



## OPEN ACCESS

## EDITED BY

Raghu Machiraju,  
The Ohio State University, United States

## REVIEWED BY

Yun Zhang,  
University of Texas Southwestern  
Medical Center, United States  
Feiyang Ma,  
University of Michigan, United States

## \*CORRESPONDENCE

Jeffrey A. Thompson,  
jthompson21@kumc.edu

## SPECIALTY SECTION

This article was submitted to Genomic  
Analysis,  
a section of the journal  
Frontiers in Bioinformatics

RECEIVED 09 March 2022

ACCEPTED 22 August 2022

PUBLISHED 21 September 2022

## CITATION

Neums L, Koestler DC, Xia Q, Hu J,  
Patel S, Bell-Glenn S, Pei D, Zhang B,  
Boyd S, Chalise P and Thompson JA  
(2022), Assessing equivalent and inverse  
change in genes between  
diverse experiments.  
*Front. Bioinform.* 2:893032.  
doi: 10.3389/fbinf.2022.893032

## COPYRIGHT

© 2022 Neums, Koestler, Xia, Hu, Patel,  
Bell-Glenn, Pei, Zhang, Boyd, Chalise  
and Thompson. This is an open-access  
article distributed under the terms of the  
[Creative Commons Attribution License  
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or  
reproduction in other forums is  
permitted, provided the original  
author(s) and the copyright owner(s) are  
credited and that the original  
publication in this journal is cited, in  
accordance with accepted academic  
practice. No use, distribution or  
reproduction is permitted which does  
not comply with these terms.

# Assessing equivalent and inverse change in genes between diverse experiments

Lisa Neums<sup>1,2</sup>, Devin C. Koestler<sup>1,2</sup>, Qing Xia<sup>1,2</sup>, Jinxiang Hu<sup>1,2</sup>,  
Shachi Patel<sup>1,2</sup>, Shelby Bell-Glenn<sup>1,2</sup>, Dong Pei<sup>1,2</sup>, Bo Zhang<sup>1</sup>,  
Samuel Boyd<sup>1,2</sup>, Prabhakar Chalise<sup>1,2</sup> and  
Jeffrey A. Thompson<sup>1,2\*</sup>

<sup>1</sup>Department of Biostatistics and Data Science, University of Kansas Medical Center, Kansas City, KS, United States, <sup>2</sup>University of Kansas Cancer Center, Kansas City, KS, United States

**Background:** It is important to identify when two exposures impact a molecular marker (e.g., a gene's expression) in similar ways, for example, to learn that a new drug has a similar effect to an existing drug. Currently, statistically robust approaches for making comparisons of equivalence of effect sizes obtained from two independently run treatment vs. control comparisons have not been developed.

**Results:** Here, we propose two approaches for evaluating the question of equivalence between effect sizes of two independent studies: a bootstrap test of the Equivalent Change Index (ECI), which we previously developed, and performing Two One-Sided t-Tests (TOST) on the difference in log-fold changes directly. The ECI of a gene is computed by taking the ratio of the effect size estimates obtained from the two different studies, weighted by the maximum of the two *p*-values and giving it a sign indicating if the effects are in the same or opposite directions, whereas TOST is a test of whether the difference in log-fold changes lies outside a region of equivalence. We used a series of simulation studies to compare the two tests on the basis of sensitivity, specificity, balanced accuracy, and F1-score. We found that TOST is not efficient for identifying equivalently changed gene expression values (F1-score = 0) because it is too conservative, while the ECI bootstrap test shows good performance (F1-score = 0.95). Furthermore, applying the ECI bootstrap test and TOST to publicly available microarray expression data from pancreatic cancer showed that, while TOST was not able to identify any equivalently or inversely changed genes, the ECI bootstrap test identified genes associated with pancreatic cancer. Additionally, when investigating publicly available RNAseq data of smoking vs. vaping, no equivalently changed genes were identified by TOST, but ECI bootstrap test identified genes associated with smoking.

**Conclusion:** A bootstrap test of the ECI is a promising new statistical approach for determining if two diverse studies show similarity in the differential expression of genes and can help to identify genes which are similarly

**Abbreviations:** DO, Disease Ontology; DGN, DisGeNET; ECI, Equivalent Change Index; PBMC, peripheral blood mononuclear cells; TOST, Two One-Sided t-Tests.

influenced by a specific treatment or exposure. The R package for the ECI bootstrap test is available at <https://github.com/Hecate08/ECIbootstrap>.

#### KEYWORDS

differential gene expression, equivalence test, study comparison, diverse experiments, molecular marker

## 1 Introduction

In whole genome differential gene expression studies, the difference in the expression of thousands of genes among groups (e.g., treatment group vs. control group) is investigated with the motivation of finding underlying mechanisms of different conditions such as cancer or to observe treatment effects (Gilbert, 2000). An emerging use for gene expression data is to identify genes that are affected in similar or opposing ways across different studies. There are various reasons for doing so, including validating the results of a study (Fu et al., 2021), finding common underlying mechanisms of a disease (Shen et al., 2016; Blake et al., 2020), or investigating similar treatment effects of different drugs (Hollenbach et al., 2010). One of the important challenges in comparing studies is variable study conditions such as technology, environment, and personnel (Goh et al., 2017). For these reasons, gene expression levels are challenging to directly compare across studies. In many cases though, the focus is on the nature of the change in gene expression between treatment or disease conditions (Gilbert, 2000). Therefore, there is a need for statistical methods that can be used to validate the similarity in expression changes across studies.

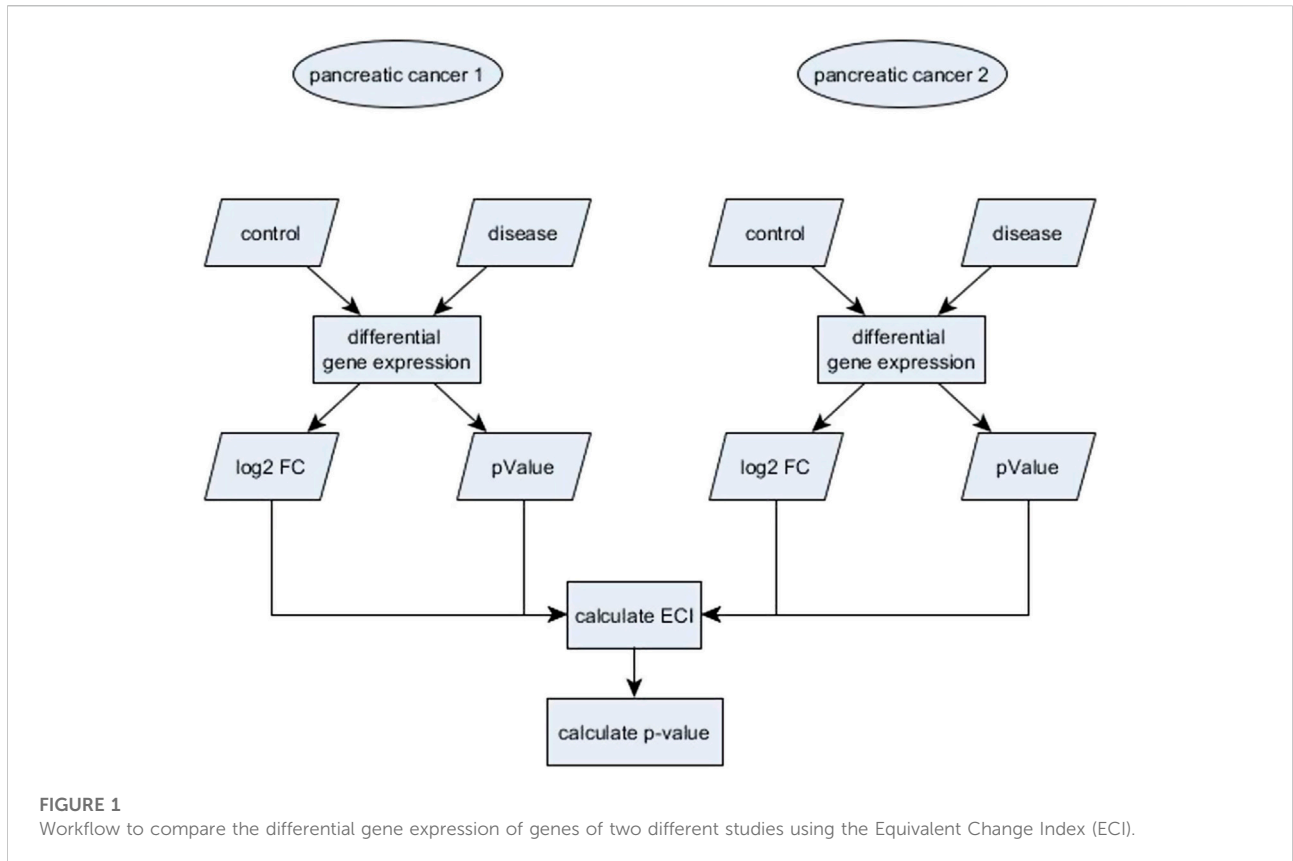
Given the lack of statistical methods that can test such equivalence hypotheses, researchers currently use naïve methods for determining if treatments have similar effects on gene expression. The most-commonly employed method focused on intersection, which simply involves finding the intersection of differentially expressed genes across studies, without determining the probability of such intersections occurring by chance. One example is the e-cigarette study from Shen et al. (2016), which investigated if similar pathways were enriched in cigarette smokers. In the process of analyzing their results, they observed that the same or different genes were differentially expressed at several time points of their study and used this result as well as a subsequent gene enrichment study to form their conclusion. Another example is a study from Blake et al. (2020), which compares gene expression across tissues in different animal species. Again, they define genes of interest as being both equivalently changed in the same direction and statistically significant in a pairwise comparison of tissue types. However, because this naïve approach is not statistically motivated, it is prone to false positives, i.e., declaring a gene equivalently changed when the differential expression of the gene between two studies is in the same direction only by chance. This could lead to a large number of candidate genes requiring validation,

