengths of the arrows in Figure 8 are smaller than the lengths of the arrows in Figure 6, showing the smaller increases in PC values at 20,000 iterations versus 10,000 iterations. By running the 22 cases at 20,000 iterations, two of the 22 cases reach or surpass the 50.00% PC threshold using either the Default percentile or the Alternative percentile.

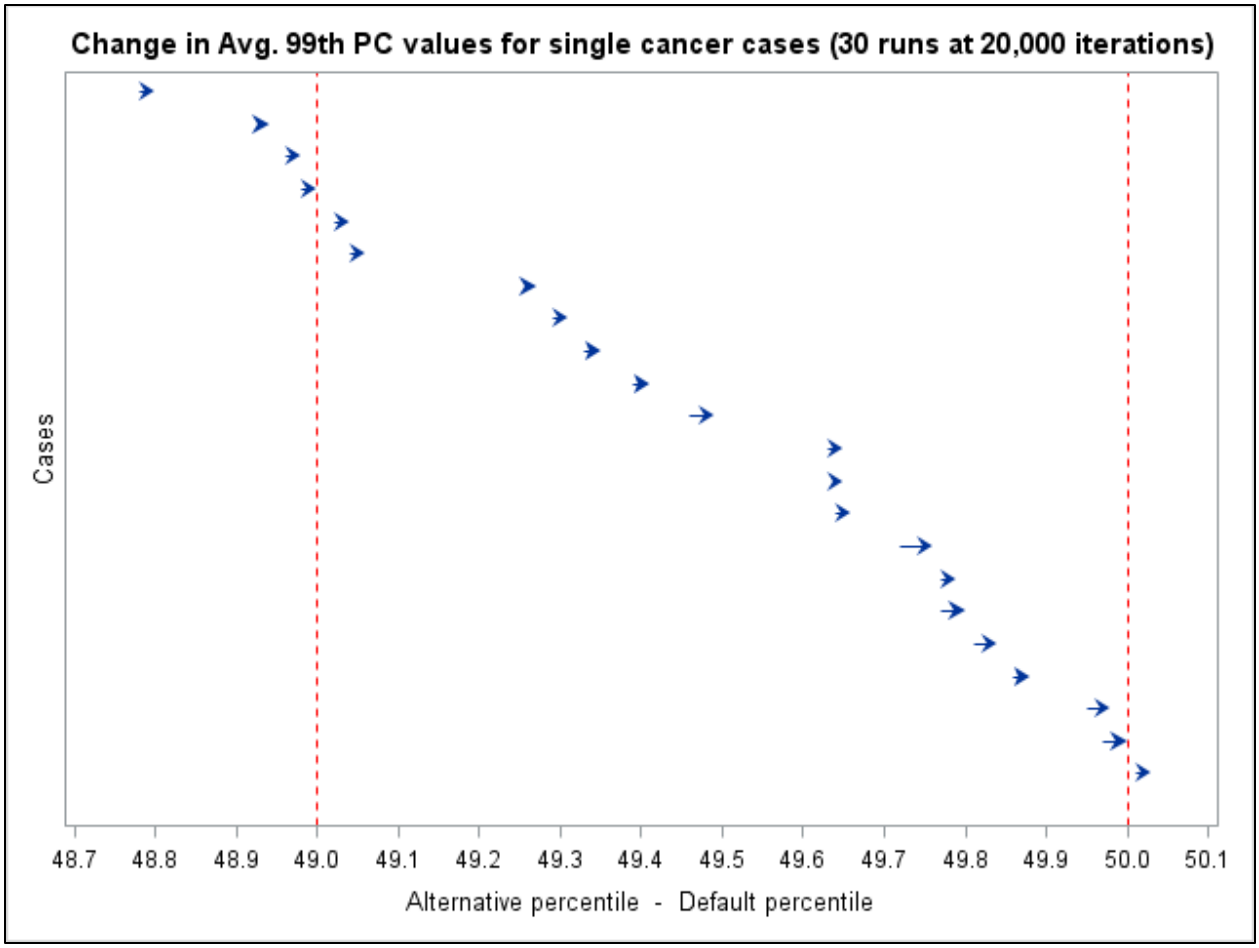

Figure 8. The increase in the Avg. 99th PC values from using the Default percentile versus using the Alternative percentile, for the 22 single cancer cases with PC values between 49.00% and 49.99%, based on 30 runs at 20,000 iterations.

Figure 9 shows the Avg. 99th PC values computed using both the Default percentile and the Alternative percentile for all 22 cases based on 30 runs at 30,000 iterations. The lengths of the arrows in Figure 9 have slightly smaller magnitudes than the lengths of the arrows in Figure 8, showing even smaller increases in PC values at 30,000 iterations versus 20,000 iterations. By running the 22 cases at 30,000 iterations, two of the 22 cases reach or surpass the 50.00% threshold using either the Default percentile or the Alternative percentile.

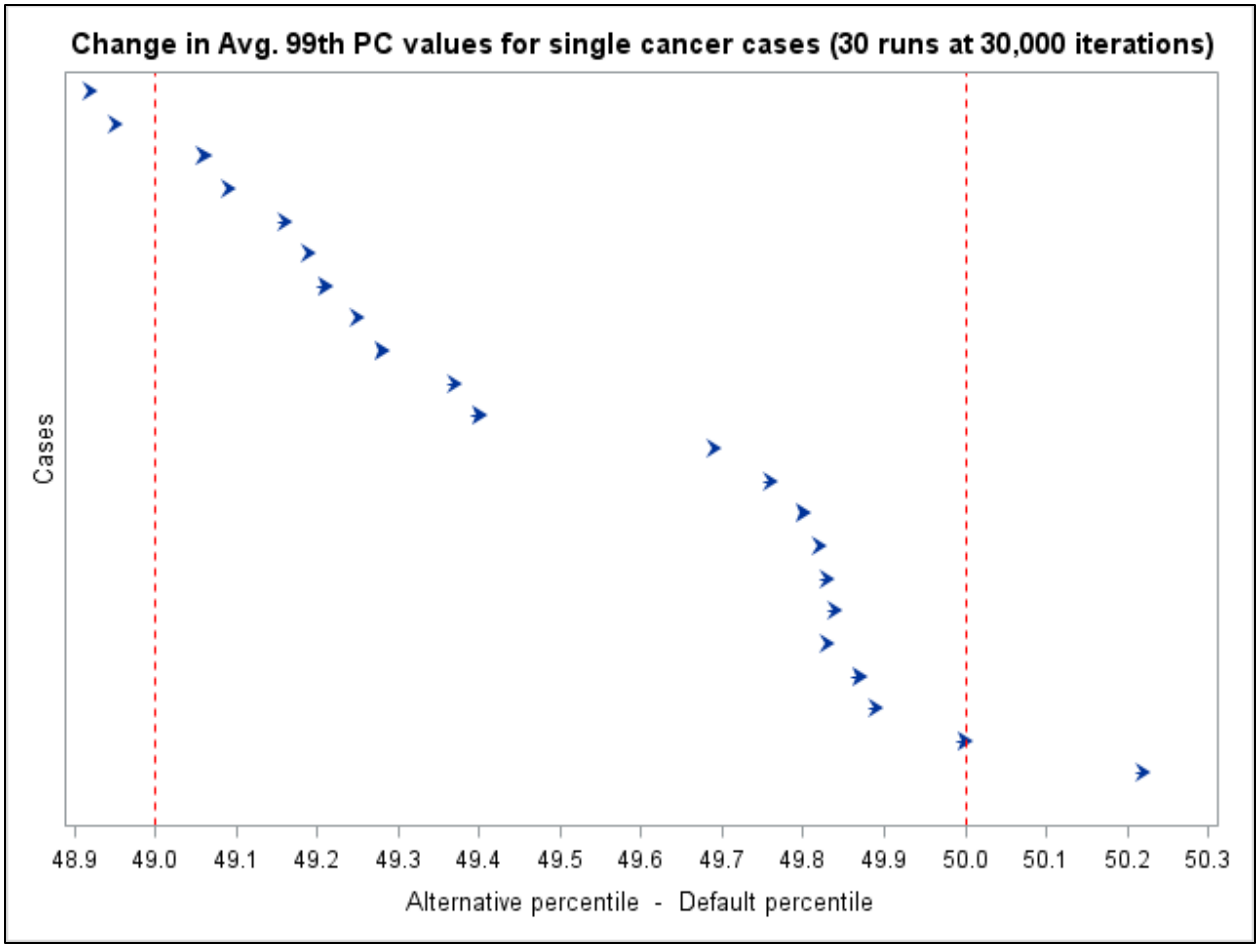

Figure 9. The increase in the Avg. 99th PC values from using the Default percentile versus using the Alternative percentile, for the 22 single cancer cases with PC values between 49.00% and 49.99%, based on 30 runs at 30,000 iterations.

Figure 10 shows a needle plot with the increase in the Avg. 99th PC value from using the Default percentile to using the Alternative percentile at 10,000, 20,000, and 30,000 iterations for all 22 cases. The plot also shows the reference lines corresponding to the average difference between the Alternative percentile and the Default percentile at 10,000, 20,000, and 30,000 iterations. While the differences in the Avg. 99th PC values from the Default percentile to the Alternative percentile at 10,000 iterations are almost always greater than the differences at 20,000 iterations, that's not necessarily true for the differences at 20,000 versus 30,000 iterations, with several cases having the same difference in the Avg. 99th PC at 20,000 and 30,000 iterations.

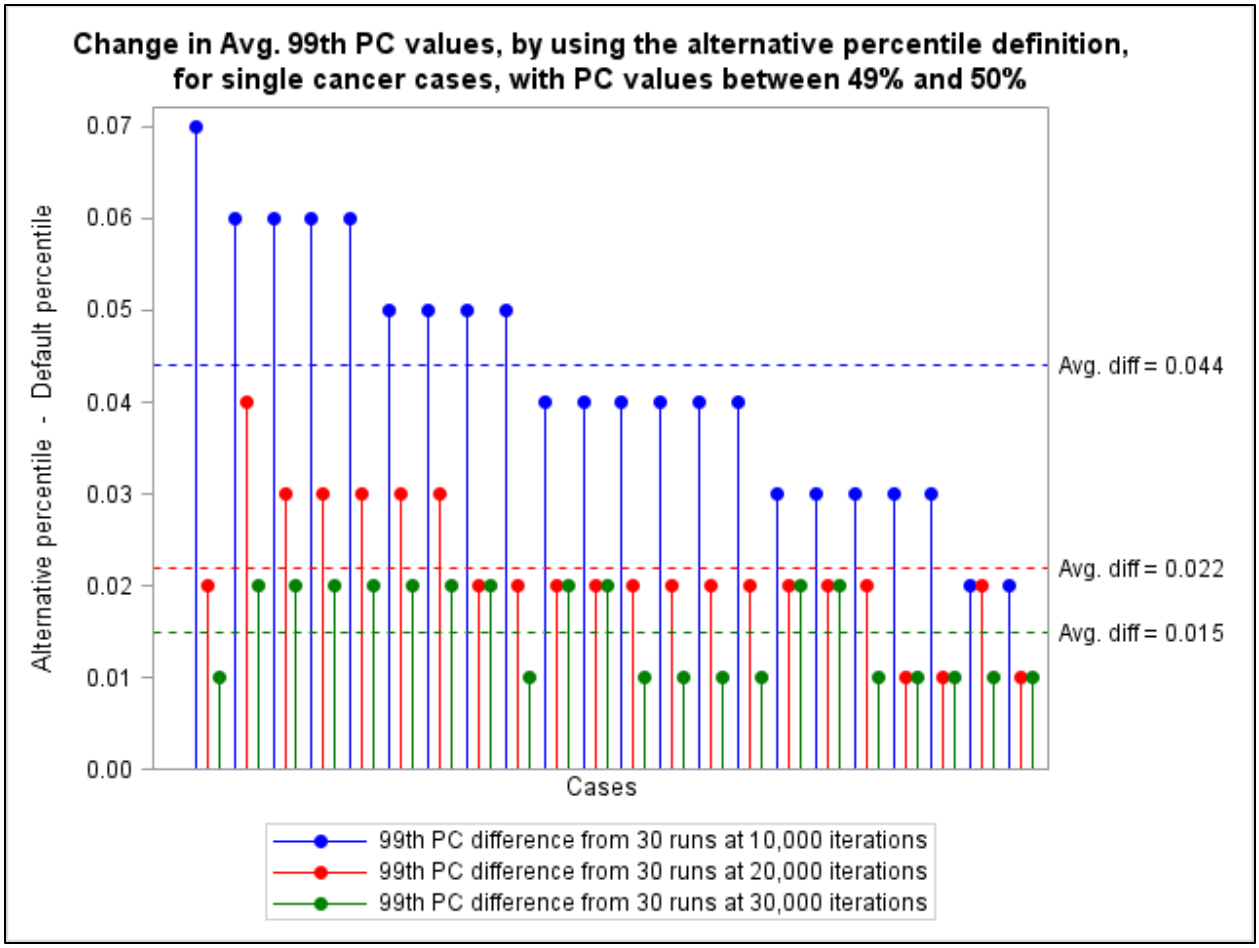

Figure 10. Increase in the Avg. 99th PC values from the Default percentile to the Alternative percentile at 10,000, 20,000, and 30,000 iterations, for the 22 single cancer cases with Avg. 99th PC values between 49.00% and 49.99%.

