

RESEARCH ARTICLE

# Hypothesis Testing of Inclusion of the Tolerance Interval for the Assessment of Food Safety

Hungyen Chen<sup>1,2</sup>, Hirohisa Kishino<sup>2\*</sup>

**1** National Research Institute of Fisheries Science, Fisheries Research Agency, Kanagawa, 236–8648, Japan, **2** Graduate School of Agricultural and Life Sciences, The University of Tokyo, Tokyo, 113–8657, Japan

\* [kishino@lbn.ab.a.u-tokyo.ac.jp](mailto:kishino@lbn.ab.a.u-tokyo.ac.jp)



**OPEN ACCESS**

**Citation:** Chen H, Kishino H (2015) Hypothesis Testing of Inclusion of the Tolerance Interval for the Assessment of Food Safety. PLoS ONE 10(10): e0141117. doi:10.1371/journal.pone.0141117

**Editor:** Guangyuan He, Huazhong University of Science and Technology, CHINA

**Received:** March 26, 2015

**Accepted:** October 5, 2015

**Published:** October 28, 2015

**Copyright:** © 2015 Chen, Kishino. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** Third-party data are available from the Food Composition Database for Safety 184 Assessment of Genetically Modified Crops as Foods and Feeds and are listed as reference [24](#) and [25](#).

**Funding:** This work was supported by funding from the Japan Society for the Promotion of Science to HK (grant number 25280006). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

## Abstract

In the testing of food quality and safety, we contrast the contents of the newly proposed food (genetically modified food) against those of conventional foods. Because the contents vary largely between crop varieties and production environments, we propose a two-sample test of substantial equivalence that examines the inclusion of the tolerance intervals of the two populations, the population of the contents of the proposed food, which we call the target population, and the population of the contents of the conventional food, which we call the reference population. Rejection of the test hypothesis guarantees that the contents of the proposed foods essentially do not include outliers in the population of the contents of the conventional food. The existing tolerance interval ( $TI_0$ ) is constructed to have at least a pre-specified level of the coverage probability. Here, we newly introduce the complementary tolerance interval ( $TI_1$ ) that is guaranteed to have at most a pre-specified level of the coverage probability. By applying  $TI_0$  and  $TI_1$  to the samples from the target population and the reference population respectively, we construct a test statistic for testing inclusion of the two tolerance intervals. To examine the performance of the testing procedure, we conducted a simulation that reflects the effects of gene and environment, and residual from a crop experiment. As a case study, we applied the hypothesis testing to test if the distribution of the protein content of rice in Kyushu area is included in the distribution of the protein content in the other areas in Japan.

## Introduction

The safety assessment of genetically modified (GM) foods was confirmed as an important issue in the Organization for Economic Cooperation and Development (OECD) discussion resumed in 1988. Substantial equivalence has been a starting point for the safety assessment for GM foods which is used worldwide since this approach was first suggested in 1993 [1]. Substantial equivalence embodies the concept that if a new food or food component is found to be substantially equivalent to an existing food or feed component, it can be treated in the same manner

with respect to safety [2]. To decide if a modified product is substantially equivalent, the product is tested by the manufacturer for unexpected changes in a limited set of components such as toxins, nutrients, or allergens that are present in the unmodified food. Piaggio et al. [3] gave a clear framework of reporting of equivalence randomized trials. Ennis and Ennis [4,5] used an open interval to define equivalence and provided methods for testing a null hypothesis of non-equivalence. McNally et al. [6] proposed applying the generalized test function method in comparison to the confidence interval for assessing population bioequivalence. Herman and Price [7] examined research that has occurred over the past two decades relative to the mechanisms that affect crop composition in GM and traditionally bred crops.

In substantial equivalence tests of the population means, it is impossible to prove exact equality, so a buffer margin ( $c$ ) for the treatment effect is defined. The equivalence is defined as the treatment effect being between  $c$  and  $-c$ .

$$\begin{aligned} H_0 &: |\mu_1 - \mu_2| \geq c \\ H_1 &: |\mu_1 - \mu_2| < c \end{aligned} \tag{1}$$

A broad range of factors affect crop compositions, such as the genetic background [8,9], environmental factors [10,11], and agronomic practices [12,13]. Ricroch et al. [14] reviewed the published studies regarding the effect of genetic modification in comparison with the environmental and intervariety variations. Because the contents vary largely between crop varieties and production environments, the test of substantial equivalence should examine the inclusion of the tolerance intervals of the samples from the two populations, the population of the contents of the proposed food or feed, which we call the target population and denote as  $POP_{tar}$ , and the population of the contents of the conventional food or feed, which we call the reference population and denote as  $POP_{ref}$  (Fig 1).