which could be time-consuming and cost prohibitive (Provenzano and Mocellin, 2007; Garrido et al., 2020).

In this study, we introduce and compare two statistical tests of equivalent change in gene expression between two studies. An advantage to testing for an equivalent change in differential gene expression is that the log<sub>2</sub>-fold changes are likely to be far more comparable across studies than the gene expression itself due to variable study conditions such as technology, environment, and personnel (Rudy and Valafar, 2011; Walsh et al., 2015). It should be noted that these tests work equally well to find significant opposing changes in gene expression which, for example, could be used to suggest treatments that might reverse changes associated with a disease or to identify genes affected in opposite ways in a gene knockout and overexpression experiment. The first test is an adaptation of the two one-sided t-tests (TOST) applied to the fold-changes from a differential gene expression analysis. Although traditionally the TOST approach has been used to establish bioequivalence of drugs (e.g., to approve generics) (Schuirmann, 1987), it has been adapted to several other cases (Dixon et al., 2018; Leichsenring et al., 2018; Wu et al., 2018), including finding equivalently expressed genes in the same study (Qiu and Cui, 2010). Nevertheless, it has never been adapted to differential gene expression analysis. The second test is based on the Equivalent Change Index (ECI), introduced by Thompson and Koestler (2020) in connection with gene enrichment analysis. Although they introduced the statistic itself, no statistical test was available. Here, we present a bootstrap procedure for the ECI statistic to calculate confidence intervals. In the following, we will explain the mechanisms of both approaches followed by a simulation study to compare and assess the performance of the two methods. This is followed by real-world data applications using different publicly available gene expression studies of pancreatic cancer, Alzheimer's disease, and smoking vs. vaping to demonstrate if the tests can identify biologically plausible results.

## 2 Methods

In the following, we give a definition of the Equivalent Change Index (ECI), the ECI bootstrap test, and the Two One-Sided t-Tests (TOST). Furthermore, we explain the design of the simulation study and give information about the publicly available data sets used in our real data analysis. For the purpose of this study, we need an effect size  $\hat{\beta}_1$ , the standard deviation (for TOST) of  $\hat{\beta}_1$ , and a measurement of statistical



significance for  $\hat{\beta}_i$  (for the ECI bootstrap test) for each differentially expressed gene  $i$ . We decided to use the log2 fold change (log2 FC) as effect size, although other measures could also be used, for example the standardized mean difference.

### 2.1 Equivalent change index

The Equivalent Change Index (ECI), proposed by Thompson and Koestler (2020), is a measure of the degree of equivalent or inverse change of attributes of the same type across two diverse studies. The ECI  $\lambda_i$  of a gene  $i$  is calculated as a ratio of the minimum and maximum of the absolute effect sizes  $\hat{\beta}_{ik}$  from the two studies ( $k = [1, 2]$ ) multiplied by a sign, where the sign reflects whether the differential gene expression of the two studies was in the same direction (positive sign) or opposite direction (negative sign). Furthermore, the ECI is weighted by the maximum of the  $p$ -values  $p_{ik}$  of the two effect sizes.

$$\lambda_i = \frac{\text{sgn}(\hat{\beta}_{i1} \times \hat{\beta}_{i2}) \min(|\hat{\beta}_{i1}|, |\hat{\beta}_{i2}|)}{\max(|\hat{\beta}_{i1}|, |\hat{\beta}_{i2}|)} \cdot (1 - \max(p_{i1}, p_{i2}))$$

$\lambda_i$  is in the range of  $[-1, 1]$ , where  $-1$  indicates that the effect size was exactly opposite between the two studies and  $1$  indicates that the effect sizes between the two studies were identical. Hence,  $\lambda_i$  indicates the degree of equivalence or inverseness of the expression of a gene compared between two separate experiments.

### 2.2 ECI bootstrap test

In this section we describe our proposed bootstrap procedure to use the ECI statistic to test for equivalence. We performed all computations in the R environment (R Project for Statistical Computing, RRID:SCR\_001905) version 3.5. We first calculate the ECI values for all genes using the function `getECI()` from the R package ECEA (Thompson and Koestler, 2020) with the log2 FC as effect size and corresponding  $p$ -values as measurement of statistical significance (Figure 1). To obtain a test measure for equivalent change, we proceed as follows:

- create bootstrap samples within treatment and control groups for each study
- recalculate the differential gene expression for each study separately, e.g., the log2 FC for each gene along with its  $p$ -value

- recalculate ECI values between studies for each gene

We repeated this procedure 1,000 times.

### 2.2.1 Confidence interval

We used a 95% bias-corrected and accelerated (BCa) bootstrap interval as a confidence interval for the ECI values since it corrects for both bias and skewness (Wicklin, 2017; Jung et al., 2019). The BCa assumes that the data are independent and identically distributed and is calculated as  $CI[\hat{\beta}_L^*, \hat{\beta}_U^*]$ , where  $\hat{\beta}_L^*$  denotes the  $L$ th quantile and  $\hat{\beta}_U^*$  denotes the  $U$ th quantile, i.e., lower limit and upper limit. The indices  $L$  and  $U$  are defined as  $L = a_1 \cdot B$  and  $U = a_2 \cdot B$ , where  $B$  is the number of bootstrap samples, here  $B = 1000$ , and  $a_1$  and  $a_2$  are defined as:

$$a_1 = \Phi\left(\hat{z}_0 + \frac{\hat{z}_0 + z^{(\alpha/2)}}{1 - \hat{a}(\hat{z}_0 + z^{(\alpha/2)})}\right)$$

$$a_2 = \Phi\left(\hat{z}_0 + \frac{\hat{z}_0 + z^{(1-\alpha/2)}}{1 - \hat{a}(\hat{z}_0 + z^{(1-\alpha/2)})}\right)$$

Here,  $\Phi$  is the standard normal cumulative distribution function,  $z^{(\alpha/2)}$  is the  $100x(\frac{\alpha}{2})$  th percentile of a standard normal distribution, and  $\hat{z}_0$  and  $\hat{a}$  are the bias-corrections and acceleration factors, respectively. The bias-correction factor is computed as the inverse of the standard normal cumulative distribution function of the proportion of bootstrap effect sizes  $\hat{\beta}^*$  smaller than the original effect size  $\hat{\beta}$ :

$$\hat{z}_0 = \Phi^{-1}\left(\frac{1}{B} \sum_{b=1}^B I[\hat{\beta}^* < \hat{\beta}]\right)$$

The acceleration factor uses “leave one out” (jackknife) resampling as follows:

$$\hat{a} = \frac{\sum_{i=1}^n \left(\hat{\beta}_{(i)} - \hat{\beta}_{(-i)}\right)^3}{6 \left[\sum_{i=1}^n \left(\hat{\beta}_{(i)} - \hat{\beta}_{(-i)}\right)^2\right]^{3/2}}$$

Where,  $\hat{\beta}_{(-i)}$  is an effect size of a jackknife sample and  $\hat{\beta}_{(i)}$  is the average of the effect sizes of all jackknife samples.