3.4. Multiple Cancer Cases with 99th PC values between 49.00% and 49.99%

In addition to the single cancer cases with 99th PC values between 49.00% and 49.99%, all the 36 multiple cancer cases with Combined 99th PC values between 49.00% and 49.99% were also selected for testing purposes. The number of cancers varies between 2 and 13 cancers, for the 36 multiple cancer cases selected.

Each of the 36 cases selected had the Combined Avg. 99th PC value obtained using the IREP Enterprise Edition version between 49.00% and 49.99%, based on the IREP EE version in use at the time the case was processed. Since there were multiple IREP versions used for all the selected cases, each of the 36 cases was rerun on IREP EE v.5.8.2 version. Rerunning the cases

on the IREP EE v.5.8.2 version would result in different PC values than the PC values on record for the cases that were run on IREP versions before the IREP EE v.5.8.2 version was introduced. As a result of this, some of the PC values obtained on IREP EE v.5.8.2 version might not fall in the 49.00% to 49.99% range. All the cases with the Combined Avg. 99th PC value between 49.00% and 49.99% were originally run at 10,000 iterations and using a set of 30 different random seeds, and for a given case the same sets of 30 random seeds used for each cancer were also used when that case was rerun on IREP EE v.5.8.2 version. In addition to each case being rerun at 10,000 iterations, it was also run at 20,000 iterations on IREP EE v.5.8.2. For each case and each number of iterations used, the Combined Avg. 99th PC value computed using the Default percentile and the Combined Avg. 99th PC value computed using the Alternative percentile were recorded.

Table 4 shows summary statistics for the differences between the Alternative percentile and the Default percentile for the Combined Avg. 99th PC values for the 36 cases, computed for each number of iterations. The Alternative percentile is greater than the Default percentile for all cases and each number of iterations. The difference between Alternative percentile and the Default percentile has an average increase of 0.05, with a minimum increase of 0.03, and maximum increase of 0.09 at 10,000 iterations, and an average increase of 0.03, with a maximum increase of 0.05 at 20,000 iterations.

Table 4. Avg. 99th PC difference (Alternative percentile - Default percentile) for multiple cancer cases with PC values between 49.00% and 49.99%.

Number of Iterations	N	Minimum	25th %ile	Median	Mean	75th %ile	Maximum
10,000	36	0.03	0.04	0.05	0.05	0.06	0.09
20,000	36	0.01	0.02	0.02	0.03	0.03	0.05

These statistics shown in Table 4 for the multiple cancer cases with Combined Avg. 99th PC values between 49.00% and 49.99% are slightly higher than the statistics shown in Table 3 for the single cancer cases with Avg. 99th PC values between 49.00% and 49.99%, which suggests that the increase between the Default and Alternative percentile is going to be higher for the multiple cancer cases, due to the compounding effect of combining the larger PC values obtained using the Alternative percentile versus combining the smaller PC values obtained using the Default percentile.

Figure 11 shows the Avg. 99th PC values computed using both the Default percentile and the Alternative percentile for all 36 cases, based on 30 runs at 10,000 iterations. The lengths of the arrows in Figure 11 are larger than the lengths of the arrows in Figure 6, showing the larger increases in PC values for the multiple cancer cases versus the single cancer cases. As it can be seen from the plot, three of the PC values recomputed on the IREP EE v.5.8.2 fall outside the

49.00% to 49.99% range, and two cases reach or surpass the 50.00% PC threshold using the Alternative percentile.

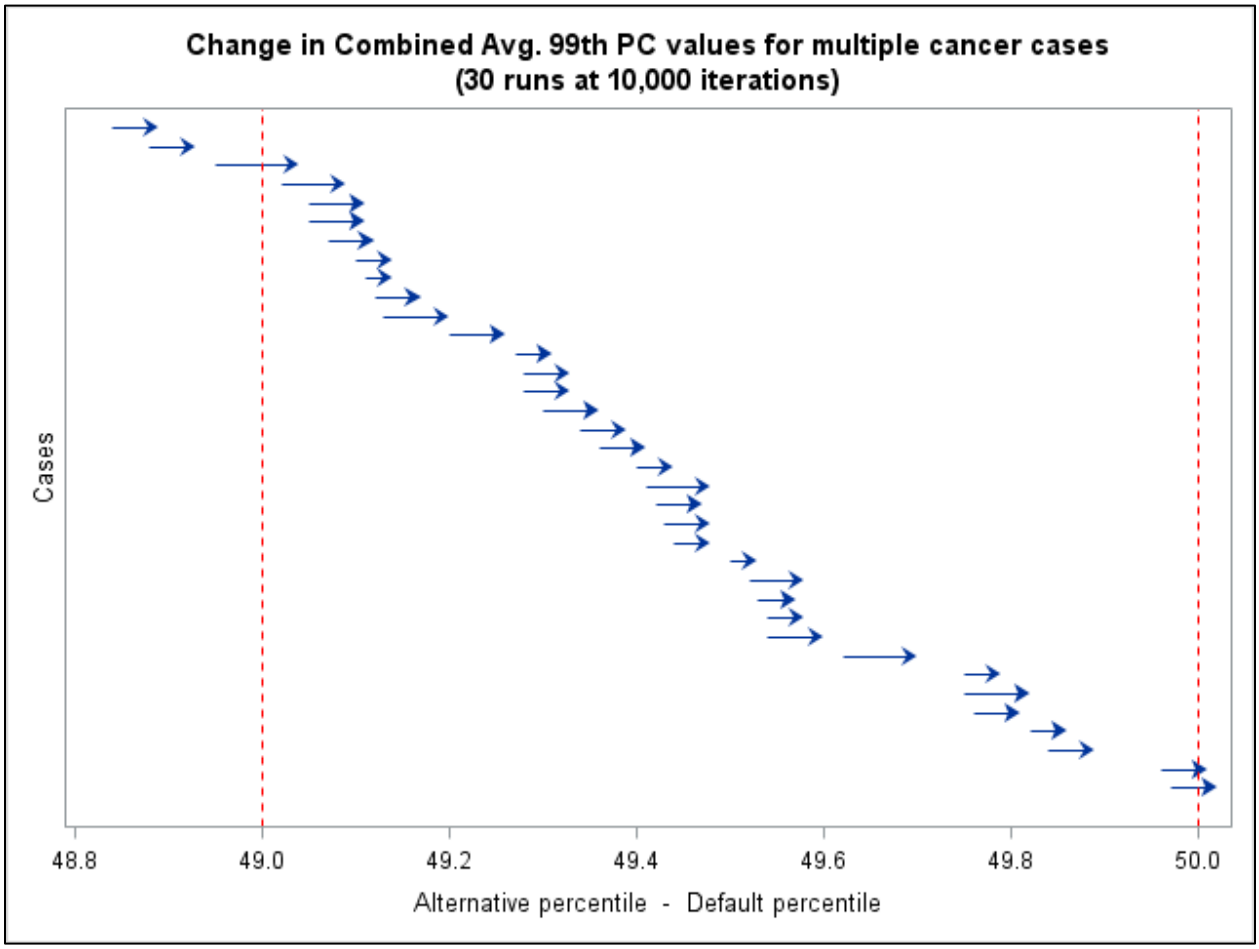

Figure 11. The increase in the Combined Avg. 99th PC values from using the Default percentile versus using the Alternative percentile, for the 36 multiple cancer cases with Combined Avg. 99th PC values between 49.00% and 49.99%.

Figure 12 shows a needle plot with the increase in the Combined Avg. 99th PC value from using the Default percentile to using the Alternative percentile at 10,000, and 20,000 iterations for all 36 cases. The plot also shows the reference lines corresponding to the average difference between the Alternative percentile and the Default percentile at 10,000, and 20,000 iterations. For all cases except one, the difference in the Combined Avg. 99th PC values from the Default percentile to the Alternative percentile at 10,000 iterations is greater than the difference at 20,000 iterations.

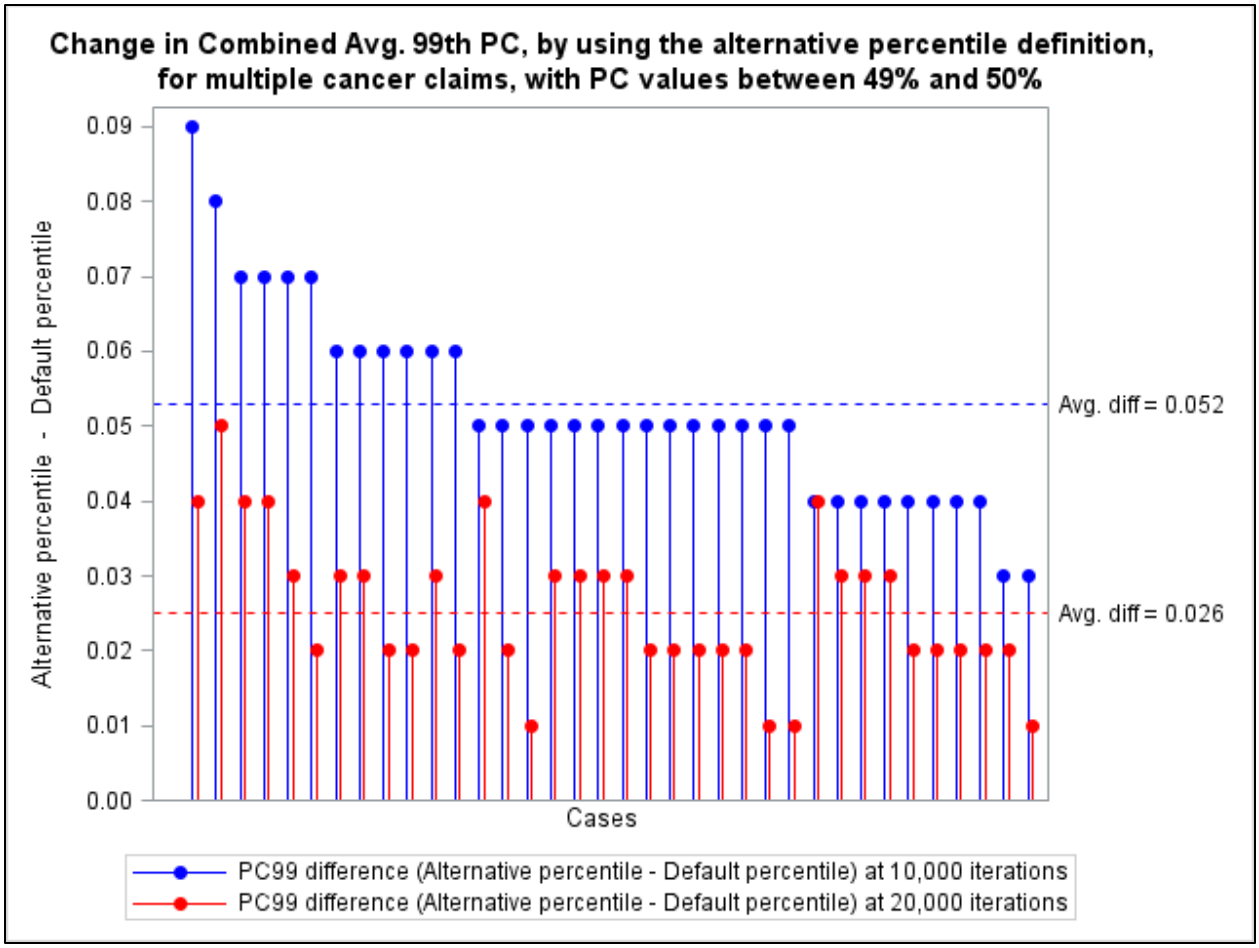

Figure 12. Increase in the Combined Avg. 99th PC values from the Default percentile to the Alternative percentile at 10,000, and 20,000 iterations, for the 36 multiple cancer cases with Combined Avg. 99th PC values between 49.00% and 49.99%.

3.5. Single Cancer Cases with 99th PC values between 50.00% and 51.99%

Of the 475 single cancer cases with 99th PC values between 50.00% and 51.99%, 25 cases were randomly selected for testing purposes.

Each of the 25 cases selected had the Avg. 99th PC value obtained using the IREP Enterprise Edition version between 50.00% and 51.99%, based on the IREP version in use at the time the case was processed. Since there were multiple IREP versions used for all the selected cases, each of the 25 cases was rerun on IREP EE v.5.8.2 version. Rerunning the cases on the IREP EE v.5.8.2 version would result in different PC values than the PC values on record for the cases that were run on IREP EE versions before the IREP EE v.5.8.2 version was introduced. As a result of

this, some of the PC values obtained on IREP EE v.5.8.2 version might not fall in the 50.00% to 51.99% range. All the cases with the Avg. 99th PC value between 50.00% and 51.99% were originally run at 10,000 iterations and using a set of 30 different random seeds, and the same set of 30 random seeds were used when each case was rerun on IREP EE v.5.8.2 version. In addition to each case being rerun at 10,000 iterations, it was also run at 20,000 and 30,000 iterations on IREP EE v.5.8.2. For each case and number of iterations used, the Avg. 99th PC value computed using the Default percentile and the Avg. 99th PC value computed using the Alternative percentile were recorded.