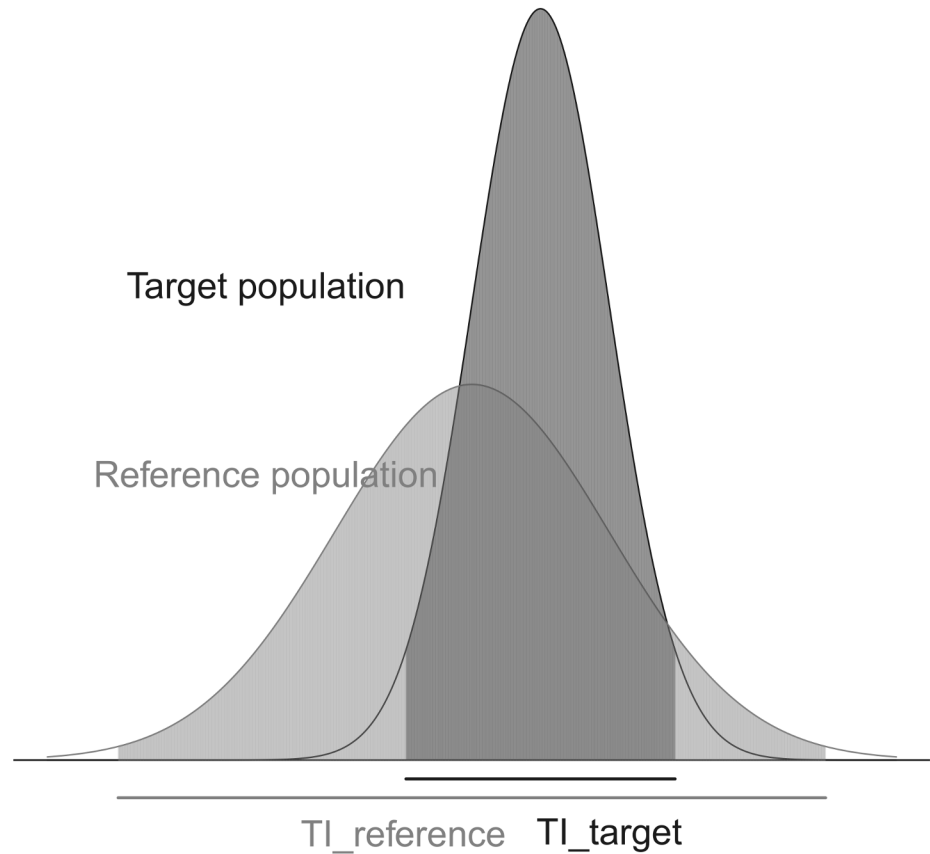
Statistical tolerance intervals are useful in practical applications in many areas and the construction of tolerance intervals has been extensively studied [15]. Formula for tolerance intervals (regions) for known and unknown mean and variance was given by Proschan [16] for univariate normal distribution and by Chew [17] for multivariate normal distribution. The tolerance interval procedure was developed for balanced one-way random model [18], general linear mixed models for balanced data [19] and unbalanced data [20]. A  $(1 - \gamma, 1 - \alpha)$  tolerance interval ( $TI_0$ ) based on a sample is constructed so that it would include at least a proportion  $1 - \gamma$  of the sampled population with confidence  $1 - \alpha$  [21]. Such a tolerance interval is usually referred to as  $(1 - \gamma)$ -content- $(1 - \alpha)$ -confidence (coverage) tolerance interval.

We introduced the complementary tolerance interval that is guaranteed to have at most a pre-specified level of the coverage probability. A  $(1 - \gamma, 1 - \alpha)$  tolerance interval ( $TI_1$ ) based on a sample is constructed so that it would include at most a proportion  $1 - \gamma$  of the sampled population with confidence  $1 - \alpha$ . By applying  $TI_0$  and  $TI_1$  to the samples from the target population and the reference population respectively, the rejection of the test guarantees that the target population essentially does not include outliers in the reference population.

## Material and Methods

### Two complementary tolerance intervals and two-sample hypothesis testing

We consider a sample  $X$  from a Gaussian population  $N(\mu, \sigma^2)$ . When the sample is collected by simple random sampling, the sample mean  $\hat{\mu}$  follows  $N(\mu, \sigma^2/R_0)$ , and the sample variance  $\hat{\sigma}^2$  follows  $(\sigma^2/m) \chi_m^2$ .  $R_0$  is the sample size, and the degree of freedom,  $m$ , is  $R_0 - 1$ . By allowing for the uncertainty of the sample mean and the sample variance, the conventional two-sided



**Fig 1. The distributions of two normal populations, the target population (POP<sub>tar</sub>) and the reference population (POP<sub>ref</sub>), and the tolerance intervals TI( $\gamma_{tar}$ , POP<sub>tar</sub>) and TI( $\gamma_{ref}$ , POP<sub>ref</sub>). As an example,  $\gamma_{tar}$  and  $\gamma_{ref}$  were set to 0.05 and 0.01 respectively.**

doi:10.1371/journal.pone.0141117.g001

(1 -  $\gamma$ )-content, (1 -  $\alpha$ )-confidence tolerance interval is defined as

$$TI_0 = \left( \hat{\mu} - \hat{\sigma} \sqrt{\frac{m \cdot \chi_{1;1-\gamma}^2(\text{ncp} = 1/R_0)}{\chi_{m;\alpha}^2(\text{ncp} = 0)}}, \hat{\mu} + \hat{\sigma} \sqrt{\frac{m \cdot \chi_{1;1-\gamma}^2(\text{ncp} = 1/R_0)}{\chi_{m;\alpha}^2(\text{ncp} = 0)}} \right), \tag{2}$$

where  $\chi_{1;1-\gamma,1/R_0}^2$  denotes the upper 100(1 -  $\gamma$ )% point of the non-central chi-squared distribution with degree of freedom one and non-centrality parameter  $1/R_0$ , and  $\chi_{m;\alpha}^2$  denotes the upper 100 $\alpha$ % point of chi-squared distribution with degree of freedom  $m$  [22]. The notation ncp stands for non-centrality parameter.  $R_0$  is the ratio of  $\sigma^2$  over the variance of  $\hat{\mu}$ , and  $m$  represents the ratio of  $2(\hat{\sigma}^2)^2$  over variance of  $\sigma^2$ . The tolerance interval  $TI_0$  covers at least 1 -  $\gamma$  of the population with the probability of 1 -  $\alpha$ .

Here, we introduce a new tolerance interval  $TI_1$  defined by

$$TI_1 = \left( \hat{\mu} - \hat{\sigma} \sqrt{\frac{m \cdot \chi_{1;1-\gamma}^2(\text{ncp} = 1/R_0)}{\chi_{m;1-\alpha}^2(\text{ncp} = 0)}}, \hat{\mu} + \hat{\sigma} \sqrt{\frac{m \cdot \chi_{1;1-\gamma}^2(\text{ncp} = 1/R_0)}{\chi_{m;1-\alpha}^2(\text{ncp} = 0)}} \right), \tag{3}$$

where  $\chi_{m;1-\alpha}^2$  denotes the lower 100 $\alpha$ % point of Chi-squared distribution with degree of freedom  $m$ . As is seen below, the tolerance interval  $TI_1$  covers at most 1 -  $\gamma$  of the population with

the probability of  $1 - \alpha$ . By increasing the sample size, the two complementary tolerance intervals both converge to the population tolerance interval.