We used the bias corrected confidence interval to test the null hypothesis that the ECI value is not different from zero, in other words, there is no equivalent or inverse change between the effect sizes of the attributes of the two studies. We reject the null hypothesis when the confidence interval of the ECI values does not include zero.

### 2.2.2 P-value (approximate)

According to Asparouhov and Muth'en (Tihomir Asparouhov aBMe, 2021), assuming that  $\hat{a} = 0$ , we have

$$a_1 = \Phi\left(\hat{z}_0 + \hat{z}_0 + z^{(\alpha/2)}\right) = \Phi\left(2\hat{z}_0 + z^{(\alpha/2)}\right)$$

$$a_2 = \Phi\left(\hat{z}_0 + \hat{z}_0 + z^{(1-\alpha/2)}\right) = \Phi\left(2\hat{z}_0 + z^{(1-\alpha/2)}\right)$$

When we concentrate only on positive ECI values, we would reject the null hypothesis of non-equivalence when  $\hat{\beta}_L^* > 0$  or  $L > M$  where

$$M = \sum_{b=1}^B I[\hat{\beta}^* < 0]$$

In the special case of  $p$ -value = 0.05 ( $L = M$ ) we have

$$0.5 = 2 \cdot 0.025 = 2 \cdot \Phi(-1.96)$$

$$= 2 \cdot \Phi\left(-2\hat{z}_0 + 2\hat{z}_0 - 1.96\right)$$

$$= 2 \cdot \Phi\left(-2\hat{z}_0 + \Phi^{-1}\left(\Phi\left(2\hat{z}_0 - 1.96\right)\right)\right)$$

$$= 2 \cdot \Phi\left(-2\hat{z}_0 + \Phi^{-1}\left(a_1 \cdot B/B\right)\right)$$

$$= 2 \cdot \Phi\left(-2\hat{z}_0 + \Phi^{-1}(L/B)\right)$$

Therefore, the two-sided  $p$ -value can be computed, approximately, by

$$p\text{-value} = 2 \cdot \Phi\left(-2\hat{z}_0 + \Phi^{-1}(M/B)\right)$$

where  $M$  is the number of ECI values, which are smaller than zero. In case of the ECI value being smaller than zero we compute

$$p\text{-value} = 1 - 2 \cdot \Phi\left(-2\hat{z}_0 + \Phi^{-1}(M/B)\right)$$

## 2.3 Two one-sided t-tests

What follows is a brief description of how the TOST method was applied to these data. Let  $\Delta$  represent some fold change which is considered to be unimportant. To test for equivalent change, we use a null hypothesis of non-equivalence vs. an alternative hypothesis of equivalence or inverse relationship. In particular, we must perform two hypothesis tests, namely:

$$H_{01}: \beta_{i1} - \beta_{i2} \geq \Delta \text{ vs } H_{a1}: \beta_{i1} - \beta_{i2} < \Delta$$

and,

$$H_{02}: \beta_{i1} - \beta_{i2} \leq -\Delta \text{ vs } H_{a2}: \beta_{i1} - \beta_{i2} > -\Delta$$

where  $\beta_{ik}$  is the effect size of gene  $i$  of study  $k$ . The  $t$ -test statistics for the two tests are:

$$t_i^{*1} = \frac{\hat{\beta}_{i1} - \hat{\beta}_{i2} - \Delta}{s_i}$$

and,

$$t_i^{*2} = \frac{\hat{\beta}_{i1} - \hat{\beta}_{i2} + \Delta}{s_i}$$

To accommodate for unequal variances we define  $s_i$  according to Welch's  $t$ -test, where  $s_i = \sqrt{\frac{sd_1^2}{N_1} + \frac{sd_2^2}{N_2}}$ ,  $sd_k$  is the standard deviation of  $\hat{\beta}_{ik}$ , and  $N_k$  is the size of study  $k$ . To perform the  $t$ -test we used Satterthwaite formula for degrees of freedom:

$$df = \frac{\left(\frac{sd_1^2}{N_1} + \frac{sd_2^2}{N_2}\right)^2}{\frac{\left(\frac{sd_1^2}{N_1}\right)^2}{N_1-1} + \frac{\left(\frac{sd_2^2}{N_2}\right)^2}{N_2-1}}$$

Only when the  $p$ -value for test one and the  $p$ -value for test two are smaller than the significance level of  $\alpha = 0.05$ , is a gene considered to be equivalently changed. Likewise, to test for inverse change, we use the same test statistics except that we multiply  $\hat{\beta}_{i1}$  by  $-1$ .

## 2.4 Multiple testing

We performed a test of equivalence for each gene. To adjust the  $p$ -value for the multiple testing we used the adjusted false discovery rate (FDR) approach by Benjamini and Hochberg (1995). Here, the FDR correction  $q_i = p_i N/i$ , where  $p_i$  is the  $i$ th  $p$ -value in a sorted list of ascending  $p$ -values and  $N$  is the total number of  $p$ -values, is the ratio of expected false positives and the total number of accepted positives. To adjust for the non-monotony of the  $q_i$  value we replace the  $q_i$  value with the lowest  $q_i$  value among all  $q_f$ , where  $f \geq i$  (Yekutieli and Benjamini, 1999).

## 2.5 Simulation

We conducted a simulation study to compare the two different tests for equivalent change. This simulation study consisted of two simulated studies of differential gene expression, for which we aimed to test equivalent change of gene expression. The two simulated studies, each consisting of 1,000 genes, are constructed so that 30% of the differentially expressed genes of Simulated Study 1 are equivalently changed between the two studies. The simulation process has two levels. First, we simulated descriptive features of each simulated study (mean and standard deviation), which we used in the second level to draw random gene expression values for each sample. We then used the simulated studies to performed differential gene expression and the equivalence testing as described above. In the following we will explain how each study is set up step by step.

### 2.5.1 Simulated study 1

The first stage is to simulate a study with two groups. The group is denoted by  $k$ , with controls as  $k = 0$  and cases or treatment as  $k = 1$ . The simulation derives certain values from a reference study of pancreatic cancer (see section Biological Data) with case control

data, including samples that were either tumor or tumor adjacent normal tissue (GSE16515), as described below.

- a. The mean expression for gene  $g$  in the simulated control group ( $k = 0$ ) is drawn from a gamma distribution with scale and shape parameters extracted from the reference dataset using the function `egamma()` of the R package `EnvStats` (Millard, 2013):

$$m_g \sim \Gamma(\alpha = 8.24, \beta = 0.66)$$

- b. For each gene  $g$  from the reference dataset we obtained the mean difference in expression, denoted by  $\delta_g$ .
- c. Genes with  $-1 \geq \delta_g \leq 1$  are removed. For this simulation we are only interested in genes with a difference between groups.
- d. From the filtered down gene set we obtained a new gene set by sampling with replacement, where each gene  $g$  has:
  - i. The standard deviation for gene  $g$  in group  $k$  of the reference dataset, denoted by  $s_{gk}$
  - ii. The mean difference  $\delta_g$
- e. The genes were divided into three subgroups: equivalently changed genes ( $f = 1$ ), non-equivalently changed genes 1 ( $f = 2$ ), and non-equivalently changed genes 2 ( $f = 3$ ). The non-equivalently changed genes were divided into two subgroups to have the differential expression of those genes be balanced between the two studies: One half of the genes is differentially expressed in one study while non-differentially expressed in the other study and vice versa.
- f. The expression value for gene  $g$ , observation  $i$ , group  $k$  is denoted  $x_{gik}$  and is drawn from a truncated normal distribution:

$$x_{gik} \sim N(\gamma_g, s_{gk}); 0 < x_{gik}$$

- With

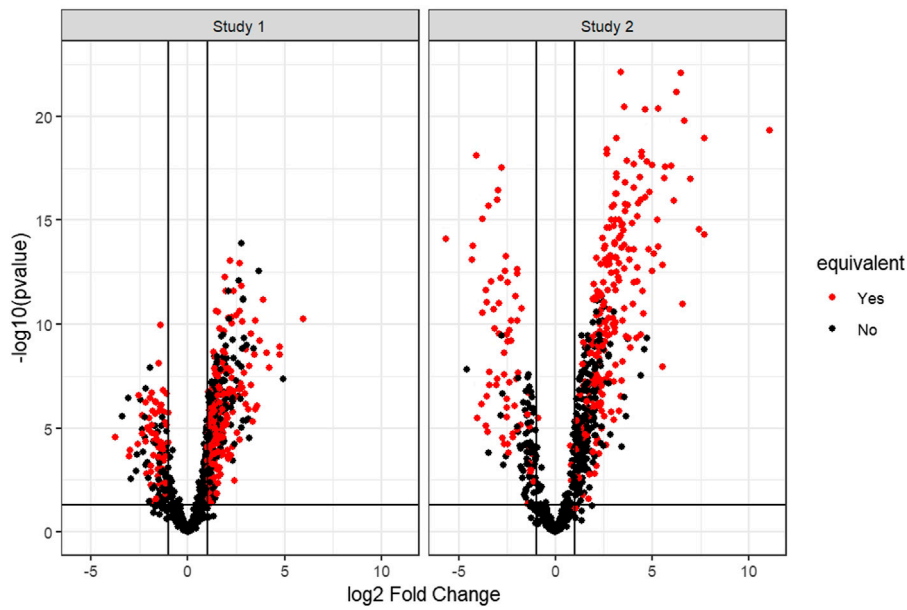
$$\gamma_g = \begin{cases} m_g + k\delta_g & \text{if } f = 1 \\ m_g & \text{if } f = 2 \\ m_g + k\delta_g & \text{if } f = 3 \end{cases}$$

- g. With the simulated control group and treatment group, we calculated the differential expression for each gene as described in section "Differential Gene Expression".

### 2.5.2 Simulated study 2

The next stage is to simulate gene expression from a second, similar study. At this point, we will determine genes which are equivalently changed across the two studies. We will choose 30% of the differentially expressed genes ( $abs(\log_2FC) > 1$  and  $p\text{-value} < 0.05$ ) from Simulated Study 1 to be equivalently changed.

- a. The mean expression for a gene in the simulated control group is equal to the mean expression in Simulated Study 1.
- b. Equivalently changed genes are further simulated to not always be perfectly equivalently changed (even on average). This is done by having a modifier for the change in gene  $g$ , denoted  $\theta_g$  and is in  $[1, 2.5]$ .



**FIGURE 2**  
 Example volcano plots of the two simulated gene expression studies of one simulation iteration. Marked red are all genes set to be equivalent between the two studies. Here, the differential expression of the genes to be set equivalently changed between the two simulated studies must pass a threshold ( $\text{abs}(\log_2\text{FC}) > 1$  and  $p\text{-value} < 0.05$ ) in *Simulated Study 1* but the differential expression of the same genes is altered in *Simulated Study 2* to mirror equivalent change and thus can fail the threshold for differential gene expression.

c. For *Simulated Study 2*, the expression value for gene  $g$ , observation  $j$ , group  $k$  is then drawn from a truncated normal distribution:

$$y_{gik} \sim N(\gamma_g, s_{gk}); 0 < y_{gik}$$

With

$$\gamma_g = \begin{cases} m_g + k\theta_g\delta_g & \text{if } f = 1 \\ m_g + k\delta_g & \text{if } f = 2 \\ m_g & \text{if } f = 3 \end{cases}$$

We used the pairing of  $s_{gk}$  and  $\delta_g$  to be able to simulate the fact that differentially expressed genes with low difference in means between the treatment groups often have small standard deviations.

Figure 2 shows the distribution of differential gene expression of the two simulated studies for one iteration of the simulation. The degree of equivalence of change in gene expression between the two studies for one iteration of the simulation is visualized in Supplemental Figure S1.

## 2.6 Performance measures

To assess the performance of each test we used sensitivity, specificity, balanced accuracy and F1 score, which is the harmonic mean of precision and sensitivity.

**TABLE 1** Specifics for the different pancreatic cancer studies used in this project.

Tissue type	Accession #	# Tumor	# Control
Disease	GSE16515	36	16
Disease	GSE22780	8	8
PBMC	GSE74629	36	14

## 2.7 Biological data

The biological data analysis is based on publicly available pancreatic cancer microarray expression data, Alzheimer’s disease microarray expression data, and smoking RNAseq data obtained from the NCBI-GEO database (Edgar et al., 2002; Barrett et al., 2013) (Gene Expression Omnibus (GEO), RRID:SCR\_005012). Given that these datasets are drawn from similar populations, we expect to see equivalently changed genes among them.

For pancreatic cancer, one dataset was comprised of samples of peripheral blood mononuclear cells (PBMC) (accession number GSE74629). The other two datasets are comprised of samples of tumor tissue (tumor tissue data set 1: GSE16515, tumor tissue data set 2: GSE22780) (Pei et al., 2009; Ellsworth et al., 2013; Li et al., 2016) (Table 1). The control

data for the PBMC study were from gender, age, and habit matched healthy participants and the control data for the two disease tissue studies are from adjacent normal tissue of the cancer patients. After reading in the raw files, the count data were background corrected, quantile normalized, and the expression values were calculated using the function `rma()` from the package `affy` (Gautier et al., 2004) for Affymetrix microarray data, or the function `neqc()` from the package `limma` (Ritchie et al., 2015) for Illumina microarray data in the R statistical environment.

For Alzheimer's disease, the studies GSE1297 and GSE29378 were compared. In GSE1297 only severe cases vs. control were considered since GSE29378 only included advanced Alzheimer's disease cases. In GSE29378 we only used the CA1 regions, in order to match what was used for GSE1297. GSE1297 was background subtracted and normalized using the `rma()` function from the `affy` package and GSE29378 was background corrected and normalized using `neqc()` function from the `limma` package.

For the smoking data set, the study GSE169757 was used where smokers were compared to control and vapers. We used the function `cpm()` from the `edgeR` packet (Robinson et al., 2010; McCarthy et al., 2012) to filter expressed genes by a threshold of 0.3, which roughly corresponds to a count of 10. To normalize the data, we used the function `calcNormFactors()` from the `edgeR` package and `voom()` from the `limma` package.

## 2.8 Differential gene expression

We used the package `limma` to perform differential gene expression analysis. The model was fitted by using the functions `lmfit()` and `eBayes()`, which uses moderated t-statistics for ranking the genes. The standard deviation (sd) for gene  $g$  was extracted from the output by multiplying the square root of the unscaled covariance  $\vartheta_{gj}$  and the posterior residual variance  $\tilde{s}_g^2$ ,  $sd = \sqrt{\vartheta_{gj} \cdot \tilde{s}_g^2}$  as recommended by Smyth (2004).

## 2.9 Disease enrichment analysis

To investigate the association of equivalently changed genes with disease types we used the R package `DOSE` (Yu et al., 2015), which calculates a  $p$ -value using the hypergeometric distribution to determine whether the number of genes associated with a disease is larger than expected. We used the function `enrichDO()`, which supports Disease Ontology (DO) data (Schriml et al., 2012), and `enrichDGN`, which supports DisGeNET (Pinero et al., 2015) (DGN), with a minimal size of genes annotated by NCG category for testing of 5 and a  $q$ -value cut off of 0.05.

# 3 Results

## 3.1 Simulation

Simulation 1 was comprised of `Simulated Study 1` and `Simulated Study 2`. Each simulated study contained a simulated treatment and control. The two simulated studies were created such that 30% of the genes were equivalently changed with varying degrees of equivalence. [Figure 3](#) and [Table 2](#) show the decision of significance of the two tests (ECI bootstrap test and TOST) for one single iteration of the simulation. As can be observed, for a sample size of the comparator groups (e.g., case vs. control) of 20 people in each of the control group and case group, the ECI bootstrap test is able to identify the majority of equivalently changed genes (in this example sensitivity = 0.987) but also falsely identifies not equivalently changed genes (specificity <1). TOST on the other hand, does not misclassify any non-equivalently changed genes but identifies only 5 equivalently changed genes.