Table 5 shows summary statistics for the differences between the Alternative percentile and the Default percentile for the Avg. 99th PC values for the 25 cases, computed for each number of iterations. The Alternative percentile is greater than the Default percentile for all cases at 10,000 and 20,000 iterations, and greater or equal at 30,000 iterations. The difference between Alternative percentile and the Default percentile has an average increase of 0.05, with a minimum increase of 0.02, and maximum increase of 0.11 at 10,000 iterations, an average increase of 0.02, with a maximum increase of 0.05 at 20,000 iterations, and an average increase of 0.016, with a maximum increase of 0.04 at 30,000 iterations. These statistics shown in Table 5 for the cases with 99th PC values between 50.00% and 51.99% are comparable to the statistics shown in Tables 2 and 3, for the single cancer cases with 99th PC values between 45.00% and 48.99%, and respectively 49.00% and 49.99%.

Table 5. Avg. 99th PC difference (Alternative percentile - Default percentile) for cases with PC values between 50.00% and 51.99%.

Number of Iterations	N	Minimum	25th %ile	Median	Mean	75th %ile	Maximum
10,000	25	0.02	0.03	0.04	0.05	0.06	0.11
20,000	25	0.01	0.02	0.02	0.02	0.02	0.05
30,000	25	0.00	0.01	0.01	0.016	0.02	0.04

Figure 13 shows the Avg. 99th PC values computed using both the Default percentile and the Alternative percentile for all 25 cases, based on 30 runs at 10,000 iterations.

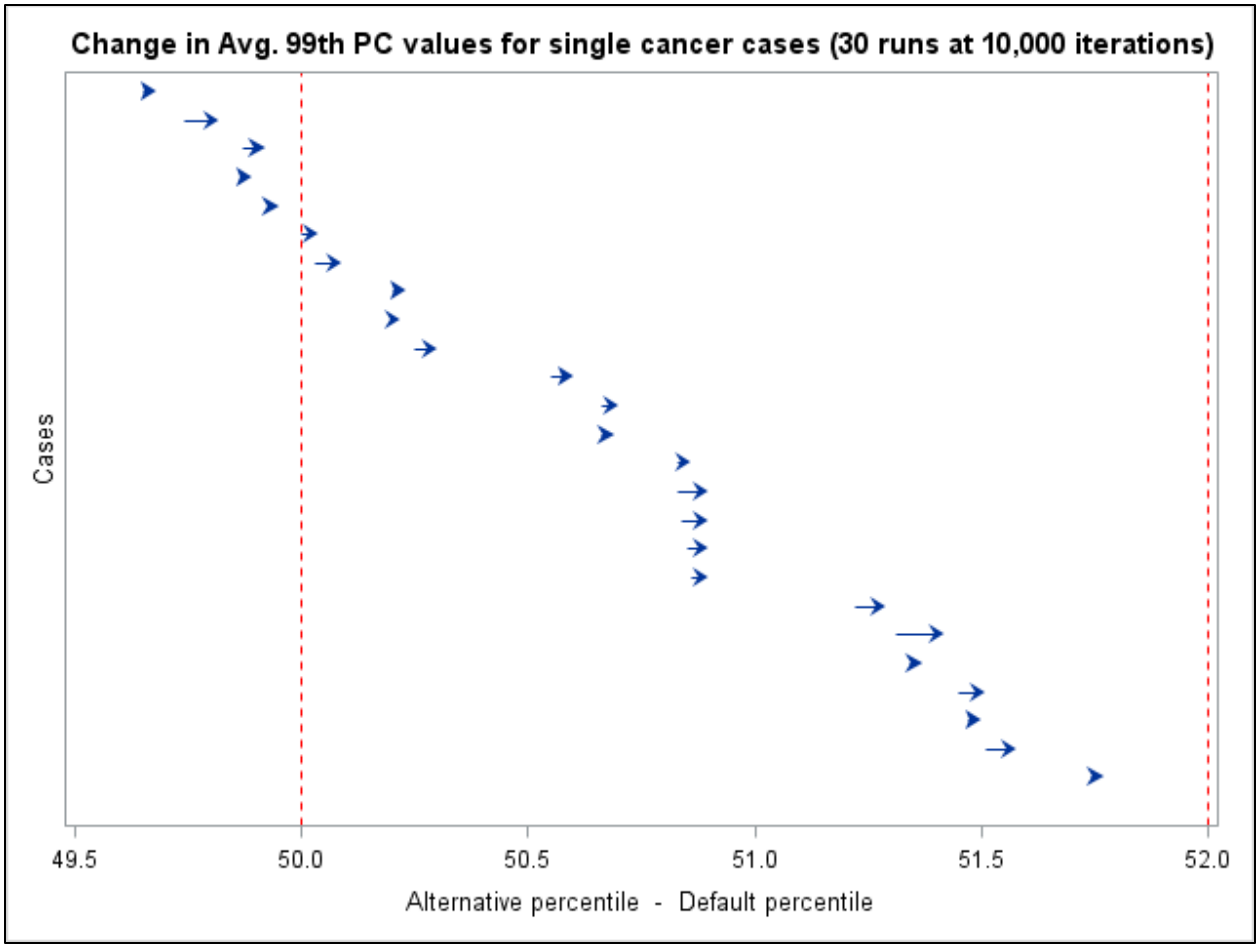

Figure 13. The increase in the Avg. 99th PC values from using the Default percentile versus using the Alternative percentile, for the 25 cases with Avg. 99th PC values between 50.00% and 51.99%.

Figure 14 shows the magnitude of the increase in the Avg. 99th PC value from using the Default percentile versus using the Alternative percentile for all 25 cases, sorted from the lowest increase to the highest increase. The plot also shows the IREP Cancer Model for each of the 25 cases.

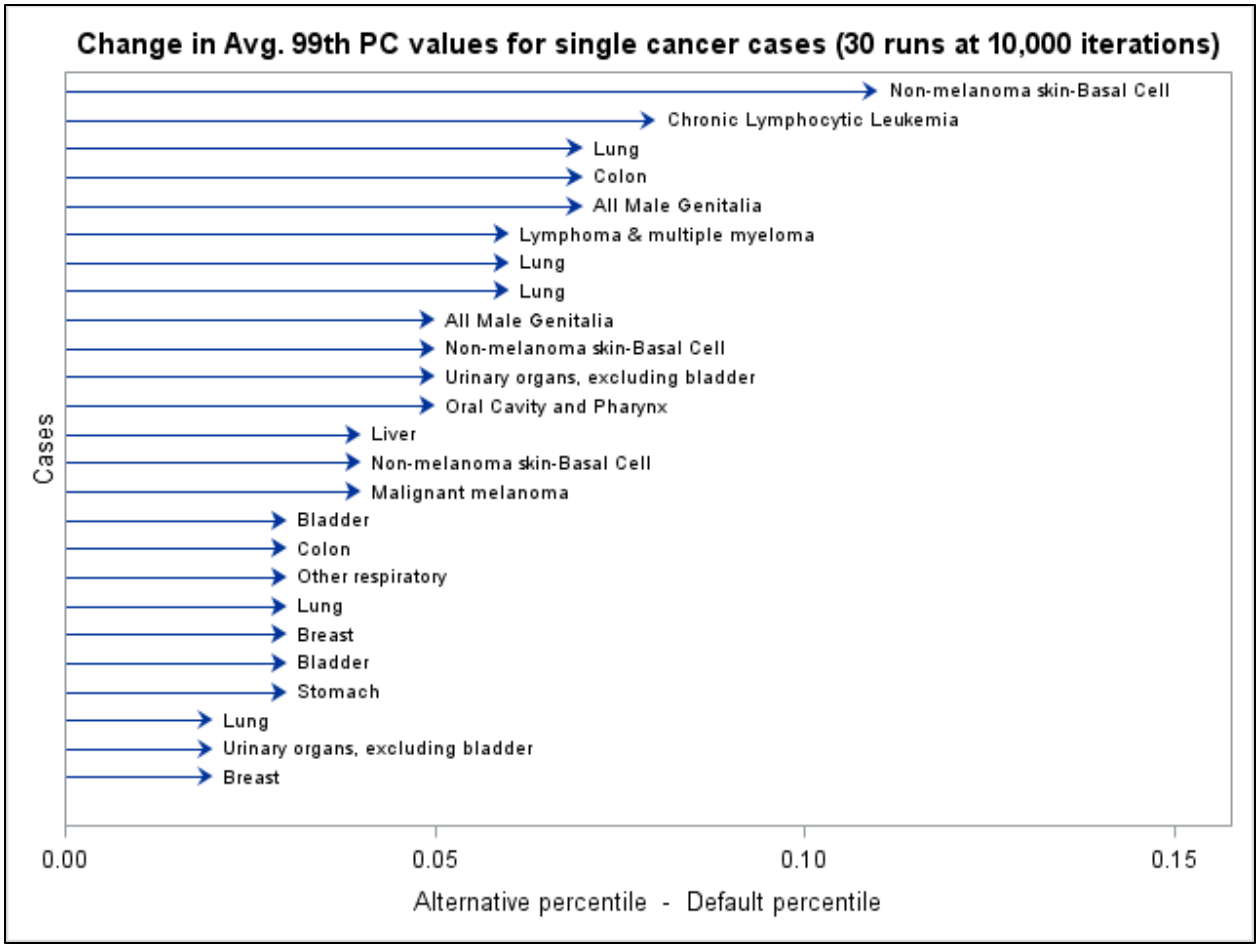

Figure 14. The increase in the Avg. 99th PC value from using the Default percentile versus using the Alternative percentile, for the 25 cases with Avg. 99th PC values between 50.00% and 51.99%.

Figure 15 shows a needle plot with the increase in the Avg. 99th PC value from using the Default percentile to using the Alternative percentile at 10,000, 20,000, and 30,000 iterations for all 25 cases. The plot also shows the reference lines corresponding to the average difference between the Alternative percentile and the Default percentile at 10,000, 20,000, and 30,000 iterations. While the differences in the Avg. 99th PC values from the Default percentile to the Alternative percentile at 10,000 iterations are mostly strictly greater than the differences at 20,000 iterations, that's not necessarily true for the differences at 20,000 versus 30,000 iterations, with multiple cases having the same difference in the Avg. 99th PC at 20,000 and 30,000 iterations.

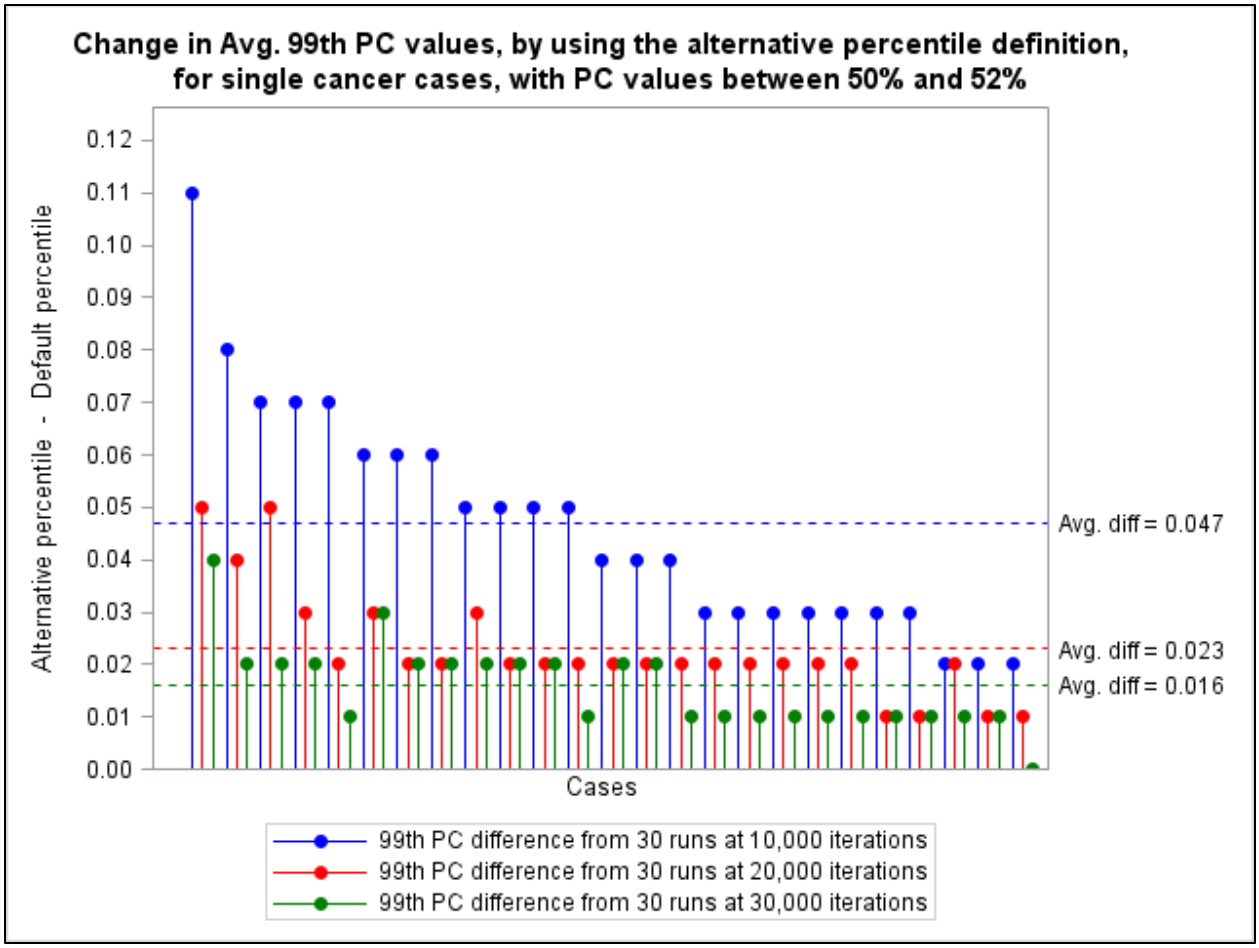

Figure 15. Increase in the Avg. 99th PC values from the Default percentile to the Alternative percentile at 10,000, 20,000, and 30,000 iterations, for the 25 cases with Avg. 99th PC values between 50.00% and 51.99%.