We contrast the tolerance interval of the target population,  $POP_{tar}$ , with the tolerance interval of the reference population,  $POP_{ref}$ . Given the values of  $\gamma_{tar}$  and  $\gamma_{ref}$  ( $\gamma_{tar} > \gamma_{ref}$ ), the null hypothesis is that the tolerance interval of  $POP_{tar}$  is not included in the tolerance interval of  $POP_{ref}$ . The alternative is that the tolerance interval of  $POP_{tar}$  is included in the tolerance interval of  $POP_{ref}$ . To make the dependence of the tolerance intervals on the sample  $X$  and population  $P$  explicit, we express them as  $TI_0(\alpha, \gamma, X)$ ,  $TI_1(\alpha, \gamma, X)$ , and  $TI(\gamma, P)$ . Our framework of testing substantial equivalence is to test the null hypothesis,  $H_0$ , against the alternative hypothesis,  $H_1$ .

$$\begin{aligned} H_0 &: TI(\gamma_{tar}, POP_{tar}) \not\subset TI(\gamma_{ref}, POP_{ref}) \\ H_1 &: TI(\gamma_{tar}, POP_{tar}) \subset TI(\gamma_{ref}, POP_{ref}) \end{aligned} \tag{4}$$

We define the test statistic as,

$$\begin{aligned} \alpha_{TI01} &= \alpha_{TI01}(X_{tar}, X_{ref}) \\ &= \arg \min\{\alpha \mid TI_0(\alpha/2, \gamma_{tar}, X_{tar}) \subset TI_1(\alpha/2, \gamma_{ref}, X_{ref})\} \end{aligned} \tag{5}$$

where  $X_{tar}$  and  $X_{ref}$  are the sample from  $POP_{tar}$  and  $POP_{ref}$  respectively. The p value is obtained by locating the test statistic on its distribution for the case of  $TI(\gamma_{tar}, POP_{tar}) = TI(\gamma_{ref}, POP_{ref})$ .

### Mixed effect model and the coverage probabilities of the tolerance intervals

The two complementary tolerance intervals (Eqs (2) and (3)) can be generalized for the non-iid samples. The effective sample size,  $R_0$ , is obtained by comparing the variance of the estimated global mean with the total variance:  $V[\hat{\mu}] = (\hat{\sigma}_T^2)/R_0$ . The effective degree of freedom,  $m$ , is obtained by equating the estimated variance of the estimated total variance and the expected variance of the estimated total variance by the Satterthwaite's chi-square approximation:  $V[\hat{\sigma}_T^2] \approx 2(\hat{\sigma}_T^2)^2/m$ .

As an example, we consider the hypothetical samples with random genetic and environmental effects. The hypothetical samples reflect the maize samples of 61 lines from eight multi-site field studies. The field sites represented 47 unique environments in the commercial maize-growing regions of the United States, Canada, Chile and Argentina [23]. The experimental design used at each field site was a randomized complete block design containing three or four blocks. Variances of random components of concentrations of two analytes (tryptophan and oleic acid) were used to generate the simulated data. Table 1 shows the variances of random components of tryptophan and oleic acid. The variance component of environmental effect is large for tryptophan, whereas the genetic effect is the major component for oleic acid.

**Table 1. Variance of random components of a maize experiment.**

| Analyte    | % Total variance |      |     |     |            |
|------------|------------------|------|-----|-----|------------|
|            | G                | E    | GxE | B   | $\epsilon$ |
| Tryptophan | 6.7              | 71.6 | 3.5 | 0.1 | 18.1       |
| Oleic acid | 55.6             | 16.4 | 8.7 | 0.1 | 19.3       |

G, genotype; E, environment; GxE, interaction of genotype and environment; B, block;  $\epsilon$ , residual.

doi:10.1371/journal.pone.0141117.t001

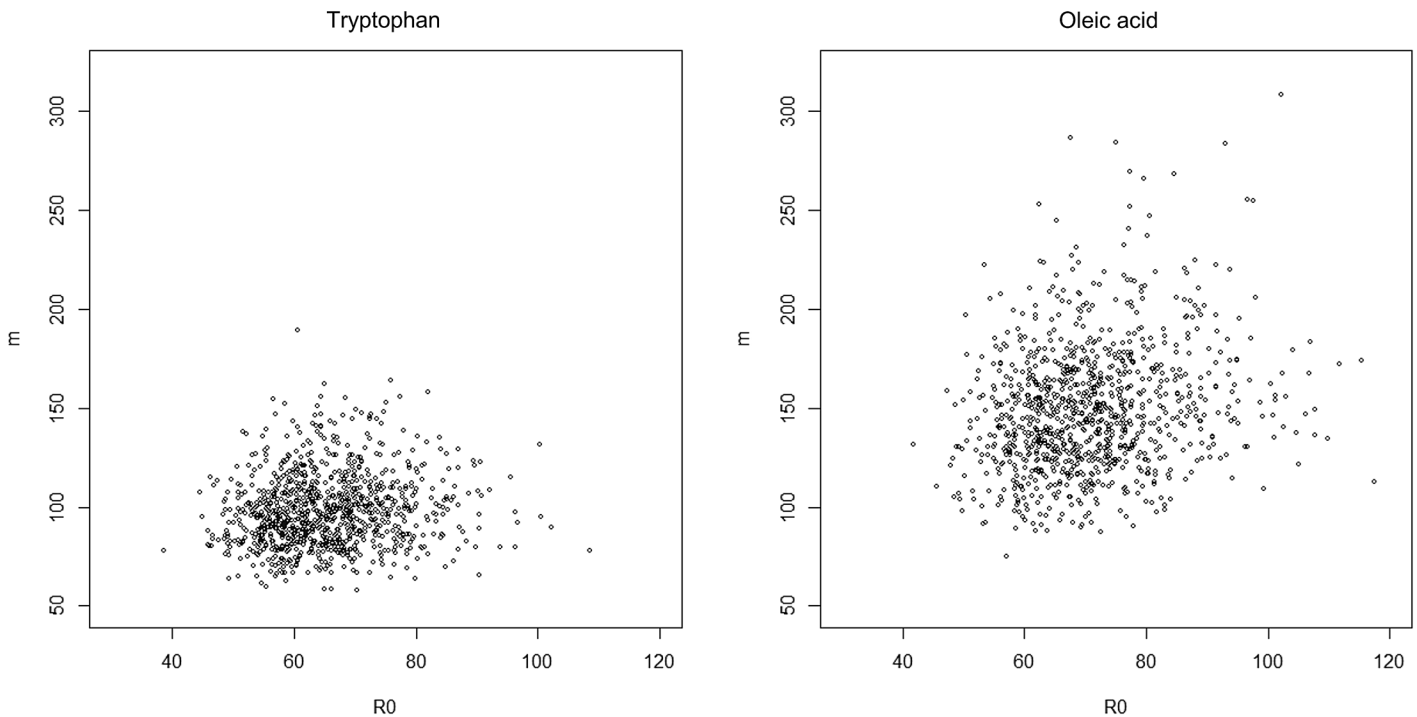
**Table 2. Design setup for the simulation study.**