We repeated the previously described simulation 1,000 times and calculated the average of sensitivity, specificity, balanced accuracy, and F1 score for different sample sizes of the comparator groups, namely 5, 7, 10, 20, 50, and 100 ([Figure 4](#)). As can be seen, the sensitivity and F1 score for TOST are less than 0.2 and balanced accuracy is close to 0.5 irrespective of group size, and the specificity is close to 1 for all group sizes. The ECI bootstrap test shows that all performance metrics are dependent on the group size (as expected). Nevertheless, the performance of the ECI bootstrap test is overall very good. For a group size of 10, the balanced accuracy averages 0.940 and the F1 score averages 0.925 ([Supplemental Table S1](#)).

## 3.2 Biological data

In the following, the naïve approach is defined as declaring a gene equivalently changed when it is differentially expressed in the same direction in both studies. For the ECI bootstrap test and TOST we report a gene as significant when it differentially expressed in at least one of the two studies and has a  $q$ -value less than 0.05.

### 3.2.1 Pancreatic cancer

We used three publicly available data sets that were created to study pancreatic cancer. Two of the datasets contained gene expression measurements in pancreatic tumor tissue with adjacent normal tissue as control and the third data set contained gene expression measurements from peripheral blood mononuclear cells (PBMCs) with gene expression data from gender, age, and habits matched healthy patients as controls (hereafter referred to as the PBMC data set).

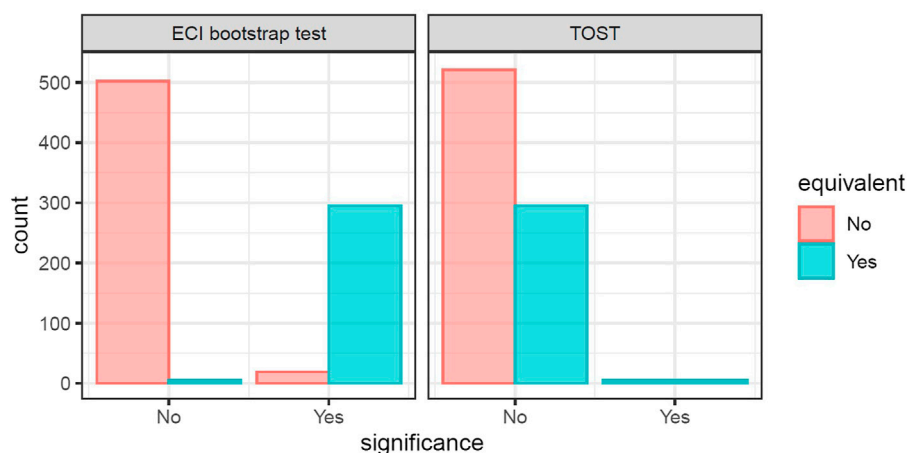


FIGURE 3

Example plots of decisions based on the q-value by the two tests of one simulation iteration. Left is the significance decision of the ECI bootstrap test, right is the significance decision of the TOST.

TABLE 2 Performance measurements for one single simulation with group size 20. Decisions of equivalence are made using the q-value of 0.05 for ECI bootstrap test and TOST. The TOST is too conservative to identify any equivalently changed genes, while the ECI bootstrap test shows good performance.

Performance	ECI bootstrap test	TOST
Sensitivity	0.987	0
Specificity	0.951	1
Balanced accuracy	0.969	0.5
F1-score	0.953	0

We performed bootstrap ECI test, TOST, and naïve approach on the biological data. Using TOST, none of the genes were significantly equivalently changed between the two tumor tissue data sets. When using the naïve approach, 150 genes were identified as equivalent. While using bootstrap ECI, the comparison of the two tumor tissue data sets led to the identification of 251 genes which were equivalently changed. From the 150 genes defined equivalent by the naïve approach, 127 were found equivalently changed using ECI bootstrap test (Supplemental Table S2). Significant genes identified by the ECI bootstrap test are enriched with annotations for pancreatic cancer ( $p\text{-adj} = 5.5e-04$ ), pancreatic carcinoma ( $p\text{-adj} = 3.4e-03$ ), and pancreatic adenocarcinoma ( $p\text{-adj} = 7.6e-03$ ) when using the DO database (Supplemental Table S7). When we use the DGN database, multiple disease descriptions were related to cancer such as adenocarcinoma of pancreas ( $p\text{-adj} = 2.0e-04$ ), or gastric adenocarcinoma ( $p\text{-adj} = 6.4e-03$ ) (Supplemental Table S8). The top 10 equivalently changed genes include some with known or suspected tumor relevance such as

CYP3A5 (Noll et al., 2016), STEAP1 (Gomes et al., 2012), and PROX1 (Saukkonen et al., 2016), which are being investigated as drug targets for pancreatic cancer, GMNN (Kim et al., 2012; Kushwaha et al., 2016), which is associated with cancer pathophysiology and development, and SORBS2 (Alsafadi et al., 2011; Zhao et al., 2018; Lv et al., 2020; Lv et al., 2021), which is known to be associated with metastatic relapse.

Using TOST we identified two equivalently changed genes (SPICE1 and MCEMP1) and one inverse changed gene (STAMBPL1) by comparing tumor tissue data set 1 with the PBMC data set. When using the naïve approach, 14 genes were found equivalently changed. When using the ECI bootstrap test, we identified 181 equivalently changed genes and 94 inversely changed genes and all genes equivalently changed using the naïve approach were also identified as equivalently changed using the ECI bootstrap test (Supplemental Table S3). The inversely changed genes were not found to be associated with diseases listed in either the DO database or the DGN database. Significant equivalently changed genes identified by the ECI bootstrap test are enriched with annotations for pancreatic cancer ( $p\text{-adj} = 3.2e-03$ ), pancreatic carcinoma ( $p\text{-adj} = 2.1e-02$ ), and pancreatic ductal adenocarcinoma ( $p\text{-adj} = 1.3e-02$ ) when using the DO database (Supplemental Table S9). When using the DGN database, multiple disease descriptions related to pancreatic cancer were identified such as pancreatic ductal adenocarcinoma ( $p\text{-value} = 1.3e-05$ ), and adenocarcinoma of pancreas ( $p\text{-value} = 2.3e-03$ ) (Supplemental Table S10).

The comparison of tumor tissue data set 2 with the data set from PBMC led to no equivalently changed genes being identified when using TOST, two equivalently changed genes identified when using naïve approach, and we identified nine equivalently changed genes and eight inversely changed genes when using the