4.0 IREP ENTERPRISE EDITION RUNNING TIMES

Since increasing the number of iterations from 10,000 iterations to 20,000 or 30,000 iterations would require more computing time, the IREP EE running times obtained for each number of iterations was recorded for all 72 single cancer cases with PC values between 45.00% and 51.99%. Figure 16 shows a plot with the running times at 10,000, 20,000, and 30,000 iterations versus the number of exposures for all 72 cases with Avg. 99th PC values between 45.00% and 51.99%. The largest number of exposures for the 72 cases is 724 exposures. The filled circles in the plot correspond to cases with Lung cancers, which usually take longer to complete due to the two lung cancer models that are run in parallel. While these running times were obtained using a research version of IREP EE, which is faster than the public IREP EE version, it is expected that the running times will be up to one order of magnitude higher when the cases will be run on the public IREP EE version.

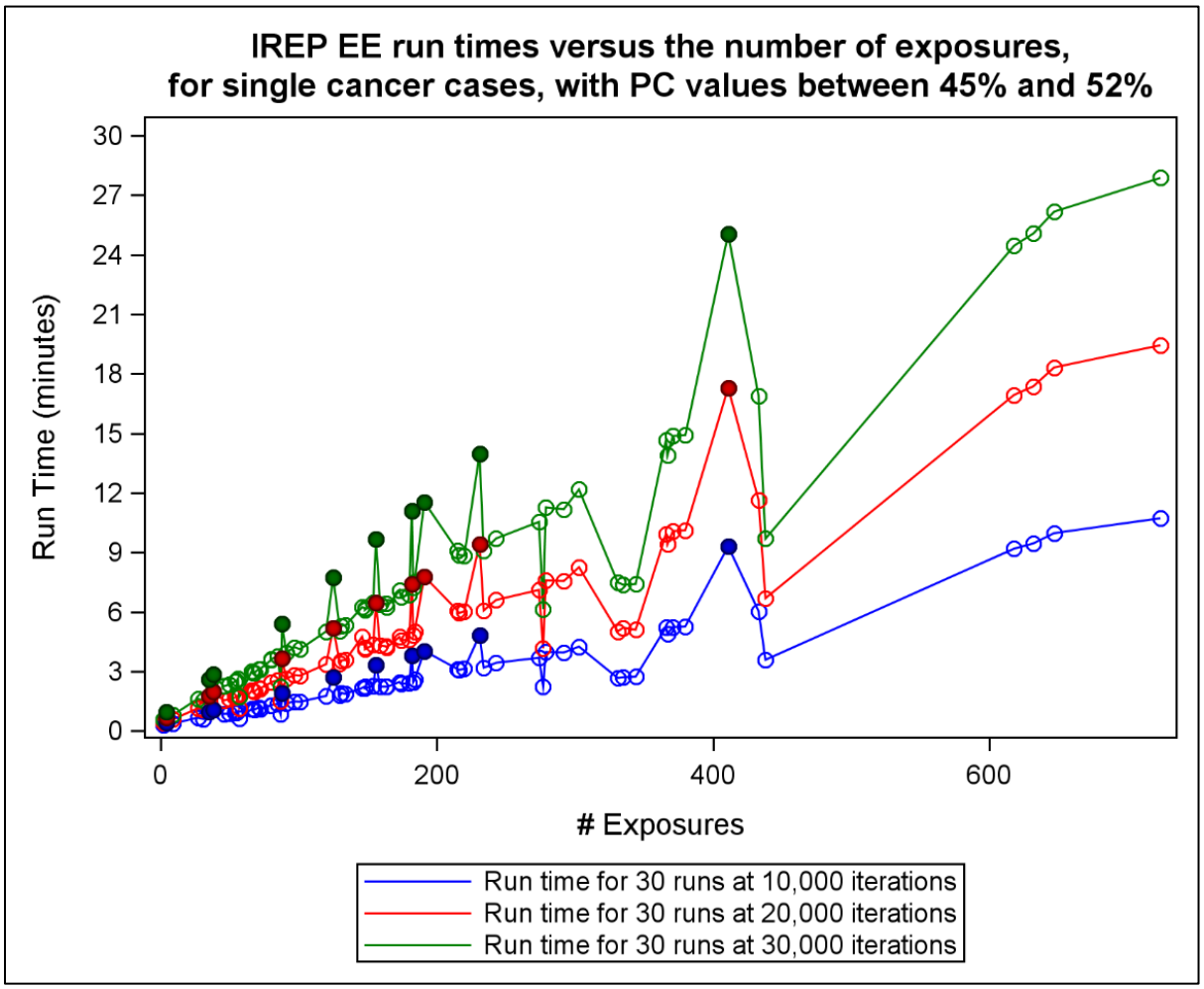

Figure 16. Run times (in minutes) at 10,000, 20,000, and 30,000 iterations versus the number of exposures for the 72 cases with Avg. 99th PC values between 45.00% and 51.99%.

5.0 PC VARIABILITY AT DIFFERENT NUMBERS OF ITERATIONS

Each case in NOCTS is run in IREP using 2,000 iterations and a default random seed of 99, and if the resulting 99th PC value is between 45.00% and 51.99%, the case is rerun on IREP EE with 30 runs at 10,000 iterations, and each run having a different random seed. Estimates of 99th percentile of PC obtained at 2,000 iterations can vary significantly depending on the random seed used. To better determine the variability of 99th PC values obtained at 2,000 iterations, an analysis was conducted for all cases with the 99th PC values between 45.00% and 51.99%. The cases selected for this analysis include the 25 single cancer cases with PC values between 45.00% and 48.99%, the 22 single cancer cases with PC values between 49.00% and 49.99%, the 25 single cancer cases with PC values between 50.00% and 51.99%, and an additional 5 cases from the category of multiple cancer cases with PC values between 49.00% and 49.99% (for each of these 5 cases, only one cancer with the PC value between 45.00% and 48.99% was selected). Altogether, a total of 77 cases with 99th PC values between 45.00% and 51.99% were selected for this analysis, and each of the 77 cases was run on IREP EE with 2,000, 10,000, 20,000, and 30,000 iterations.

Table 6 shows statistics for the standard deviations and observed ranges of the individual 99th PC values from each set of 30 runs, for all 77 cases with PC values between 45.00% and 51.99% and all number of iterations used. There is a lot of variability in the individual 99th PC values obtained from each set of 30 runs, with a considerable higher variability at 2,000 iterations, and decreasing variability as the number of iterations increases to 10,000, 20,000, and 30,000 iterations. The results shown in Table 6 are for the 99th PC values computed using the Default percentile, and the 99th PC values computed using the Alternative percentile show a similar variability.

Table 6. Statistics showing the variability of the individual 99th PC values for the 77 cases with PC values between 45% and 52%.

Number of Iterations	Statistic	N	Minimum	Maximum
2,000	Std. Dev.	77	0.80	3.56
2,000	Range	77	2.99	15.77
10,000	Std. Dev.	77	0.40	1.79
10,000	Range	77	1.78	8.24
20,000	Std. Dev	77	0.27	1.14
20,000	Range	77	0.93	4.54
30,000	Std. Dev	77	0.27	1.14
30,000	Range	77	0.98	5.11

For each set of 30 runs performed for a case at any number of iterations, the individual 99th PC values are approximately normally distributed, with the mean given by the Avg. 99th PC value.

This is shown in Figures 17 and 18, for two of the tested cases. In both Figure 17 (20,000 iterations) and in Figure 18 (2,000 iterations), the individual 99th PC values are normally distributed, as it can be seen from the q-q plot, as well as from the normality test p-value. Most of the tested cases show that the individual 99th PC values obtained at all iterations are normally distributed or can be considered approximately normally distributed. The cases for which the individual 99th PC values are only approximately normally distributed are usually obtained at a lower number of iterations (2,000 iterations, as shown in Figure 18), since one or a few outliers create a slight departure from normality.

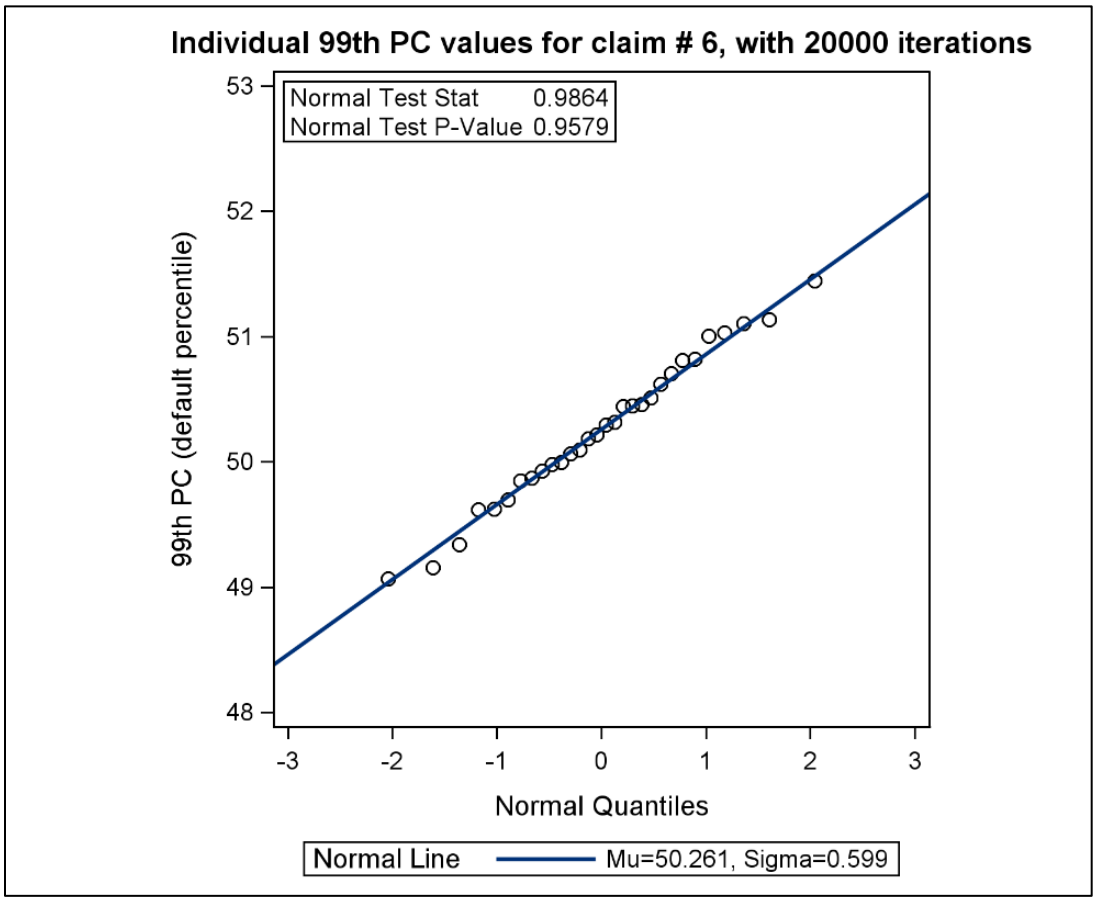

Figure 17: Normality of the individual 99th PC values obtained at 20,000 iterations, for a case with Avg. 99th PC value between 50% and 52%.