| Environment | Genotype |         |         |         |         |
|-------------|----------|---------|---------|---------|---------|
|             | G01–G10  | G11–G20 | G21–G30 | G31–G40 | G41–G50 |
| E01–E10     | ✓        | ✓       |         |         |         |
| E11–E20     |          | ✓       | ✓       |         |         |
| E21–E30     |          |         | ✓       | ✓       |         |
| E31–E40     |          |         |         | ✓       | ✓       |
| E41–E50     | ✓        |         |         |         | ✓       |

doi:10.1371/journal.pone.0141117.t002

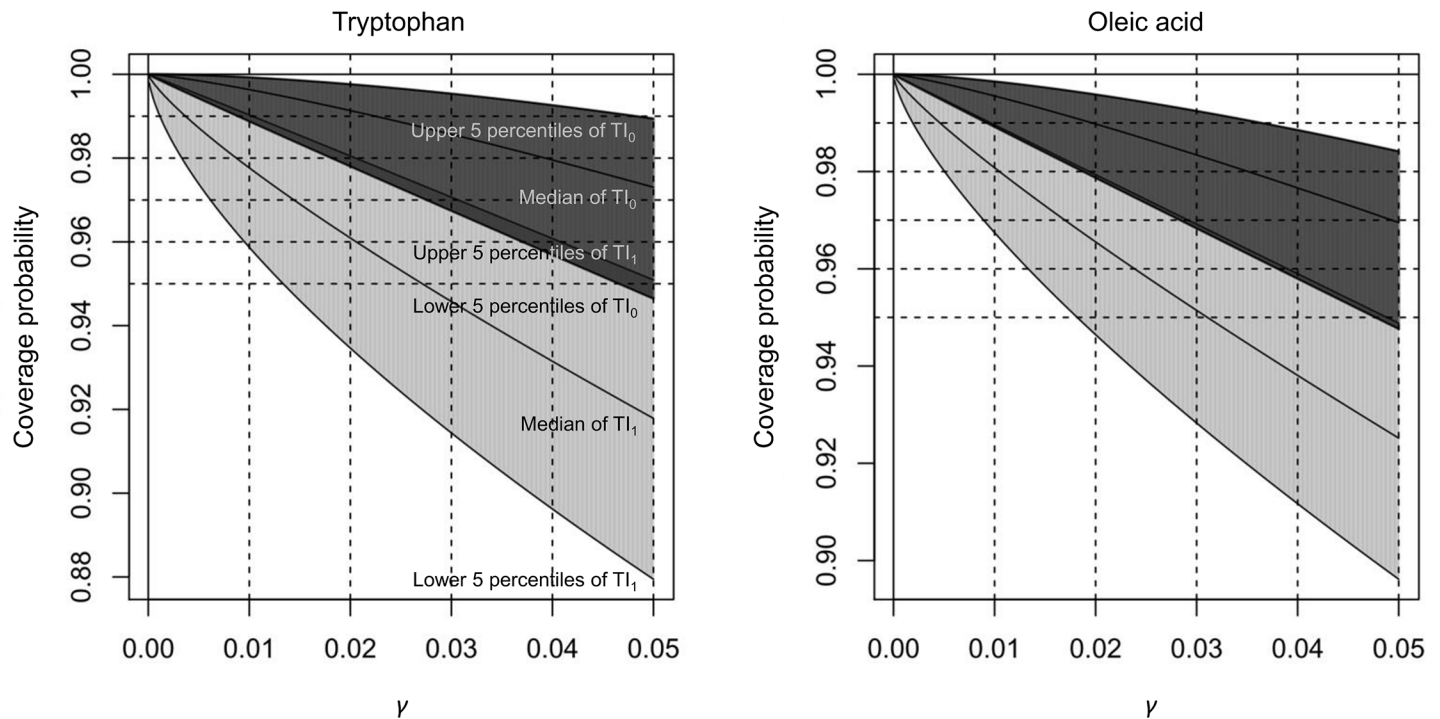
Table 2 shows the simulation setting with total number of environment,  $n_E = 50$ , total number of genotype,  $n_G = 50$  and number of blocks per environment,  $n_B = 4$ . We generated 1,000 sample datasets by normal random numbers with the variances in Table 1. We applied a linear mixed model to each of the dataset, and estimated the total mean and the variance components. The variance of total mean and the total variance, which are required for the calculation of  $R_0$  and  $m$ , were estimated by the variance among the 100 runs of parametric bootstrap.

The estimated  $m$  and  $R_0$  were distributed widely (Fig 2). The means of the estimated  $m$  were 98.0 for tryptophan and 149.3 for oleic acid. The means of the estimated  $R_0$  were 65.4 for tryptophan and 70.8 for oleic acid. Fig 3 shows the median, lower and upper 5 percentiles of the coverage probabilities of the tolerance intervals,  $TI_0$  and  $TI_1$ . The coverage probability of  $TI_0$  is larger than the nominal coverage probability  $(1 - \gamma)$  with probability 0.95. For the value of  $\gamma = 0.01$  (see the first dotted vertical line from the left on both panels), with probability 95%, the lower 5 percentiles of coverage probabilities of  $TI_0$  were larger than 98.9% and 98.9% for tryptophan and oleic acid respectively; the upper 5 percentiles of coverage probabilities of  $TI_0$  were



**Fig 2. The distributions of the estimated effective sample size ( $R_0$ ) and the effective degree of freedom ( $m$ ) for tryptophan and oleic acid separately.**

doi:10.1371/journal.pone.0141117.g002



**Fig 3. The median, lower and upper 5 percentiles of the coverage probabilities of the tolerance intervals,  $TI_0$  (dark gray area) and  $TI_1$  (light gray area).** For each area, middle, lower, and upper curves represent the median, lower and upper 5 percentiles of the coverage probabilities respectively. The 5 dotted horizontal lines represent the nominal coverage probability  $(1 - \gamma)$  for the 5 marked  $\gamma$  (0.01, 0.02, 0.03, 0.04, and 0.05).