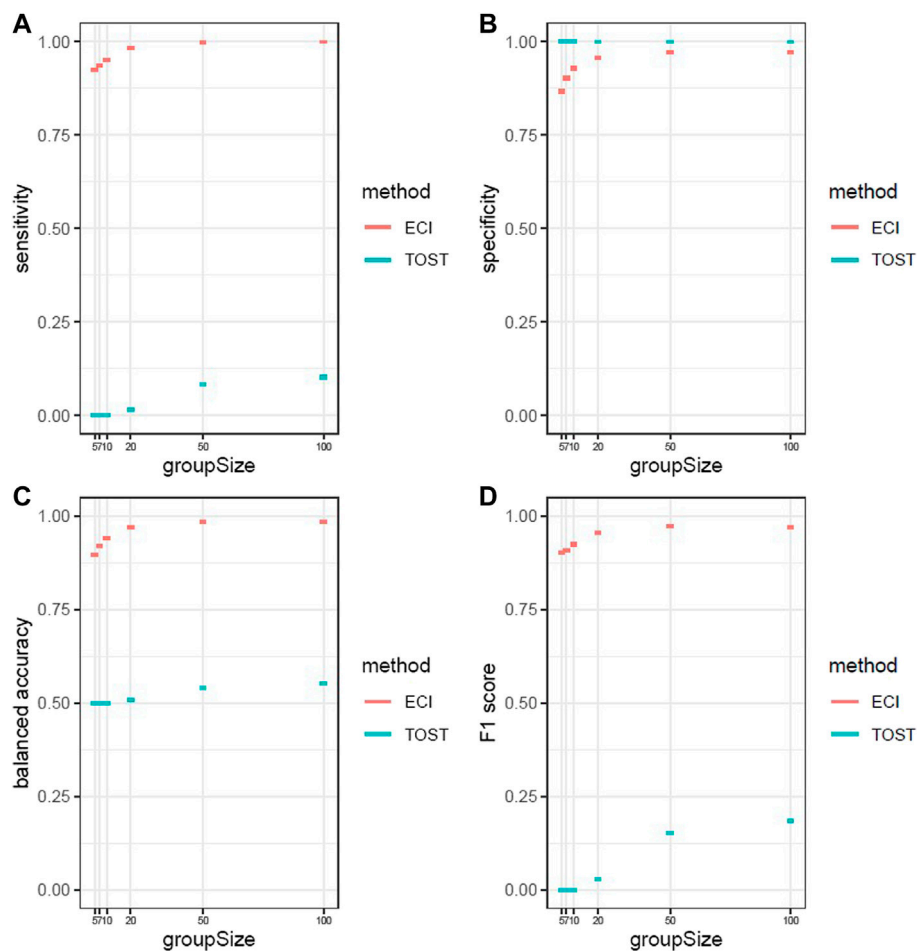


FIGURE 4

Performance metrics with 95% confidence intervals of 1,000 iterations for ECI bootstrap test and two one-sided  $t$ -test (A) Sensitivity (B) specificity (C) balanced accuracy (D) F1 score. The ECI bootstrap test shows increasing performance with increasing group size and has overall a good performance. The TOST shows only in specificity a better performance than the ECI bootstrap test and shows no performance for all the other measurements.

ECI bootstrap test. The two genes identified as equivalently changed when using the naïve approach were also significantly equivalently changed using the ECI bootstrap test (Supplemental Table S4). Using the DO database as well as the DGN database resulted in no disease descriptions found to be associated with the equivalently changed genes, likely due to the small number identified.

### 3.2.2 Alzheimer's disease

We compared two publicly available data sets for Alzheimer's disease, where 147 genes were differentially expressed in at least one of the two studies. We found no genes to be equivalently changed using TOST and three genes identified as equivalently changed using the naïve approach. Using the ECI bootstrap test we found 20 genes to be equivalently changed, where two were differentially expressed in both studies (ANO3 and SST), and

1 gene to be inverse changed (Supplemental Table S5). The equivalently changed genes are enriched with an annotation for Non-Functioning Pituitary Gland Neoplasm ( $p_{adj} = 0.039$ ) when using the DGN database (Supplemental Table S11), where down-regulation of the Pituitary Gland is associated with Alzheimer's disease (Ishii and Iadecola, 2015).

The top equivalently changed genes identified by the ECI bootstrap test are: SST, FXYD7, PCP4, RASL12, SLC14A1, INA, SERPINA3, ANO3, APLNR, AEBP1. SST is directly linked to AD (SolarSKI et al., 2018). FXYD7, APLNR, INA, AEBP1 and SERPINA3 were found to be differentially expressed in a meta-analysis of AD data sets (Su et al., 2019). PCP4 was found to be dysregulated in the forebrain of mice with AD (Renelt et al., 2014). SLC14A1 is one of significantly altered transcripts in APP<sup>swe</sup>/PS1<sup>dE9</sup> transgenic mice during the development of beta amyloid protein ( $A\beta$ ) pathology (Wirz

et al., 2013). Finally, ANO3 is significantly differentially expressed in relation to Alzheimer's disease (Vastrand and Vastrand, 2021).

### 3.2.3 Smoking vs. vaping

We used one publicly available data set of a study of three groups: smoking, vaping, control. We compared the control group against both smoking and vaping for differential gene expression and finally compared the two sub-studies (smoking vs. control and vaping vs. control) using the TOST and ECI bootstrap test. No genes were found to be equivalently changed using the TOST and 382 genes were found to be equivalently changed using the naïve approach. Using the ECI bootstrap test we found 1,090 genes to be equivalently changed and 0 to be inversely changed (Supplemental Table S6). Out the 1,090 genes, 380 genes were differentially expressed in both studies, 404 were protein coding genes and 348 could be used for disease enrichment analysis because they had an entrez id. The disease enrichment analysis delivered 0 associated diseases for both the DO database and the DGN database.

The top 10 equivalently changed genes are: TERB2, TM4SF1, PHKA1, GLYATL1, AC073610.2, PF4, HECW1, CALD1, APOBEC4, NPHP3-ACAD11. TERB2 is associated with Spermatogenic Failure 59 (GeneCards, 2022). (Dai et al., 2015) found that smoking is associated with impaired spermatogenesis. (Fu et al., 2020). found TM4SF1 to be associated with smoking. PHKA1, an inflammatory-associated transcript, is significantly altered with prenatal nicotine exposure (Zhou et al., 2021). GLYATL1 is associated with HCC (Guan et al., 2020). PF4 is found to be significantly higher expressed when smoking (Zevin et al., 2001). CALD1 is associated with *in utero* tobacco smoke exposure (Breton et al., 2014). And two SNPs nearest to APOBEC4 were associated with fluorescent oxidation products (FIOPs) accounting for tobacco smoking status in adults (Orsi et al., 2022).

## 4 Discussion

In this work, we compared two options to test for equivalently changed genes between two studies, namely the proposed ECI bootstrap test and TOST. We were able to show that the ECI bootstrap performs well in identifying equivalently changed genes with respect to balanced accuracy and F1, which were on average both over 0.9, even for the smallest analyzed group size of ten, while maintaining a high specificity. Furthermore, we found that TOST greatly underperformed the ECI bootstrap test with balanced accuracy close to 0.5 and F1 score close to 0.

By using a threshold of 0.05 for the q-value, for the ECI bootstrap test, we were able to identify many equivalently changed genes between two pancreatic cancer tumor tissue studies where most of the genes were related to cancer progression. When

comparing two Alzheimer's data sets, the majority of the equivalently changed genes were associated with Alzheimer's disease. The lack of significant enrichment for Alzheimer's disease itself could simply be due to non-comprehensive annotation in the databases used. Additionally, when we were comparing smoking to vaping most of the top equivalently changed genes were related to smoking, though the enrichment analysis delivered no associated disease. This is not surprising, given that vaping is not necessarily expected to lead to similar health outcomes as smoking. Nevertheless, the results show that the ECI bootstrap test met our expectation of identifying equivalently changed genes in studies of the same disease type or exposure and, furthermore, is able to identify equivalently changed genes which are functionally related to the disease type or exposure and could lead to more reproducible or robust results. Additionally, we were able to show that several genes in non-cancer tissue (peripheral blood) of patients of pancreatic cancer showed equivalently changed behavior to genes of tumor tissue of the same disease which implies the systemic impact of cancer. Those results may open the way for identifying reliable blood markers for cancer, and for new investigative approaches into the field of systemic changes of gene expression of cancer patients that may be addressed in future studies. It must be mentioned that the difference in the number of significant genes could be due to the group sizes of the data sets.

When comparing the ECI bootstrap test to the naïve approach we found several genes which were declared equivalently changed using the naïve approach to be not significant using the ECI bootstrap test. In particular, 23 genes out of 150 genes differentially expressed in both pancreatic tumor tissue studies were not significantly equivalently changed when using the ECI bootstrap test. This shows that simply identifying genes as equivalently changed because they are differentially expressed in the same direction might lead to false conclusions. On the other hand, in the comparison of vaping vs. smoking only 2 out of 382 genes differentially expressed in both studies were not significantly equivalently changed using the ECI bootstrap test. One reason for that could be an inherent bias due to using the same control as comparison. Additionally, the ECI bootstrap test is more powerful because it can identify statistically equivalently changed genes even when they are not statistically significantly changed in one of the studies. Furthermore, although we have demonstrated the situation in which we test for any degree of equivalent change, it is just as simple to test for the desired level of equivalent change, which the naïve method cannot do. Overall, the ECI bootstrap test gives more reliable results when identifying equivalently changed genes in comparison to the naïve approach.

Limitations of this study include the differing nature of cases and controls in some datasets, the different group sizes, and the lack of a ground truth in biological data. In future work, the plan is to create a framework for power calculation to address the influence of different group sizes on the ECI bootstrap test performance.

## 5 Conclusion

In this study we demonstrated the use of the ECI bootstrap test in a setting of differentially expressed genes to provide researchers with a statistical approach to identify genes which are similarly influenced by a specific treatment or exposure. Furthermore, statistically identified equivalently changed genes reduces the cost for validating those genes and offers the option of identifying new possible treatment targets.