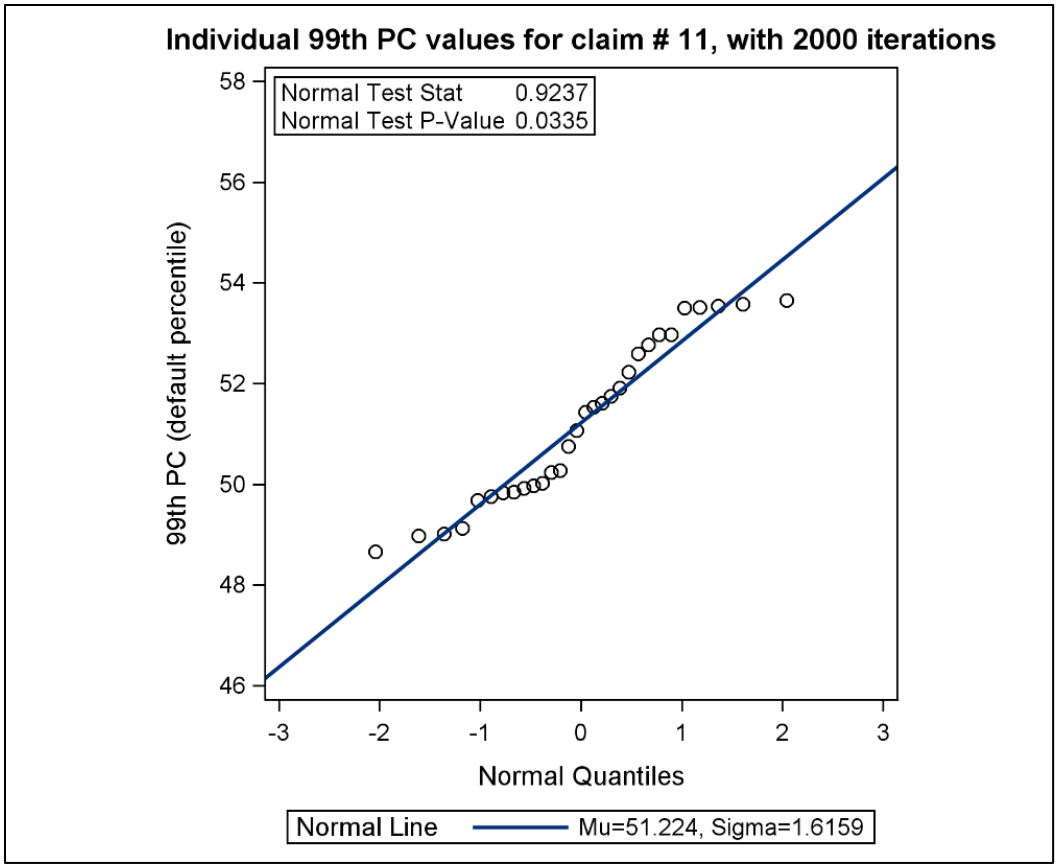

Figure 18. Approximate normality of the individual 99th PC values obtained at 2,000 iterations, for a case with Avg. 99th PC value between 50% and 52%.

Figure 19 shows the fitted normal distributions of the individual 99th PC values at 2,000, 10,000, 20,000, and 30,000 iterations for one case with the Avg. 99th PC value at 30 runs and 10,000 iterations selected between 45.00% and 48.99% (the Avg. 99th PC value on record in NOCTS for this case is 46.99%). The Avg. 99th PC values for each number of iterations are shown as vertical lines in the figure. This example shows that there is a high probability that one 99th PC value at 2,000 iterations could have been easily generated below 45%, making this case not eligible for a run on IREP EE (which would had been clearly an incorrect decision, even though the compensability would had not been affected in this case).

Figure 20 shows the fitted normal distributions of the individual 99th PC values at 2,000, 10,000, 20,000, and 30,000 iterations for one case with the Avg. 99th PC value at 30 runs and 10,000 iterations selected between 49.00% and 49.99% (the Avg. 99th PC value on record in NOCTS for this case is 49.87%). This example shows that there is a high probability that one 99th PC value at 2,000 iterations could have been easily generated above 52%, making this case not eligible for a run on IREP EE (which would had been clearly an incorrect decision, and the compensability would had been incorrectly decided for this case).

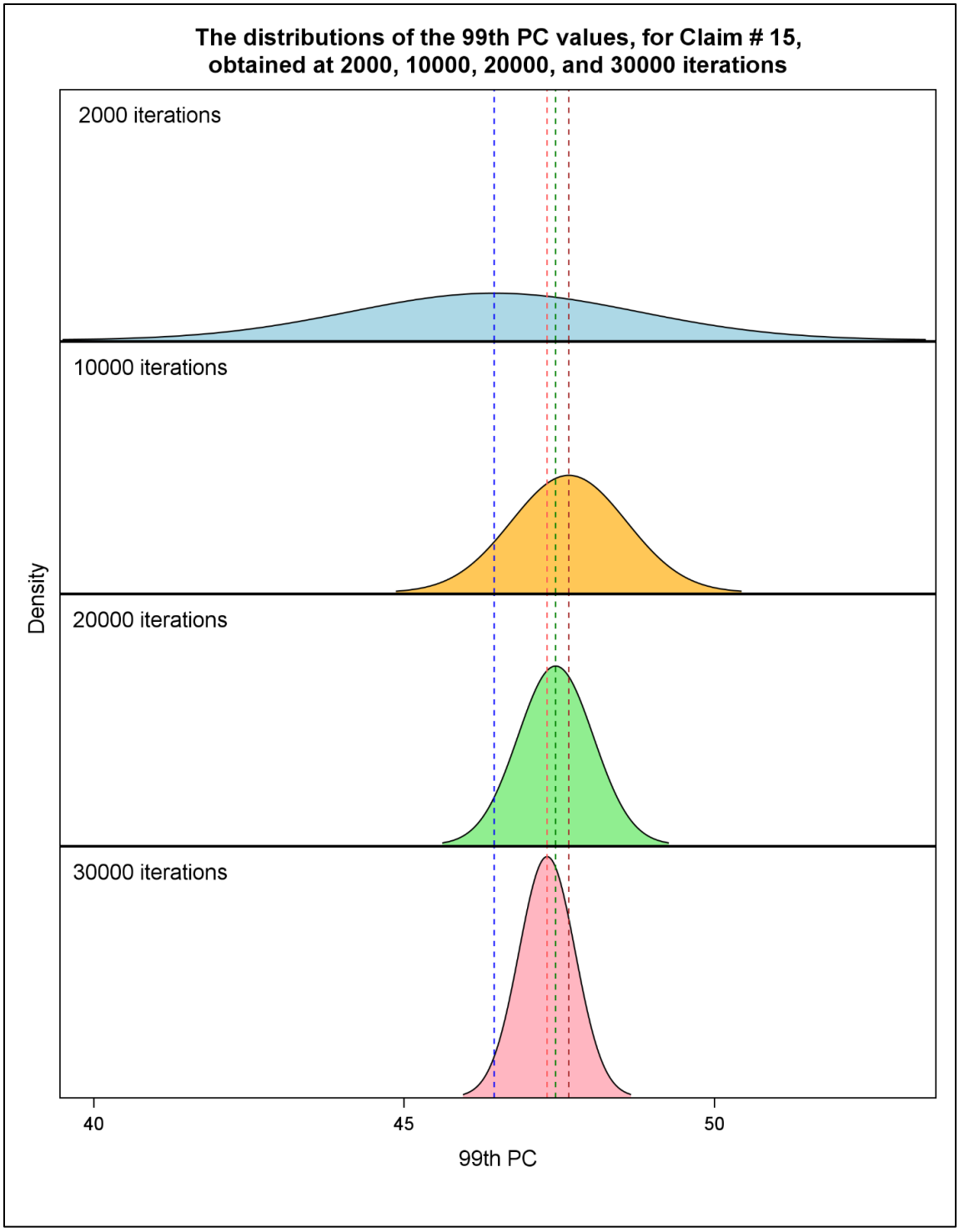

Figure 19. Fitted normal distributions for the individual 99th PC values at 2,000, 10,000, 20,000, and 30,000 iterations for a case with Avg. 99th PC value (obtained at 30 runs and 10,000 iterations) between 45% and 49%.

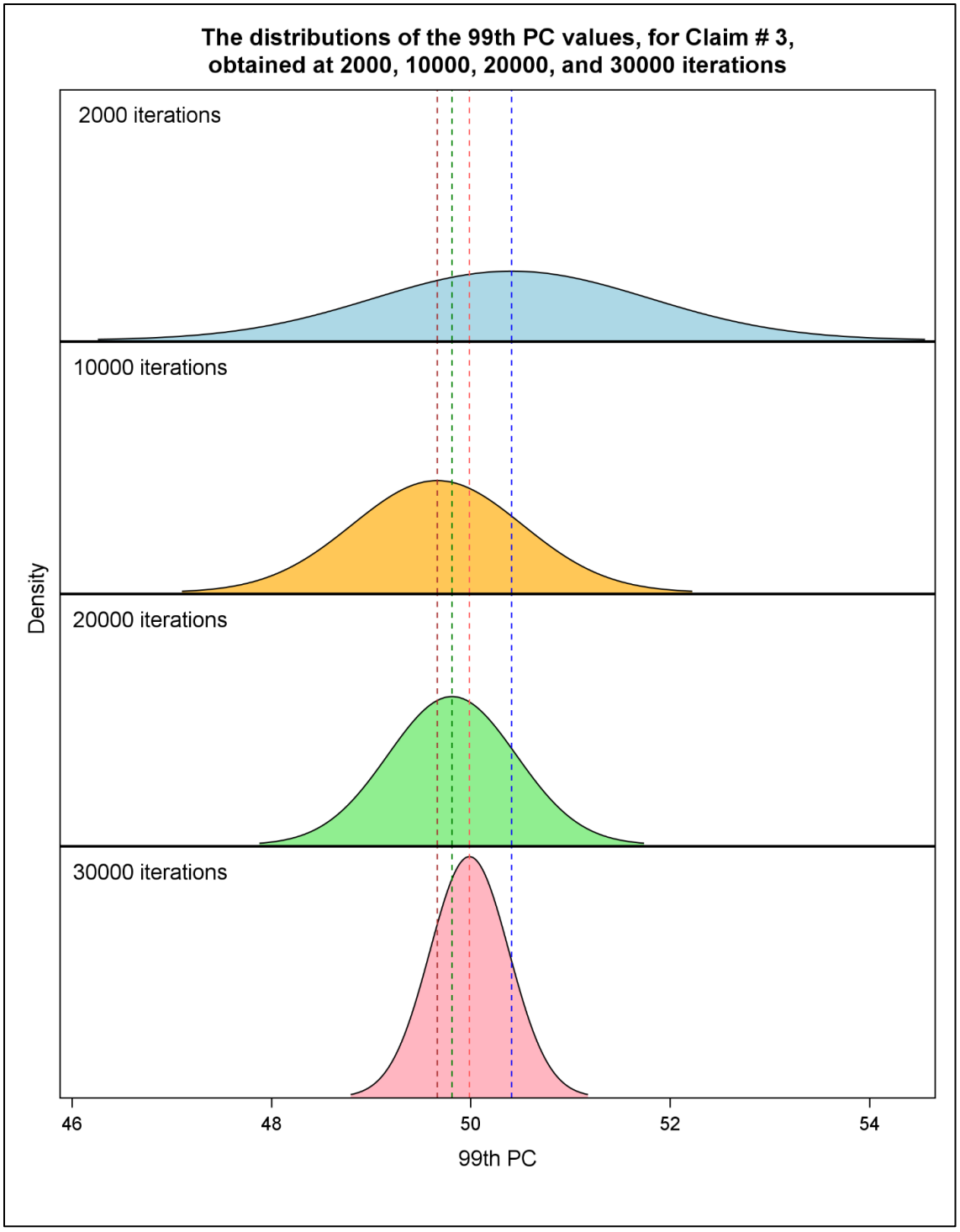

Figure 20. Fitted normal distributions for the individual 99th PC values at 2,000, 10,000, 20,000, and 30,000 iterations for a case with Avg. 99th PC value (obtained at 30 runs and 10,000 iterations) between 49% and 50%.