doi:10.1371/journal.pone.0141117.g003

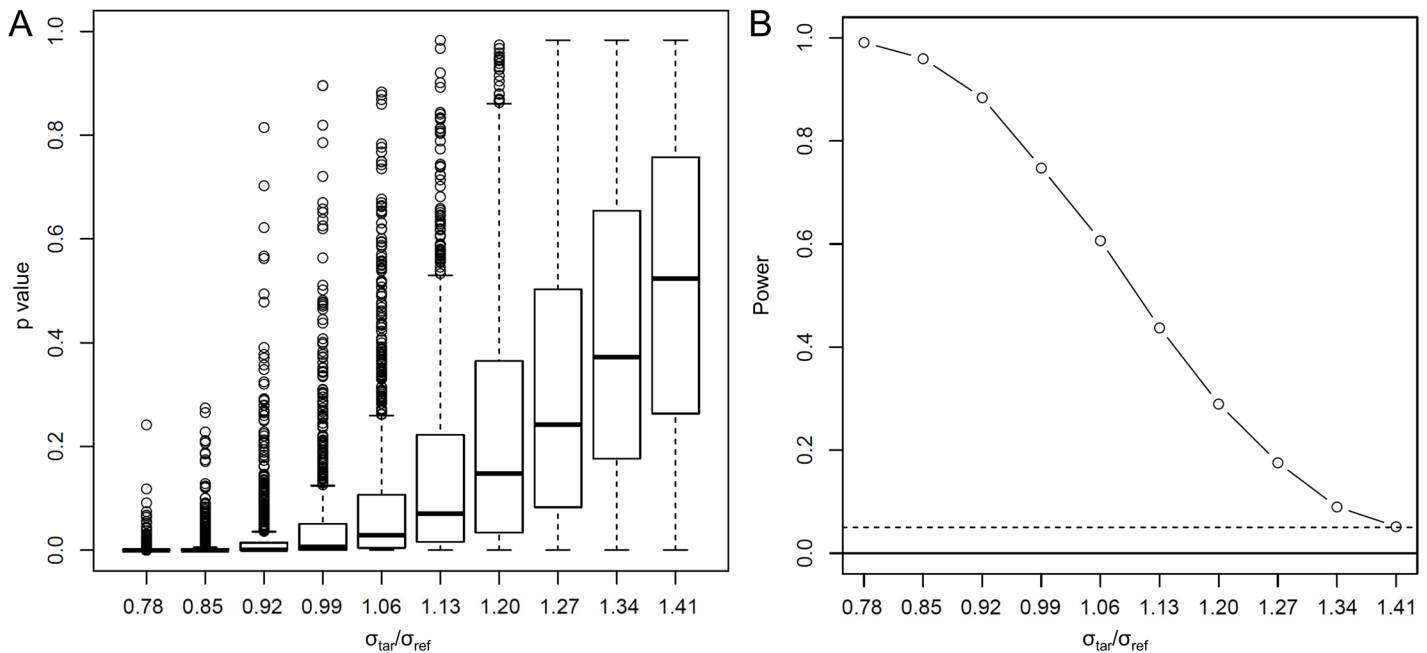
smaller than 99.9% and 99.8% for tryptophan and oleic acid respectively; the medians were 99.6% for both tryptophan and oleic acid. This means that  $TI_0$  covers at least  $1 - \gamma$  of the population with the probability 95%.

On the other hand, the coverage probability of  $TI_1$  is smaller than the nominal coverage probability  $(1 - \gamma)$  with probability 0.95. For the value of  $\gamma = 0.01$  (see the first dotted vertical line from the left on both panels), with probability 95%, the upper 5 percentiles of coverage probabilities of  $TI_1$  were smaller than 99.0% for both tryptophan and oleic acid; the lower 5 percentiles of coverage probabilities of  $TI_1$  were larger than 95.9% and 96.7% for tryptophan and oleic acid respectively; the medians were 97.8% and 98.1% for tryptophan and oleic acid respectively. This means that  $TI_1$  covers at most  $1 - \gamma$  of the population with the probability 95%.

## Results

### The p values and the power of the hypothesis test: a simulation study

To investigate the performance of the test procedure, we conducted a simulation study of contrasting two normal populations. The value of  $\gamma_{tar}$  and  $\gamma_{ref}$  were set to 0.05 and 0.01 respectively. The  $POP_{tar}$  and  $POP_{ref}$  are assumed to follow normal distribution with means  $\mu_{tar} = \mu_{ref} = 0$ . By solving the relation  $TI(\gamma_{tar}, POP_{tar}) = TI(\gamma_{ref}, POP_{ref})$ , we obtained  $\sigma_{tar0} = 1.41\sigma_{ref0}$  as the population parameter of the null hypothesis. The sample sizes were fixed to 50 for both  $POP_{tar}$  and  $POP_{ref}$ . The distribution under the null hypothesis was obtained by 10,000 simulation trials. For each value of  $\sigma_{tar} = (1 - 0.05i)\sigma_{tar0}$ ,  $i = 0, 1, 2, \dots, 10$ , 1,000 values of  $\alpha_{TI01}$  were generated randomly.



**Fig 4. The p value (A) and the power (B) of the test with sample size 50 for both target population and reference populations.** See [Material and methods](#) for the definition of the defined test statistics  $\sigma_{T101}$ . The dotted line represents the significance level of 0.05.

doi:10.1371/journal.pone.0141117.g004

Fig 4A shows the distribution of the p-values and the power of the test with the significance level of 0.05. The p-value followed mostly the uniform distribution when the null hypothesis is real ( $\sigma_{tar}/\sigma_{ref} = 1.41$ ). The power at the null hypothesis was 0.051, which was slightly larger but close to the significance level of 0.05 (Fig 4B). The power increased to 0.606 for the case of  $\sigma_{tar}/\sigma_{ref} = 1.06$ , and to 0.999 for the case of  $\sigma_{tar}/\sigma_{ref} = 0.78$ .