In addition, it is also possible, due to the non-parametric nature of the bootstrap test and the lack of assumptions on the ECI value, to adapt the ECI bootstrap test to other options such as methylation data or other types of 'omics data. In future studies we want to investigate the effectiveness of bootstrap ECI on other types of data sets.

Furthermore, we have created an R package for the ECI bootstrap test which can be obtained from the github repository at <https://github.com/Hecate08/ECIbootstrap>.

## Data availability statement

The datasets analyzed in this study can be found in the NCBI repository <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE16515>, <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE22780>, and <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE74629>.

## Author contributions

LN and JT created the new ECI bootstrap test and developed the design of the simulation study. LN analysed the data and wrote the manuscript. JT interpreted the biological data. All authors read, reviewed, and approved the final manuscript.

## Funding

This work was supported by a CTSA grant from NCATS awarded to the University of Kansas for Frontiers: University

## References

- Alsafadi, S., Scott, V., Pautier, P., Goubar, A., Lazar, V., Dessen, P., et al. (2011). P5-01-07: Identification of SORBS2 as a candidate marker to predict metastatic relapse in breast cancer. *Cancer Res.* 71, P5-01-P07. doi:10.1158/0008-5472.sabcs11-p5-01-07
- Barrett, T., Wilhite, S. E., Ledoux, P., Evangelista, C., Kim, I. F., Tomashevsky, M., et al. (2013). NCBI GEO: archive for functional genomics data sets--update. *Nucleic Acids Res.* 41, D991–D995. doi:10.1093/nar/gks1193
- Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate - a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* 57 (1), 289–300. doi:10.1111/j.2517-6161.1995.tb02031.x
- Blake, L. E., Roux, J., Hernando-Herraez, I., Banovich, N. E., Perez, R. G., Hsiao, C. J., et al. (2020). A comparison of gene expression and DNA methylation

of Kansas Clinical and Translational Science Institute (# UL1TR002366) and used the Quantitative 'Omics Core (QOC) and the National Cancer Institute (NCI) Cancer Center Support Grant P30CA168524, the Kansas IDeA Network of Biomedical Research Excellence Bioinformatics Core, supported by the National Institute of General Medical Science award P20 GM103418, and the Kansas Institute for Precision Medicine COBRE, supported by the National Institute of General Medical Science award P20 GM130423.

## Acknowledgments

We acknowledge the contribution of Mihaela Sardu, Lynn Chollet Hinton, Nanda Kumar Yellapu, Emily Nissen, Whitney Shae, and Jonah Aponsah in reviewing the paper.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fbinf.2022.893032/full#supplementary-material>

patterns across tissues and species. *Genome Res.* 30 (2), 250–262. doi:10.1101/gr.254904.119

Breton, C. V., Siegmund, K. D., Joubert, B. R., Wang, X., Qui, W., Carey, V., et al. (2014). Prenatal tobacco smoke exposure is associated with childhood DNA CpG methylation. *PLoS One* 9 (6), e99716. doi:10.1371/journal.pone.0099716

Dai, J. B., Wang, Z. X., and Qiao, Z. D. (2015). The hazardous effects of tobacco smoking on male fertility. *Asian J. Androl.* 17 (6), 954–960. doi:10.4103/1008-682x.150847

Dixon, P. M., Saint-Maurice, P. F., Kim, Y., Hibbing, P., Bai, Y., and Welk, G. J. (2018). A primer on the use of equivalence testing for evaluating measurement