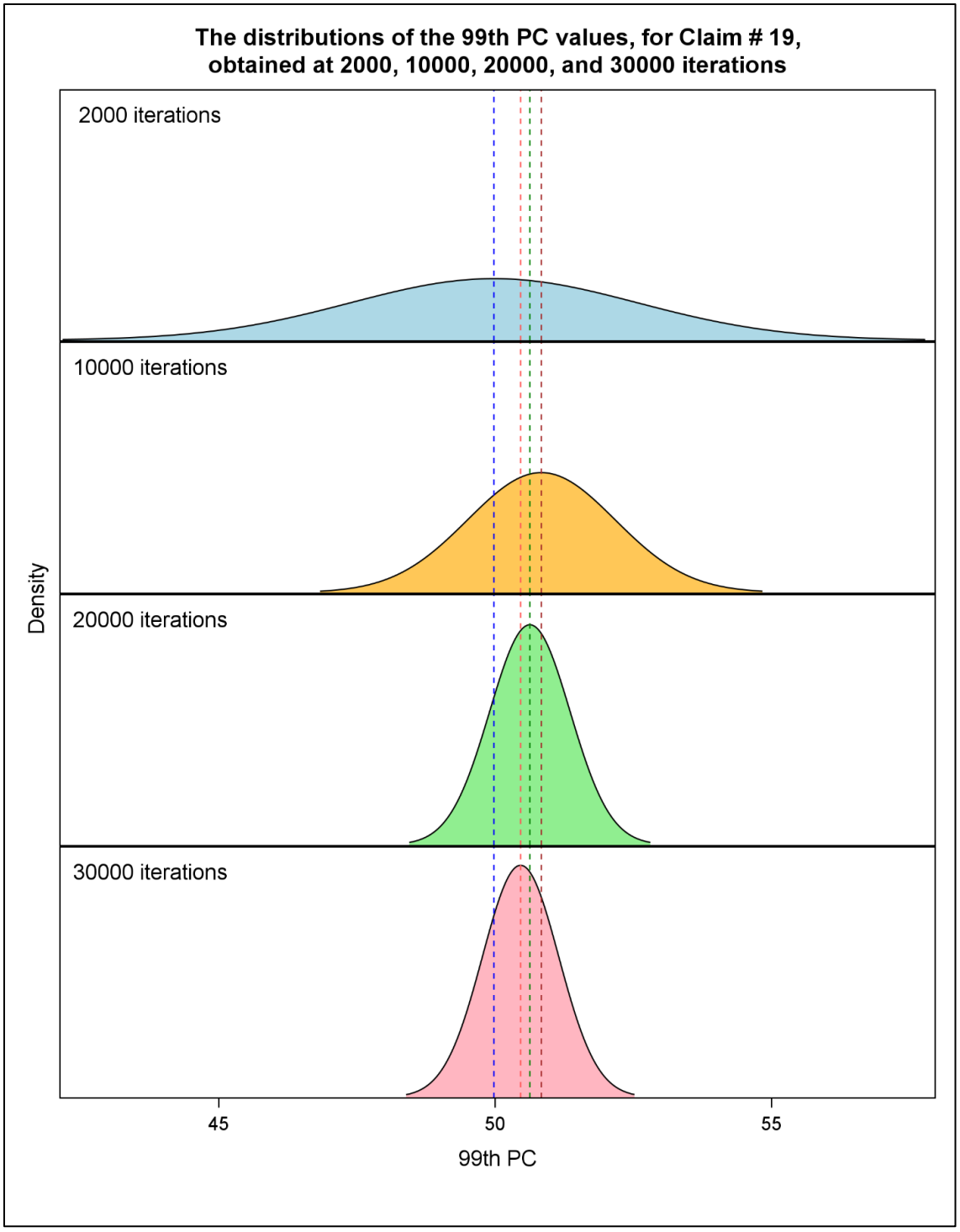

Figure 21. Fitted normal distributions for the individual 99th PC values at 2,000, 10,000, 20,000, and 30,000 iterations for a case with Avg. 99th PC value (obtained at 30 runs and 10,000 iterations) between 50% and 52%.

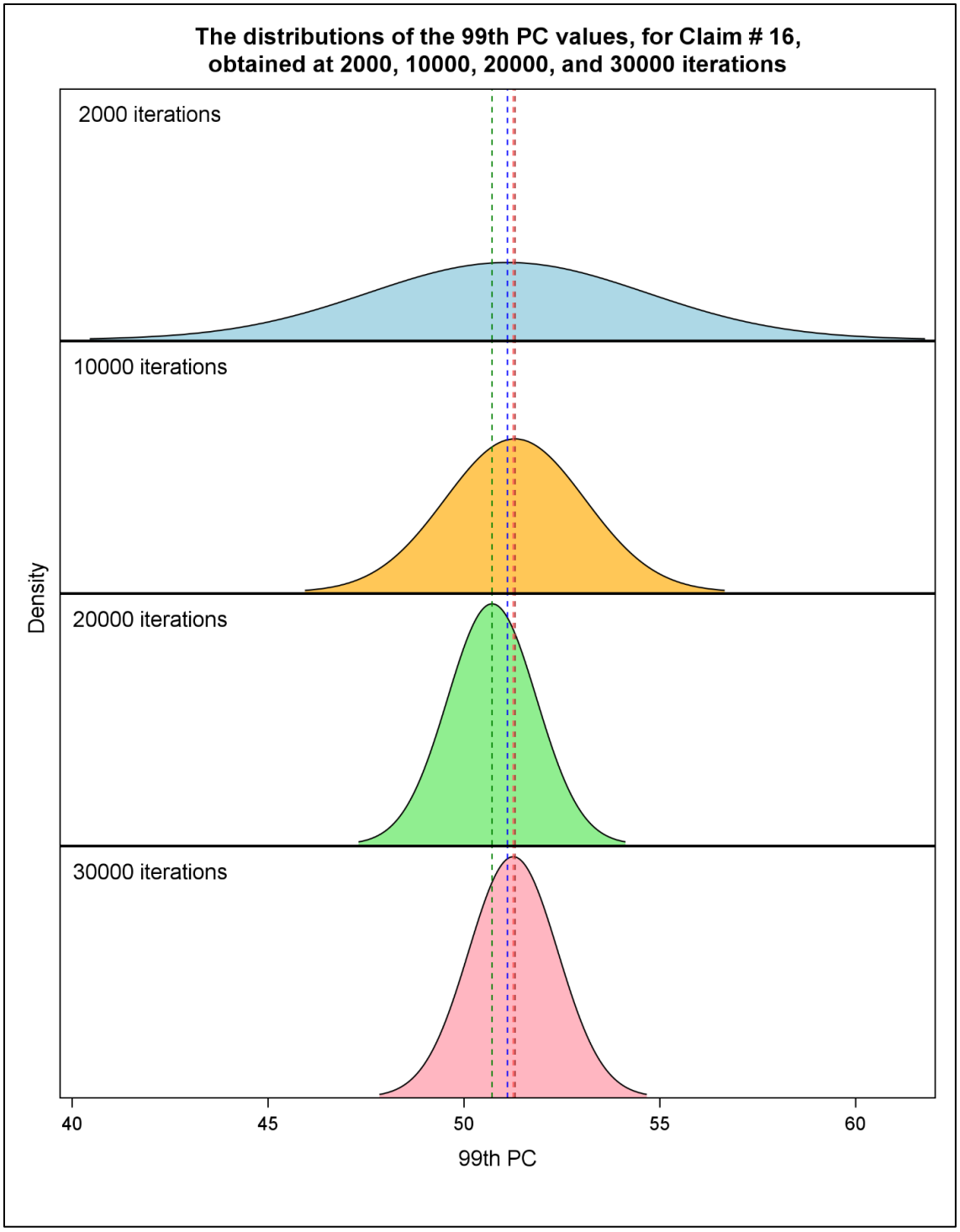

Figure 22. Fitted normal distributions for the individual 99th PC values at 2,000, 10,000, 20,000, and 30,000 iterations for a case with Avg. 99th PC value (obtained at 30 runs and 10,000 iterations) between 50% and 52%.

Figure 21 shows the fitted normal distributions of the individual 99th PC values at 2,000, 10,000, 20,000, and 30,000 iterations for one case with the Avg. 99th PC value at 30 runs and 10,000 iterations selected between 50.00% and 51.99% (the Avg. 99th PC value on record in NOCTS for this case is 50.47%). This example shows that there is a small probability that one 99th PC value at 2,000 iterations could have been generated below 45%, making this case not eligible for a run on IREP EE (which would had been clearly an incorrect decision, a decision that would had been incorrectly affected the compensability for this case).

Figure 22 shows the fitted normal distributions of the individual 99th PC values at 2,000, 10,000, 20,000, and 30,000 iterations for one case with the Avg. 99th PC value at 30 runs and 10,000 iterations selected between 50.00% and 51.99% (the Avg. 99th PC value on record in NOCTS for this case is 51.30%). This example has the largest range for the 99th PC values obtained at 2,000 iterations from all the 77 cases tested, with an observed range of 15.8, and a computed range (estimated based on the 99.7% confidence interval of the fitted normal distribution) of 21.3. This example shows that there is a large probability that one 99th PC value at 2,000 iterations could have been easily generated either below 45% or above 52%, making this case not eligible for a run on IREP EE (which would had been incorrectly affected the compensability in the first scenario for this case).

Table 6 and Figures 19, 20, 21, and 22 show that there can be a large range of values for the individual 99th PC values, especially for the individual 99th PC values obtained at 2,000 iterations. Based on the approximate normal distributions fitted to the 30 individual 99th PC values, the computed 99.7% confidence intervals for the individual 99th PC values (corresponding to 3 times the standard deviations around the Avg. 99th PC value) are actually larger than the observed ranges shown in Table 6, and are shown in Table 7; based on the 99.7% confidence intervals, the maximum observed range is 21.3 at 2,000 iterations, 10.7 at 10,000 iterations, 6.8 at 20,000 and 30,000 iterations.

Table 7. Maximum observed and computed ranges of the individual 99th PC values for cases with PC values between 45% and 52%.

Number of Iterations	Statistic	N	Maximum
2,000	Observed Range	77	15.8
2,000	Computed Range	77	21.3
10,000	Observed Range	77	8.2
10,000	Computed Range	77	10.7
20,000	Observed Range	77	4.5
20,000	Computed Range	77	6.8
30,000	Observed Range	77	5.1
30,000	Computed Range	77	6.8

These ranges were determined based on the 30 individual 99th PC values obtained for each case and is expected that they will vary if additional sets of 30 random seeds were used for each case.

While the individual 99th PC values have a very large variability, the Avg. 99th PC values have a much smaller variability, as it can be seen from the standard errors of the Avg. 99th PC values in Table 8. These standard errors of the Avg. 99th PC values show the accuracy of the Avg. 99th PC value obtained from a sample of 30 values (corresponding to the 30 random seeds used for each case) compared with the theoretical 99th PC value.

Table 8. Variability of the Std. Errors of the Avg. 99th PC values for cases with PC values between 45.00% and 51.99%.

Number of Iterations	Statistic	N	Minimum	Median	Mean	Maximum
2,000	Std. Error	77	0.15	0.29	0.31	0.65
10,000	Std. Error	77	0.07	0.14	0.15	0.33
20,000	Std. Error	77	0.05	0.10	0.10	0.21
30,000	Std. Error	77	0.05	0.08	0.08	0.21

Figures 23, 24, and 25 show the Std. Errors of the Avg. 99th PC values at different numbers of iterations, for the 30 cases with PC values between 45.00% and 48.99%, for the 22 cases with PC values between 49.00% and 49.99%, and respectively for the 25 cases with PC values between 50.00% and 51.99%. Each of the three figures shows a significant decrease in the Std. Errors of the Avg. 99th PC values as the number of iterations increases from 2,000 to 10,000 iterations, but only modest decreases from 20,000 to 30,000 iterations.

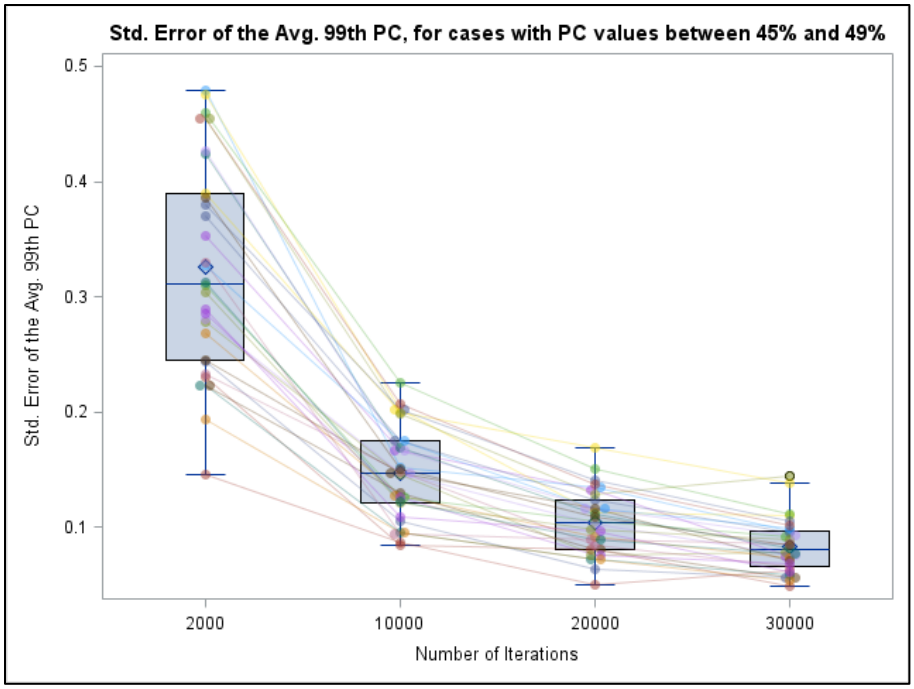

Figure 23. Std. Errors of the Avg. 99th PC values at different number of iterations, for 30 cases with PC values between 45.00% and 48.99%.