To see the effect of sample size, we conducted the simulation for the cases of sample sizes to 100, 150 and 200 for both  $POP_{tar}$  and  $POP_{ref}$ . The power to reject the null hypothesis with the significance level of 0.05 is shown in Table 3. The power for the case of  $\sigma_{tar}/\sigma_{ref} = 1.06$  became 0.852 when the sample size was doubled, and 0.987 when the sample size was 200. On the other hand, the power for the case of  $\sigma_{tar}/\sigma_{ref} = 1.41$  stayed nearly at the value of 0.05.

### A case study of testing inclusion of tolerance intervals: Contrasting protein composition of rice in Kyushu against other areas in Japan

As an example of empirical study, we applied the hypothesis testing to test if the protein value of rice in Kyushu area (Kyushu, including prefectures Fukuoka and Kagoshima) was included in the other areas in Japan (Japan). We downloaded the rice component data for Japan from

**Table 3. The power to reject the null hypothesis with significance level of 0.05 for each combination of sample size and the size of  $\sigma_{tar}/\sigma_{ref}$ .**

| Sample size | $\sigma_{tar}/\sigma_{ref}$ |       |       |       |       |       |
|-------------|-----------------------------|-------|-------|-------|-------|-------|
|             | 1.06                        | 1.13  | 1.20  | 1.27  | 1.34  | 1.41  |
| 50          | 0.606                       | 0.449 | 0.287 | 0.171 | 0.097 | 0.051 |
| 100         | 0.852                       | 0.684 | 0.456 | 0.246 | 0.114 | 0.053 |
| 150         | 0.948                       | 0.816 | 0.578 | 0.343 | 0.164 | 0.045 |
| 200         | 0.987                       | 0.915 | 0.708 | 0.390 | 0.183 | 0.058 |

doi:10.1371/journal.pone.0141117.t003



Table 4. Number of rice sample of varieties and production areas.

| Production area<br>(Prefecture) | Variety |     |     |     |     |     |     |     |     |     |     |     |     |     | Total |
|---------------------------------|---------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-------|
|                                 | Aic     | Aki | Don | Hae | Han | Hin | Hit | Hos | Kin | Kir | Kos | Mas | Tsu | Yum |       |
| Aichi                           | 5       |     |     |     |     |     |     |     |     |     |     |     |     |     | 5     |
| Akita                           |         | 9   |     |     |     |     |     |     |     |     |     |     |     |     | 9     |
| Aomori                          |         |     |     |     |     |     |     |     |     |     |     | 2   | 4   | 2   | 8     |
| Fukui                           |         |     |     |     | 9   |     |     |     |     |     | 1   |     |     |     | 10    |
| Fukuoka                         |         |     |     |     |     | 9   |     |     |     |     |     |     |     |     | 9     |
| Hokkaido                        |         |     |     |     |     |     |     | 9   |     | 9   |     |     |     |     | 18    |
| Hyogo                           |         |     | 1   |     |     |     |     |     |     |     |     |     |     |     | 1     |
| Ibaraki                         |         |     |     |     |     |     |     |     |     |     | 8   |     |     |     | 8     |
| Iwate                           |         |     |     |     |     |     | 9   |     |     |     |     |     |     |     | 9     |
| Kagoshima                       |         |     |     |     |     | 9   |     |     |     |     |     |     |     |     | 9     |
| Miyagi                          |         |     |     |     |     |     | 9   |     |     |     |     |     |     |     | 9     |
| Nagano                          |         | 2   |     |     |     |     |     |     |     |     |     |     |     |     | 2     |
| Niigata                         |         |     | 4   |     |     |     |     |     |     |     | 9   |     |     |     | 13    |
| Shiga                           |         |     |     |     |     |     |     |     | 9   |     |     |     |     |     | 9     |
| Tochigi                         |         |     |     |     |     |     |     |     |     |     | 9   |     |     |     | 9     |
| Yamagata                        |         | 1   |     | 9   |     |     |     |     |     |     |     |     |     |     | 10    |
| Total                           | 5       | 12  | 5   | 9   | 9   | 18  | 18  | 9   | 9   | 9   | 27  | 2   | 4   | 2   | 138   |

Aic, Aichinokaori; Aki, Akitakomachi; Don, Dontokoi; Hae, Haenuki; Han, Hanaechizen; Hin, Hinohikari; Hit, Hitomebore; Hos, Hoshinoyume; Kin, Kinuhikari; Kir, Kirara397; Kos, Koshihikari; Mas, Masshigura; Tsu, Tsugaruroman; Yum, Yumeakar.