- agreement. *Med. Sci. Sports Exerc.* 50 (4), 837–845. doi:10.1249/mss.0000000000001481
- Edgar, R., Domrachev, M., and Lash, A. E. (2002). Gene expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* 30 (1), 207–210. doi:10.1093/nar/30.1.207
- Ellsworth, K. A., Eckloff, B. W., Li, L., Moon, I., Fridley, B. L., Jenkins, G. D., et al. (2013). Contribution of FKBP5 genetic variation to gemcitabine treatment and survival in pancreatic adenocarcinoma. *Plos One* 8 (8), e70216. doi:10.1371/journal.pone.0070216
- Fu, F., Yang, X., Zheng, M., Zhao, Q., Zhang, K., Li, Z., et al. (2020). Role of transmembrane 4 L six family 1 in the development and progression of cancer. *Front. Mol. Biosci.* 7, 202. doi:10.3389/fmolb.2020.00202
- Fu, Y., Thomas, A., Gasior, D., Harper, J., Gay, A., Jones, C., et al. (2021). A comparison of shared patterns of differential gene expression and gene ontologies in response to water-stress in roots and leaves of four diverse genotypes of *Lolium* and *Festuca* spp. temperate pasture grasses. *Plos One* 16 (4), e0249636. doi:10.1371/journal.pone.0249636
- Garrido, J., Aguilar, M., and Prieto, P. (2020). Identification and validation of reference genes for RT-qPCR normalization in wheat meiosis. *Sci. Rep.* 10 (1), 2726. doi:10.1038/s41598-020-59580-5
- Gautier, L., Cope, L., Bolstad, B. M., and Irizarry, R. A. (2004). affy---analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* 20 (3), 307–315. doi:10.1093/bioinformatics/btg045
- GeneCards. TERB2 gene - telomere repeat binding bouquet formation protein 2. 2022 cited 2022.
- Gilbert, S. (2000). *Developmental biology*. 6th edition. Sunderland (MA): Sinauer Associates.
- Goh, W. W. B., Wang, W., and Wong, L. (2017). Why batch effects matter in omics data, and how to avoid them. *Trends Biotechnol.* 35 (6), 498–507. doi:10.1016/j.tibtech.2017.02.012
- Gomes, I. M., Maia, C. J., and Santos, C. R. (2012). STEAP proteins: from structure to applications in cancer therapy. *Mol. Cancer Res.* 10 (5), 573–587. doi:10.1158/1541-7786.mcr-11-0281
- Guan, R., Hong, W., Huang, J., Peng, T., Zhao, Z., Lin, Y., et al. (2020). The expression and prognostic value of GLYATL1 and its potential role in hepatocellular carcinoma. *J. Gastrointest. Oncol.* 11 (6), 1305–1321. doi:10.21037/jgo-20-186
- Hollenbach, P. W., Nguyen, A. N., Brady, H., Williams, M., Ning, Y., Richard, N., et al. (2010). A comparison of azacitidine and decitabine activities in acute myeloid leukemia cell lines. *Plos One* 5 (2), e9001. doi:10.1371/journal.pone.0009001
- Ishii, M. I. C., and Iadecola, C. (2015). Metabolic and non-cognitive manifestations of Alzheimer's disease: the hypothalamus as both culprit and target of pathology. *Cell Metab.* 22 (5), 761–776. doi:10.1016/j.cmet.2015.08.016
- Jung, K., Lee, J., Gupta, V., and Cho, G. (2019). Comparison of bootstrap confidence interval methods for GSCA using a Monte Carlo simulation. *Front. Psychol.* 10, 2215. doi:10.3389/fpsyg.2019.02215
- Kim, H. E., Kim, D. G., Lee, K. J., Son, J. G., Song, M. Y., Park, Y. M., et al. (2012). Frequent amplification of CENPF, GMNN and CDK13 genes in hepatocellular carcinomas. *Plos One* 7 (8), e43223. doi:10.1371/journal.pone.0043223
- Kushwaha, P. P., Rapalli, K. C., and Kumar, S. (2016). Geminin a multi task protein involved in cancer pathophysiology and developmental process: A review. *Biochimie* 131, 115–127. doi:10.1016/j.biochi.2016.09.022
- Leichsenring, F., Abbass, A., Driessen, E., Hilsenroth, M., Luyten, P., Rabung, S., et al. (2018). Equivalence and non-inferiority testing in psychotherapy research. *Psychol. Med.* 48 (11), 1917–1919. doi:10.1017/s0033291718001289
- Li, L., Zhang, J. W., Jenkins, G., Xie, F., Carlson, E. E., Fridley, B. L., et al. (2016). Genetic variations associated with gemcitabine treatment outcome in pancreatic cancer. *Pharmacogenet. Genomics* 26 (12), 527–537. doi:10.1097/fpc.0000000000000241
- Lv, Q., Dong, F., Zhou, Y., Cai, Z., and Wang, G. (2020). RNA-binding protein SORBS2 suppresses clear cell renal cell carcinoma metastasis by enhancing MTUS1 mRNA stability. *Cell Death Dis.* 11 (12), 1056. doi:10.1038/s41419-020-03268-1
- Lv, Q., Dong, F., Zhou, Y., Cai, Z., and Wang, G. (2021). Correction to: RNA-binding protein SORBS2 suppresses clear cell renal cell carcinoma metastasis by enhancing MTUS1 mRNA stability. *Cell Death Dis.* 12 (11), 1062. doi:10.1038/s41419-021-03496-z
- McCarthy, D. J., Chen, Y., and Smyth, G. K. (2012). Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* 40, 4288–4297. doi:10.1093/nar/gks042
- Millard, S. P. (2013). *EnvStats: An R package for environmental statistics*. New York: Springer.
- Noll, E. M., Eisen, C., Stenzinger, A., Espinet, E., Muckenhuber, A., Klein, C., et al. (2016). CYP3A5 mediates basal and acquired therapy resistance in different subtypes of pancreatic ductal adenocarcinoma. *Nat. Med.* 22 (3), 278–287. doi:10.1038/nm.4038
- Orsi, L., Margaritte-Jeannin, P., Andrianjafimasy, M., Dumas, O., Mohamdi, H., Bouzigon, E., et al. (2022). Genome-wide association study of fluorescent oxidation products accounting for tobacco smoking status in adults from the French EGEA study. *Antioxidants (Basel)* 11 (5), 802. doi:10.3390/antiox11050802
- Pei, H., Li, L., Fridley, B. L., Jenkins, G. D., Kalari, K. R., Lingle, W., et al. (2009). FKBP51 affects cancer cell response to chemotherapy by negatively regulating Akt. *Cancer Cell* 16 (3), 259–266. doi:10.1016/j.ccr.2009.07.016
- Pinero, J., Queralt-Rosinach, N., Bravo, A., Deu-Pons, J., Bauer-Mehren, A., Baron, M., et al. (2015). DisGeNET: a discovery platform for the dynamical exploration of human diseases and their genes. *Database-Oxford* 2015, bav028. doi:10.1093/database/bav028
- Provenzano, M., and Mocellin, S. (2007). Complementary techniques: validation of gene expression data by quantitative real time PCR. *Adv. Exp. Med. Biol.* 593, 66–73. doi:10.1007/978-0-387-39978-2\_7
- Qiu, J., and Cui, X. (2010). Evaluation of a statistical equivalence test applied to microarray data. *J. Biopharm. Stat.* 20 (2), 240–266. doi:10.1080/10543400903572738
- Renelt, M., von Bohlen und Halbach, V., and von Bohlen und Halbach, O. (2014). Distribution of PCP4 protein in the forebrain of adult mice. *Acta Histochem.* 116 (6), 1056–1061. doi:10.1016/j.acthis.2014.04.012
- Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43 (7), e47. doi:10.1093/nar/gkv007
- Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. doi:10.1093/bioinformatics/btp1616
- Rudy, J., and Valafar, F. (2011). Empirical comparison of cross-platform normalization methods for gene expression data. *BMC Bioinforma.* 12, 467. doi:10.1186/1471-2105-12-467
- Saukkonen, K., Hagstrom, J., Mustonen, H., Juuti, A., Nordling, S., Kallio, P., et al. (2016). PROX1 and beta-catenin are prognostic markers in pancreatic ductal adenocarcinoma. *BMC Cancer* 16, 472. doi:10.1186/s12885-016-2497-5
- Schriml, L. M., Arze, C., Nadendla, S., Chang, Y. W., Mazaitis, M., Felix, V., et al. (2012). Disease ontology: a backbone for disease semantic integration. *Nucleic Acids Res.* 40, D940–D946. doi:10.1093/nar/gkr972
- Schuurmann, D. J. (1987). A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. *J. Pharmacokinet. Biopharm.* 15 (6), 657–680. doi:10.1007/bf01068419
- Shen, Y., Wolkowicz, M. J., Kotova, T., Fan, L., and Timko, M. P. (2016). Transcriptome sequencing reveals e-cigarette vapor and mainstream-smoke from tobacco cigarettes activate different gene expression profiles in human bronchial epithelial cells. *Sci. Rep.* 6, 23984. doi:10.1038/srep23984
- Smyth, G. K. (2004). Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.* 2004, 3. doi:10.2202/1544-6115.1027
- Solarski, M., Wang, H., Wille, H., and Schmitt-Ulms, G. (2018). Somatostatin in alzheimer's disease: A new role for an old player. *Prion* 12 (1), 1–8. doi:10.1080/19336896.2017.1405207
- Su, L., Chen, S., Zheng, C., Wei, H., and Song, X. (2019). Meta-analysis of gene expression and identification of biological regulatory mechanisms in alzheimer's disease. *Front. Neurosci.* 13, 633. doi:10.3389/fnins.2019.00633
- Thompson, J. A., and Koestler, D. C. (2020). Equivalent change enrichment analysis: assessing equivalent and inverse change in biological pathways between diverse experiments. *Bmc Genomics* 21 (1), 180. doi:10.1186/s12864-020-6589-x
- Tihomir Asparouhov aBMe. Bootstrap P-value computation 2021 Available from: <http://www.statmodel.com/download/FAQ-Bootstrap%20-%20Pvalue.pdf>.
- Vastrad, B., and Vastrad, C., Bioinformatics analyses of significant genes, related pathways and candidate prognostic biomarkers in Alzheimer's disease. 2021. bioRxiv preprint.
- Walsh, C. J., Hu, P., Batt, J., and Santos, C. C. (2015). Microarray meta-analysis and cross-platform normalization: Integrative genomics for robust biomarker discovery. *Microarrays (Basel)* 4 (3), 389–406. doi:10.3390/microarrays4030389
- Wicklin, R. The bias-corrected and accelerated (BCa) bootstrap interval: SAS blogs; 2017 Available from: <https://blogs.sas.com/content/iml/2017/07/12/bootstrap-bca-interval.html>.
- Wirz, K. T., Bossers, K., Stargardt, A., Kamphuis, W., Swaab, D. F., Hol, E. M., et al. (2013). Cortical beta amyloid protein triggers an immune response, but no

synaptic changes in the APPswe/PS1dE9 Alzheimer's disease mouse model. *Neurobiol. Aging* 34 (5), 1328–1342. doi:10.1016/j.neurobiolaging.2012.11.008

Wu, L. J., Gander, P. H., van den Berg, M., and Signal, T. L. (2018). Equivalence testing as a tool for fatigue risk management in aviation. *Aerosp. Med. Hum. Perform.* 89 (4), 383–388. doi:10.3357/amhp.4790.2018

Yekutieli, D., and Benjamini, Y. (1999). Resampling-based false discovery rate controlling multiple test procedures for correlated test statistics. *J. Stat. Plan. Inference* 82 (1-2), 171–196. doi:10.1016/s0378-3758(99)00041-5

Yu, G., Wang, L. G., Yan, G. R., and He, Q. Y. (2015). DOSE: an R/bioconductor package for disease ontology semantic and enrichment analysis. *Bioinformatics* 31 (4), 608–609. doi:10.1093/bioinformatics/btu684

Zevin, S., Saunders, S., Gourlay, S. G., Jacob, P., and Benowitz, N. L. (2001). Cardiovascular effects of carbon monoxide and cigarette smoking. *J. Am. Coll. Cardiol.* 38 (6), 1633–1638. doi:10.1016/s0735-1097(01)01616-3

Zhao, L., Wang, W., Huang, S., Yang, Z., Xu, L., Yang, Q., et al. (2018). The RNA binding protein SORBS2 suppresses metastatic colonization of ovarian cancer by stabilizing tumor-suppressive immunomodulatory transcripts. *Genome Biol.* 19 (1), 35. doi:10.1186/s13059-018-1412-6

Zhou, L., Tao, X., Pang, G., Mu, M., Sun, Q., Liu, F., et al. (2021). Maternal nicotine exposure alters hippocampal microglia polarization and promotes anti-inflammatory signaling in juvenile offspring in mice. *Front. Pharmacol.* 12, 661304. doi:10.3389/fphar.2021.661304