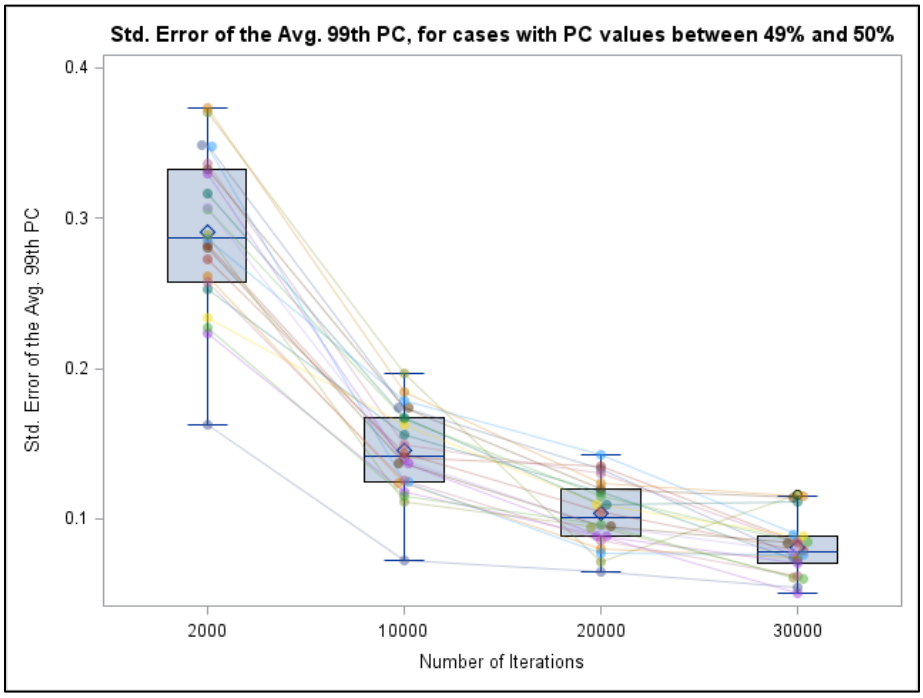

Figure 24: Std. Errors of the Avg. 99th PC values at different number of iterations, for 22 cases with PC values between 49.00% and 49.99%.

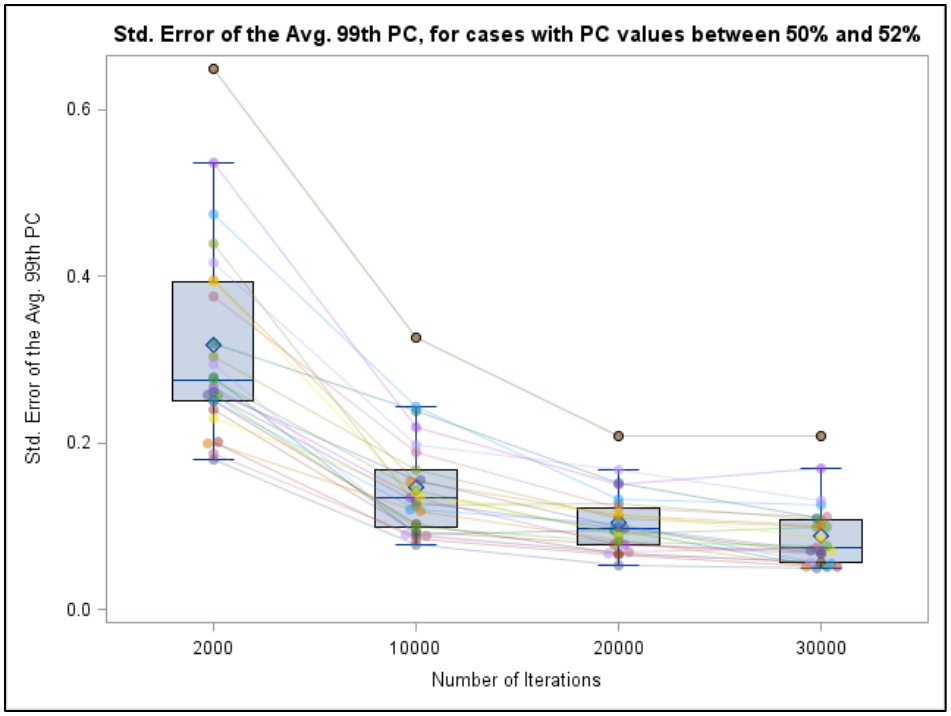

Figure 25. Std. Errors of the Avg. 99th PC values at different number of iterations, for 25 cases with PC values between 50.00% and 51.99%.

6.0 RANGES TO SELECT CASES TO BE RUN ON IREP ENTERPRISE EDITION

Table 8 and Figures 23, 24, and 25 clearly show that there is a big advantage in using the Avg. 99th PC values instead of the individual 99th PC values, due to the lower variability and the more stable values that are generated in the sampling process. To better estimate the variability of the Avg. 99th PC values for an individual case, a bootstrap analysis was conducted for 25 selected cases. The 25 cases selected for this analysis were a subset of the 77 cases used in the previous analysis and include 14 cases with PC values around the 50% compensation threshold, as well as 11 cases with the largest standard deviations at 2,000, 10,000, and 20,000 iterations. For these 25 cases, in addition to the original set of 30 random seeds used (denoted as Set 1), an additional three sets of 30 random seeds were generated (denoted as Set 2, Set 3, and Set 4), and each of the 25 cases were run on the IREP Enterprise Edition for each additional three sets of random seeds. For each of the 25 cases, the Avg. 99th PC values obtained from these three additional sets of 30 random seeds were obtained using the IREP Enterprise Edition at 2,000, 10,000, and 20,000 iterations, and the Avg. 99th PC values from the combined 120 random seeds used for each case were also computed. The four Avg. 99th PC values obtained for each case from the four random seeds sets will vary around the theoretical 99th PC value for each case, and the Avg. 99th PC values obtained from the combined 120 random seeds represents the best estimate of the theoretical 99th PC value for each case (based on the four sets of 30 random seeds used for each case). A bootstrap analysis was conducted to determine the sampling distribution of the Avg. 99th PC values obtained based on 30 runs, by repeatedly sampling with replacement from the 120 previously generated 99th PC values for each case. Figure 26 show the sampling distribution of the Avg. 99th PC value (based on 30 runs and 20,000 iterations) obtained using the bootstrap analysis (with 5,000 resamples) for a case having the Avg. 99th PC value on record in NOCTS of 49.84%. The standard deviation of the sampling distribution of the Avg. 99th PC values represents the standard error (SE) of the Avg. 99th PC value, and the standard error values obtained through the bootstrap analysis (by repeatedly resampling from the 99th PC values obtained from the 4 sets of 30 random seeds) are more accurate than the previously obtained standard errors (which were obtained based on only one set of 30 random seeds).

As it can be seen in Figure 26, the bootstrap Avg. 99th PC values from 30 runs are normally distributed around the Avg. 99th PC value from 120 runs, and this normal distribution provides us with the variability of the Avg. 99th PC values from 30 runs around the theoretical 99th PC value for this case. For this case, the standard errors of the Avg. 99th PC value obtained from the four sets of 30 random seeds are 0.14, 0.11, 0.10, and 0.09, and the standard error obtained through the bootstrap analysis is 0.11.

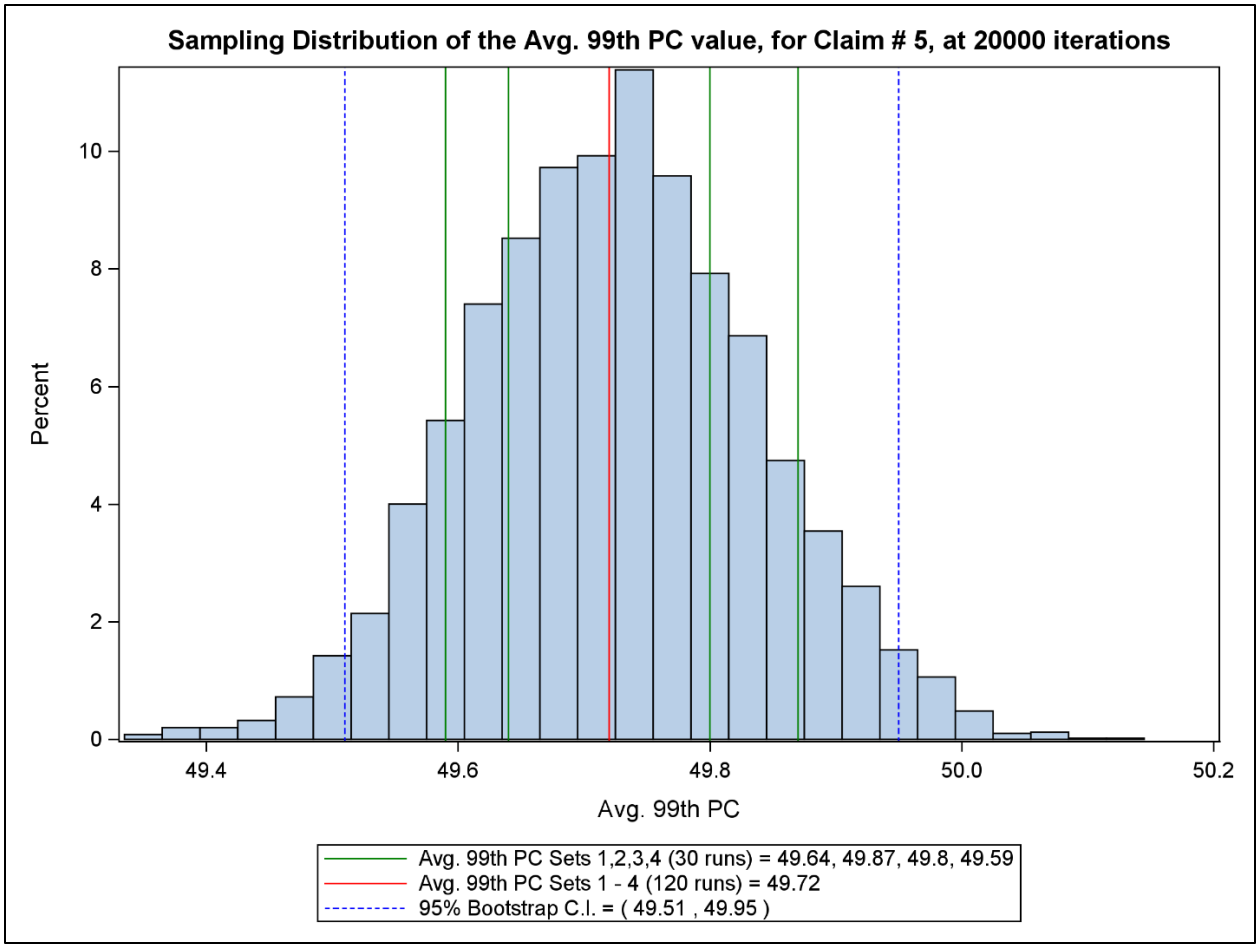

Figure 26. Sampling distribution of the Avg. 99th PC value at 20,000 iterations, for a case with Avg. 99th PC value between 49.00% and 49.99%.

The examples shown in Figures 19, 20, 21, and 22 showed the peril of choosing one individual 99th PC value in determining a case outcome, and the fact that one individual 99th PC value obtained at 2,000 iterations (as currently used the IREP version) might not be a very reliable indicator to choose which cases are run on the IREP Enterprise Edition.

The current rule in use is that if the individual 99th PC value obtained at 2,000 iterations and default random seed is between 45.00% and 51.99%, corresponding to a range of (-5.0, 2.0) around the compensation threshold of 50.00%, then the case is re-run on IREP Enterprise Edition with 30 runs at 10,000 iterations.

Based on the largest standard error of the Avg. 99th PC values at 2,000 iterations, which show how close the Avg. 99th PC value at 2,000 iterations is relative to the true but unknown 99th PC population value (a value that can be narrowed down to a very small interval as the number of iterations is increased), and based on the largest standard deviation of the individual 99th PC

values at 2,000 iterations around the Avg. 99th PC value at 2,000 iterations (obtained through the bootstrap analysis), a confidence interval (C.I.) can be determined for an individual randomly generated 99th PC value around the true 99th PC population value. The confidence interval for an individual randomly generated 99th PC value around the true 99th PC population value can be computed based on adding the two standard deviations and multiplying that value by 3 (corresponding to a 99.7% C.I. in each of the two normal distributions) or based on adding the two standard deviations and multiplying that value by 2 (corresponding to a 95% C.I. in each of the two normal distributions). Figures 27 and 28 exemplify the way the normal distributions are used to obtain the confidence interval for an individual randomly generated 99th PC value around the true 99th PC population value.