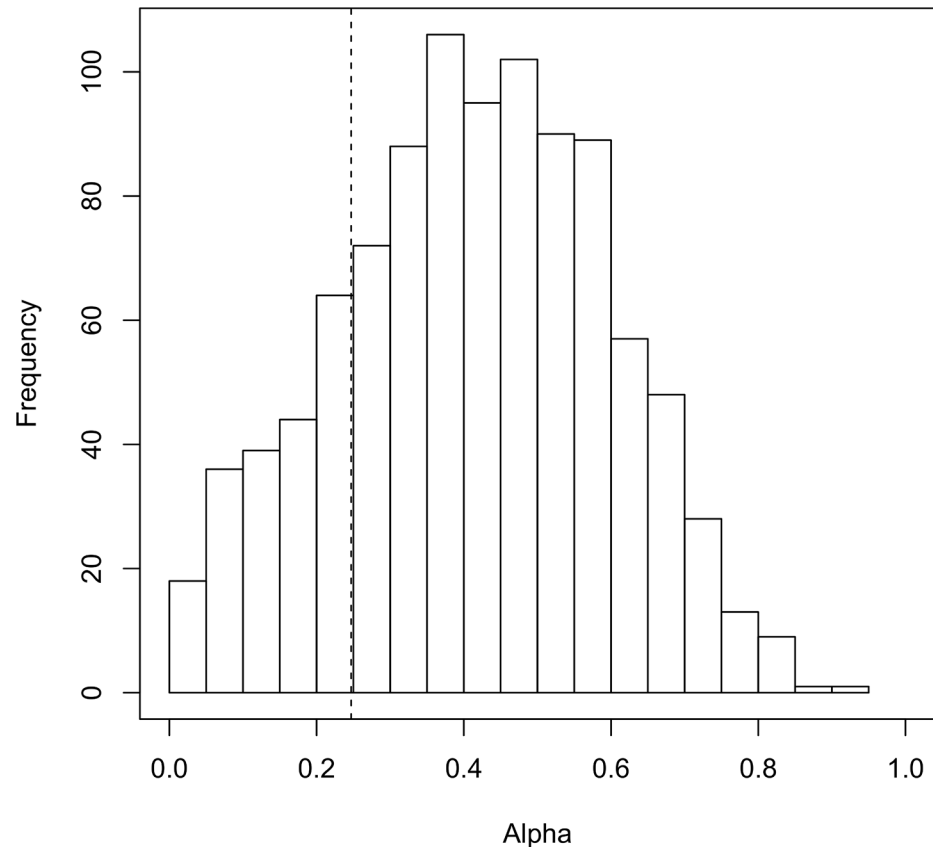
doi:10.1371/journal.pone.0141117.t004

The Food Composition Database for Safety Assessment of Genetically Modified Crops as Foods and Feeds [24,25]. It is third-party data and not owned by the authors. Major varieties of non-glutinous rice cultivated and distributed in Japan were collected from 1999 to 2009 (except for 2003 and 2004). A total of 15 or 16 samples consisting of 10–12 varieties were collected every year. The production areas are located in Japan stretching from the far north to south of the country. Table 4 shows the number of samples of different varieties and production areas. In total, the sample  $X_{Japan}$  includes 120 rice samples of varieties and the sample  $X_{Kyushu}$  includes 18 rice samples of varieties.

To test if the protein value of rice in Kyushu was included in Japan, we applied  $TI_0$  and  $TI_1$  to the samples from Kyushu and Japan respectively. The null hypothesis is that the tolerance interval of the protein of rice in Kyushu was not included in the tolerance interval of that in Japan. The alternative is that the tolerance interval of the protein of rice in Kyushu was included in the tolerance interval of that in Japan. The value of  $\gamma_{Kyushu}$  and  $\gamma_{Japan}$  were set to 0.05 and 0.01 respectively.

Using a linear mixed-effect model we estimated the total mean of  $POP_{Japan}$  as  $\mu_{Japan} = 6.60$  and random effects  $\sigma_{G,Japan}^2$ ,  $\sigma_{E,Japan}^2$  and  $\sigma_{GxE,Japan}^2 = 0.05, 0.07$  and  $0$  respectively, and the error term,  $\sigma_{\epsilon,Japan}^2 = 0.19$ . The total variance  $\sigma_{T,Japan}^2 = 0.31$ . The variance of the estimated total mean and that of the estimated total variance were estimated as  $V[\hat{\mu}_{Japan}] = 0.0128$  and  $V[\hat{\sigma}_{T,Japan}^2] = 0.00365$  respectively by 1,000 runs of parametric bootstrap. These values provide the effective sample size,  $R_{0,Japan} = \hat{\sigma}_{T,Japan}^2 / V[\hat{\mu}_{Japan}] = 24.28$  and the effective degree of freedom,  $m = 2(\hat{\sigma}_{T,Japan}^2)^2 / V[\hat{\sigma}_{T,Japan}^2] = 52.70$ . The sample from  $POP_{Kyushu}$  is assumed to be an iid sample from normal distribution with mean  $\mu_{Kyushu} = 6.92$  and variance  $\sigma_{T,Kyushu}^2 = 0.22$ . In this case,  $R_{0,Kyushu}$  is the sample size, 18, and  $m_{Kyushu}$  is  $R_{0,Kyushu} - 1 = 17$ . With these values of  $R_0$ 's





**Fig 5. Distribution of  $\alpha_{TIO1}$  obtained by 1,000 simulation trials under the null hypothesis.** The dotted line represents the value of 0.247.

doi:10.1371/journal.pone.0141117.g005

and  $m$ 's, we obtained  $TI_1(\alpha = 0.05, \gamma = 0.01, X_{Japan})$  as (5.34, 7.86) and  $TI_0(\alpha = 0.05, \gamma = 0.05, X_{Kyushu})$  as (5.59, 8.24). The latter does not include the former.

The value of the test statistic,  $\alpha_{TIO1}$ , was numerically obtained as 0.247 by solving the Eq (3). We obtained the p value by locating the value of  $\alpha_{TIO1}$  on the distribution under the null hypothesis. We generated this distribution by parametric bootstrap, assuming  $\mu_{Kyushu} = \mu_{Japan}$  and  $\sigma_{T,Kyushu} = 1.41\sigma_{T,Japan}$ . Without losing generality, we assumed  $\mu_{Kyushu} = \mu_{Japan} = 0$  and  $\sigma_{T,Japan} = 1$ . The iid sample of Kyushu was generated by normal random numbers with mean 0 and standard deviation 1.41. As for the sample of Japan, we generated the genetic effect (G), environmental effect (E), the G×E interaction and the error term, by decomposing the total variance into the variance components by the relative size of the estimated variance components. We generated 1,000 sets of the data. For each of the simulated data, we estimated the means and the total variances of Kyushu and Japan. We estimated their variances by 100 parametric bootstrap. With these estimates, we obtained  $R_0$ 's and  $m$ 's, and the value of  $\alpha_{TIO1}$ . From 1,000 values of  $\alpha_{TIO1}$ , we obtained the cumulative distribution of  $\alpha_{TIO1}$  under the null hypothesis (Fig 5). As a result, we obtained the p value as 0.195.

### Conclusion

In this study, we proposed a hypothesis test of inclusion of tolerance interval using the existing tolerance interval and a newly introduced the complementary tolerance interval. The result of simulation showed that the power of the test for the case of  $\sigma_{tar} / \sigma_{ref} = 1.41$  stayed nearly at the

value of 0.05 (Fig 4), which means that the testing procedure is almost unbiased. However, the test statistic,  $\alpha_{TI01}$ , is complex in form, and we could not attach a direct interpretation to it. We need make further effort to develop candidates of test statistics that measure the extent of coverage or non-coverage of the target population by the reference population. The mixed effect model enables unbiased estimation of the effective sample size and the effective degree of freedom, when the samples consist of subsamples collected in various conditions of genetic factors and environmental factors. However, a survey may be designed to collect the samples of matched controls. Another promising project is to develop a testing procedure for such samples.

As an alternative to the testing non-inferiority or substantial equivalence of the population mean, the proposed test examines the “range” of the distribution. A statistical test on the range of the distribution may be useful, especially when it is difficult to formulate the distribution by a simple statistical model. If a large sample is available, it is possible to construct non-parametric tolerance intervals [26,27]. The future study will investigate the statistical property of the non-parametric testing procedure.

## Acknowledgments

This work was supported by funding from the Japan Society for the Promotion of Science to HK (grant number 25280006).

## Author Contributions

Conceived and designed the experiments: HC HK. Analyzed the data: HC HK. Wrote the paper: HC HK.

## References

1. The Organisation for Economic Co-operation and Development. Safety Evaluation of Foods Derived by Modern Biotechnology: Concepts and Principles. 1993.
2. Food and Agriculture Organization of the United Nations. Joint FAO/WHO Expert Consultation on Biotechnology and Food Safety. Rome, Italy. 1996.
3. Piaggio G, Elbourne DR, Pocock SJ, Evans SJW, Altman DG. Reporting of noninferiority and equivalence randomized trials. *JAMA*. 2012; 308: 2594–2604. doi: [10.1001/jama.2012.87802](https://doi.org/10.1001/jama.2012.87802) PMID: [23268518](https://pubmed.ncbi.nlm.nih.gov/23268518/)
4. Ennis DM, Ennis JM. Hypothesis testing for equivalence based on symmetric open intervals. *Commun Stat–Theor M*. 2009; 38: 1792–1803.
5. Ennis DM, Ennis JM. Equivalence hypothesis testing. *Food Qual Prefer*. 2010; 21: 253–256.
6. McNally RJ, Iyer HK, Mathew T. Tests for individual and population bioequivalence based on generalized p-values. *Stat Med*. 2003; 22: 31–53. PMID: [12486750](https://pubmed.ncbi.nlm.nih.gov/12486750/)
7. Herman RA, Price WD. Unintended compositional changes in genetically modified (GM) crops: 20 years of research. *J Agric Food Chem*. 2013; 61: 11695–11701. doi: [10.1021/jf400135r](https://doi.org/10.1021/jf400135r) PMID: [23414177](https://pubmed.ncbi.nlm.nih.gov/23414177/)
8. Ahmadi M., Wiebold WJ, Beuerlein JE. Grain yield and mineral composition of corn as influenced by endosperm type and nitrogen. *Commun Soil Sci Plan*. 1993; 24: 2409–2426.
9. Reynolds TL, Nemeth MA, Glenn KC, Ridley WP, Astwood JD. Natural variability of metabolites in maize grain: differences due to genetic background. *J Agric Food Chem*. 2005; 53: 10061–10067. PMID: [16366695](https://pubmed.ncbi.nlm.nih.gov/16366695/)
10. Canvin DT. The effect of temperature on the oil content and fatty acid composition of the oils from several oil seed crops. *Can J Bot*. 1965; 43: 63–69.
11. Harrigan GG, Glenn KC, Ridley WP. Assessing the natural variability in crop composition. *Regul Toxicol Pharmacol*. 2010; 58: 513–520.
12. Wolfson JL, Shearer G. Amino acid composition of grain protein of maize grown with and without pesticides and standard commercial fertilizers. *Agron J*. 1981; 73: 611–613.

13. Wu P, Dai Q, Tao Q. Effect of fertilizer rates on the growth, yield, and kernel composition of sweet corn. *Commun Soil Sci Plan.* 1993; 24: 237–253.
14. Ricroch AE, Berge JB, Kuntz M. Evaluation of genetically engineered crops using transcriptomic, proteomic, and metabolomic profiling techniques. *Plant Physiol.* 2011; 155: 1752–1761. doi: [10.1104/pp.111.173609](https://doi.org/10.1104/pp.111.173609) PMID: [21350035](https://pubmed.ncbi.nlm.nih.gov/21350035/)
15. Hahn GJ, Meeker WQ. *Statistical Intervals*, New York: Wiley; 1991.
16. Prochan F. Confidence and tolerance intervals for the normal distribution. *JASA.* 1953; 48: 550–564.
17. Chew V. Confidence, prediction, and tolerance regions for the multivariate normal distribution. *JASA.* 1966; 61: 605–617.
18. Lemon GH. Factors for one-sided tolerance limits for balanced one-way ANOVA random-effects model. *JASA.* 1977; 72: 676–680.
19. Liao CT, Iyer HK. A tolerance interval for the normal distribution with several variance components. *Stat Sinica.* 2004; 14: 217–229.
20. Liao CT, Lin TY, Iyer HK. One- and two-sided tolerance intervals for general balanced mixed models and unbalanced one-way random models. *Technometrics.* 2005; 47: 323–335.
21. Krishnamoorthy K, Lian X. Closed-form approximate tolerance intervals for some general linear models and comparison studies. *JSCS.* 2012; 4: 547–563.
22. Krishnamoorthy K, Mathew T. *Statistical tolerance regions: theory, applications, and computation.* John Wileys & Sons, Inc.: Hoboken; 2009. pp. 461.
23. Hong B, Fisher T, Sult T, Maxwell C, Mickelson J, Kishino H, et al. Model-based tolerance intervals derived from cumulative historical composition data: application for substantial equivalence assessment of a genetically modified crop. *J Agric Food Chem.* 2014; 62: 9916–9926. doi: [10.1021/jf502158q](https://doi.org/10.1021/jf502158q) PMID: [25208038](https://pubmed.ncbi.nlm.nih.gov/25208038/)
24. National Agriculture and Food Research Organization. *The Food Composition Database for Safety Assessment of Genetically Modified Crops as Foods and Feeds.* 2014. Available: <http://afdb.dc.affrc.go.jp/afdb/index.asp>. Last accessed 4 March 2014.
25. Kitta K. Availability and utility of crop composition data. *J Agric Food Chem.* 2013; 61: 8304–8311. doi: [10.1021/jf400777v](https://doi.org/10.1021/jf400777v) PMID: [23718756](https://pubmed.ncbi.nlm.nih.gov/23718756/)
26. Wilks SS. Determination of sample sizes for setting tolerance limits. *Ann Math Stat.* 1941; 12: 91–96.
27. Balakrishnan N, Beutner E, Cramer E. Exact two-sample nonparametric confidence, prediction, and tolerance intervals based on ordinary and progressively type-II right censored data. *Test.* 2010; 19: 68–91.