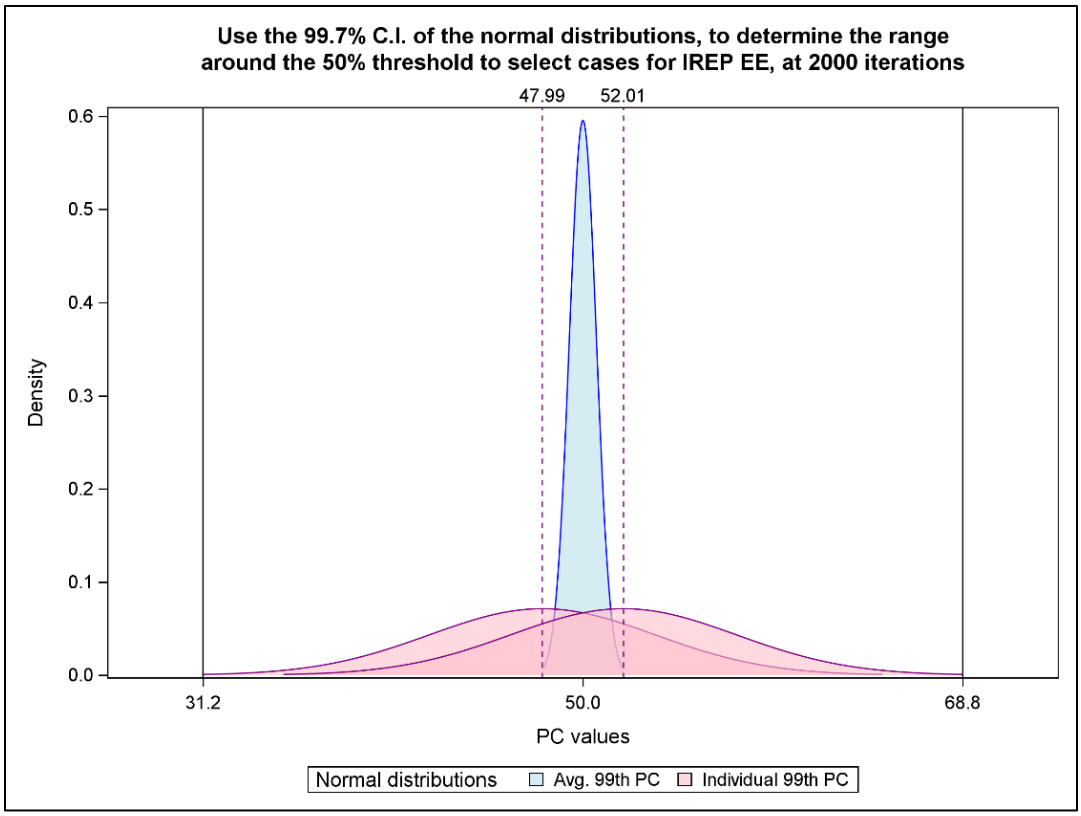

Figure 27. Use of the 99.7% C.I. of the normal distributions to determine the range for an individual randomly generated 99th PC value, to be used to select cases for an IREP EE run.

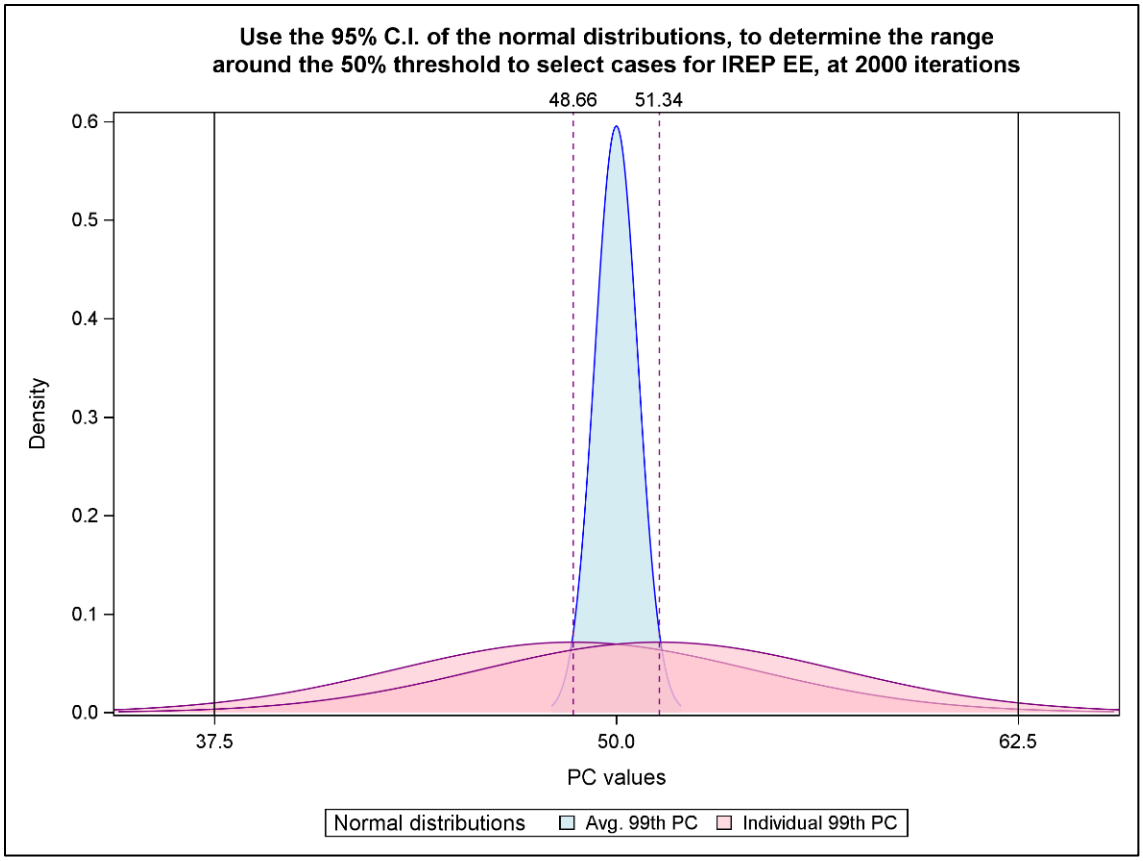

Figure 28. Use of the 95% C.I. of the normal distributions to determine the range for an individual randomly generated 99th PC value, to be used to select cases for an IREP EE run.

Another method to estimate the confidence interval for an individual randomly generated 99th PC value around the true 99th PC population value is to use the observed minimum and maximum 99th PC values from the 120 runs for the case with the largest range of individual 99th PC values. The ranges obtained using this method are not symmetrical, but it is expected that they would become more symmetrical if additional test cases would be included in the analysis.

Using each of the three computations methods described above, the confidence intervals around the compensation threshold of 50.00% to select a case to be run on the IREP Enterprise Edition, based on an individual randomly generated 99th PC value obtained at 2,000 iterations are shown in Table 9. Similar ranges can also be determined at 10,000 iterations and 20,000 iterations and they are all are summarized in Table 9. Since one outlier case (a case having some of the dose exposures modeled with lognormal distributions having GSD values close to 15) is responsible for the largest ranges, and since most of the cases don't have standard deviations as large as the outlier case, alternative ranges based on the 76 cases and obtained after the outlier case is excluded are also presented in Table 10.

Table 9. Ranges around the 50% threshold to be used in selecting which cases need to be run in the IREP Enterprise Edition, based on the 99th PC value obtained using the IREP version, computed based on data from 77 cases.

Number of iterations used in IREP	Number of iterations used in IREP EE	Method used to determine the Range around the 50% threshold to select cases to be run in IREP EE, based on the PC obtained in IREP	Range
2,000	10,000	99.7% C.I. of the normal distributions	(-18.8, 18.8)
2,000	10,000	95% C.I. of the normal distributions	(-12.5, 12.5)
2,000	10,000	Observed min and max 99th PC values	(-9.7, 10.9)
10,000	20,000	99.7% C.I. of the normal distributions	(-8.4, 8.4)
10,000	20,000	95% C.I. of the normal distributions	(-5.6, 5.6)
10,000	20,000	Observed min and max 99th PC values	(-4.0, 4.7)
20,000	30,000	99.7% C.I. of the normal distributions	(-5.6, 5.6)
20,000	30,000	95% C.I. of the normal distributions	(-3.7, 3.7)
20,000	30,000	Observed min and max 99th PC values	(-3.0, 2.5)

Table 10. Ranges around the 50% threshold to be used in selecting which cases need to be run in the IREP Enterprise Edition, based on the 99th PC value obtained using the IREP version, computed based on data from 76 cases.

Number of iterations used in IREP	Number of iterations used in IREP EE	Method used to determine the Range around the 50% threshold to select cases to be run in IREP EE, based on the PC obtained in IREP	Range
2,000	10,000	99.7% C.I. of the normal distributions	(-13.2, 13.2)
2,000	10,000	95% C.I. of the normal distributions	(-8.8, 8.8)
2,000	10,000	Observed min and max 99th PC values	(-7.0, 7.3)
10,000	20,000	99.7% C.I. of the normal distributions	(-6.6, 6.6)
10,000	20,000	95% C.I. of the normal distributions	(-4.4, 4.4)
10,000	20,000	Observed min and max 99th PC values	(-3.7, 4.6)
20,000	30,000	99.7% C.I. of the normal distributions	(-4.3, 4.3)
20,000	30,000	95% C.I. of the normal distributions	(-2.8, 2.8)
20,000	30,000	Observed min and max 99th PC values	(-2.4, 2.3)

The values in the columns with the number of iterations used in IREP EE in Tables 9 and 10 show the number of iterations that is currently used in IREP EE when a case is run on IREP at 2,000 iterations and are provided as a suggestion for the number of iterations that should be used if the number of iterations in IREP would be increased to either 10,000 or 20,000 iterations.

So, based on this analysis and using the confidence intervals obtained based on the 95% C.I. of the normal distributions, the range at 2,000 iterations around the 50% threshold to be used in selecting which cases need to be run in the IREP Enterprise Edition would need to be increased to (-12.5, 12.5), which is much wider than the current range of (-5.0, 2.0) around the compensation threshold of 50.00%. Also, if the 99th PC value are obtained using the IREP version at 10,000 iterations, then the range used should be approximately (-5.6, 5.6), and if the 99th PC value are obtained using the IREP version at 20,000 iterations, then the range used should be approximately (-3.7, 3.7). Based on the analysis with 76 cases (after excluding the outlier case), the range at 2,000 iterations around the 50% threshold to be used in selecting which cases need to be run in the IREP Enterprise Edition would need to be increased to (-8.8, 8.8), and if the 99th PC value are obtained using the IREP version at 10,000 iterations, then the range used should be approximately (-4.4, 4.4), and if the 99th PC value are obtained using the IREP version at 20,000 iterations, then the range used should be approximately (-2.8, 2.8).

7.0 CONCLUSIONS

An analysis was conducted to investigate the effect on the PC values of using an alternative percentile definition, instead of the default percentile definition implemented in the NIOSH-IREP software. The analysis was carried out using several categories of cases selected from NOCTS.

There are significant differences between 99th PC values obtained using the Default percentile currently used in the NIOSH-IREP software (computed using the Type 7 percentile definition) and those obtained using the Alternative percentile (computed using the Type 2 percentile definition), with the Alternative percentile always producing higher values than the Default percentile.

For the 45 tested cases with the 99th PC values between 44.00% and 44.99%, there is an average increase in the 99th PC values of 0.27 , with a maximum increase of 1.46 at 2,000 iterations, from using the Default percentile versus using the Alternative percentile, an average increase of 0.06, with a maximum increase of 0.27 at 10,000 iterations, and an average increase of 0.02, with a maximum increase of 0.17 at 20,000 iterations.

For the 72 tested cases with the 99th PC values between 45.00% and 51.99%, there is an average increase in the 99th PC values of 0.04, with a maximum increase of 0.11 at 10,000 iterations, from using the Default percentile versus using the Alternative percentile, an average increase of 0.02, with a maximum increase of 0.05 at 20,000 iterations, and an average increase of 0.02, with a maximum increase of 0.04 at 30,000 iterations.

To improve estimates of the 99th PC values, either the alternative percentile definition can be implemented in the future NIOSH-IREP versions, or the number of iterations used should be increased, in order to minimize the underestimation effect on the 99th PC values.

The number of iterations used in the IREP version should be increased at a minimum of 10,000 iterations, but preferably to 20,000 iterations or more. The number of iterations used in IREP Enterprise Edition should be respectively increased at a minimum of 20,000 iterations, but preferably to 30,000 iterations or more.

For cases with 99th PC values between 49.50% and 49.99%, a much larger number of iterations is recommended to reduce the variability and increase the accuracy of the Avg. 99th PC value. By increasing the number of runs in the IREP Enterprise Edition from 30 runs to 100 or more runs, and by increasing the number of iterations per run, the variability of the Avg. 99th PC around the true 99th PC value can be greatly reduced.

The ranges around the 50.0% threshold to be used in selecting cases which need to be run on the IREP Enterprise Edition, based on 99th PC value obtained using the IREP version, should be adjusted accordingly, and several ranges were provided depending on the number of iterations used in the IREP and in the IREP Enterprise Edition versions.

REFERENCES

Stancescu D., Apostoaei A. I., Thomas B.A., Comparison of Several Percentile Definitions, NIOSH, December 2019.

LIST OF TERMS, ABBREVIATIONS AND ACRONYMS

Distribution types

LN(GM, GSD)	Lognormal probability distribution defined by parameters GM and GSD
GM	Geometric Mean
GSD	Geometric Standard Deviation
N(mean, SD)	Normal probability distribution defined by parameters mean and SD
mean	Mean of a Normal (Gaussian) probability distribution
SD	Standard deviation

Other terms

99^{th} PC value	99^{th} percentile of probability of causation (PC)
Avg. 99^{th} PC value	Average of the 99^{th} percentiles of PC obtained in sets of 30, or 120 runs.
ABRWH	Advisory Board on Radiation and Worker Health
IREP	Interactive Radio-Epidemiological Program
DCAS	Division of Compensation Analysis and Support
NIOSH	National Institute of Occupational Safety and Health
NOCTS	NIOSH DCAS Claims Tracking System
PC	Probability of Causation
RB	Relative Bias
RRMSE	Relative Root-Mean Square Error