

# **Measuring the Magnitude of Health Inequality between Two Population Subgroup Proportions**

Makram Talih, Ramal Moonesinghe, and David T. Huang

Web Appendix 1

Web Figure 1

Web Tables 1–9

## Web Appendix 1

Measures of inequality between proportions based on statistical effect size

*Absolute difference.* The absolute difference is the simplest (unstandardized) measure of effect size:

$$a(p_1, p_2) = |p_1 - p_2|$$

The absolute difference can be expressed as a function of the two ratios  $R_{12}=p_1/p_2$  and  $\check{R}_{21}=q_2/q_1$  for  $q_1=1-p_1$  and  $q_2=1-p_2$ . Indeed,  $p_1 = (\check{R}_{21} - 1)R_{12}/(R_{12}\check{R}_{21} - 1)$ ,  $p_2 = (\check{R}_{21} - 1)/(R_{12}\check{R}_{21} - 1)$ , and

$$a(R_{12}, \check{R}_{21}) = \frac{(R_{12} - 1)(\check{R}_{21} - 1)}{|R_{12}\check{R}_{21} - 1|}$$

*Standardized absolute difference.* One standardized measure of effect size is the standardized absolute difference, where the difference is scaled by a standard deviation that is assumed common to both population subgroups.

Under the assumptions of independence between groups and equal group sizes,

$$D_1(p_1, p_2) = \frac{|p_1 - p_2|}{\sqrt{(p_1 q_1 + p_2 q_2)/2}}$$

Expressed as a function of the two ratios  $R_{12}=p_1/p_2$  and  $\check{R}_{21}=q_2/q_1$ , the standardized absolute difference is

$$D_1(R_{12}, \check{R}_{21}) = \sqrt{\frac{(R_{12} - 1)(\check{R}_{21} - 1)}{(R_{12} + \check{R}_{21})/2}}$$

*Standardized absolute difference with pooled variance.* Another standardized absolute difference is:

$$D_2(p_1, p_2) = \frac{|p_1 - p_2|}{2\sqrt{p_* q_*}}$$

In the above,  $p_*= (p_1+p_2)/2$  is the so-called “pooled” proportion, obtained from the (equally-weighted) average of  $p_1$  and  $p_2$ ; and  $q_*=1-p_*= (q_1+q_2)/2$ . Twice the standardized absolute difference  $D_2$  appears in classical tests of the null hypothesis that the two proportions are equal. Written as a function of the two ratios  $R_{12}=p_1/p_2$  and  $\check{R}_{21}=q_2/q_1$ , the standardized absolute difference  $D_2$  is given by

$$D_2(R_{12}, \check{R}_{21}) = \sqrt{\frac{(R_{12} - 1)(\check{R}_{21} - 1)}{(R_{12} + 1)(\check{R}_{21} + 1)}}$$

*Rescaled absolute arcsine difference.* Cohen's index of effect size for the difference in proportions is a scalar multiple of the absolute difference in the arcsine-transformed square root proportions (1, 2):

$$h(p_1, p_2) = \frac{2}{\pi} |\arcsin(\sqrt{p_1}) - \arcsin(\sqrt{p_2})|$$

The scaling here differs from the specification of Cohen's index by the factor  $1/\pi$ ; this is so that  $h_{\max}=1$  here.

While a simple explicit formula is difficult to write down, the rescaled absolute arcsine difference  $h(p_1, p_2)$  can be expressed implicitly in terms of  $R_{12}=p_1/p_2$  and  $\check{R}_{21}=q_2/q_1$ , since both  $p_1$  and  $p_2$  are functions of these two ratios.

*Absolute logit difference.* The logit transformation  $\text{logit}(p)=p/(1-p)$  is such that  $\text{logit}(0)=-\infty$  and  $\text{logit}(1)=+\infty$ . The absolute difference of the logit-transformed proportions also measures effect size:

$$\ell(p_1, p_2) = |\text{logit}(p_1) - \text{logit}(p_2)|$$

The absolute logit difference equals the absolute value of the logarithm of the odds ratio (OR):

$$\ell(p_1, p_2) = \left| \ln \left( \frac{p_1/p_2}{q_1/q_2} \right) \right| = |\ln(\text{OR})|$$

Expressed as a function of the two ratios  $R_{12}=p_1/p_2$  and  $\check{R}_{21}=q_2/q_1$ , the logit difference is simply

$$\ell(R_{12}, \check{R}_{21}) = |\ln R_{12} + \ln \check{R}_{21}|$$

*Absolute probit difference.* The probit transformation  $\text{probit}(p_j)=\Phi^{-1}(p_j)$  finds the standard normal quantile corresponding to a given proportion. By construction,  $\text{probit}(0)=-\infty$  and  $\text{probit}(1)=+\infty$ . The absolute difference of the probit-transformed proportions also measures effect size:

$$b(p_1, p_2) = |\text{probit}(p_1) - \text{probit}(p_2)|$$

As with the rescaled absolute arcsine difference, while a simple explicit formula is difficult to write down, the absolute probit difference can be expressed as an implicit function of the two ratios  $R_{12}=p_1/p_2$  and  $\check{R}_{21}=q_2/q_1$ .

Measures of inequality between proportions based on information-theoretic divergence

*Measures based on chi-squared divergence.* The Pearson chi-squared statistic tests association (or goodness-of-fit) in  $r \times c$  contingency tables (3). The rationale for chi-squared divergence is similar to that for the chi-squared statistic. On the one hand,  $P_2=(p_2, q_2)$  are looked upon as the cell probabilities that would be expected for group 1 in a  $1 \times 2$  contingency table, where the probabilities  $P_1=(p_1, q_1)$  have been “observed.” Chi-squared divergence is:

$$\chi^2(P_1, P_2) = \frac{(p_1 - p_2)^2}{p_2} + \frac{(q_1 - q_2)^2}{q_2}$$

Recognizing that  $p_1 - p_2 = q_2 - q_1$ , chi-squared divergence is written as follows:

$$\chi^2(P_1, P_2) = (p_1 - p_2)^2 \left[ \frac{1}{p_2} + \frac{1}{q_2} \right]$$

Chi-squared divergence is not symmetric in  $P_1$  and  $P_2$ , because interchanging  $P_1$  and  $P_2$ , above, yields:

$$\chi^2(P_2, P_1) = (p_1 - p_2)^2 \left[ \frac{1}{p_1} + \frac{1}{q_1} \right]$$

However, the two chi-squared divergences  $\chi^2(P_1, P_2)$  and  $\chi^2(P_2, P_1)$  can be combined to obtain a doubly symmetric measure of inequality, as shown next; see also (4, 5).

Symmetrized chi-squared measure: The symmetrized chi-squared measure is an average of  $\chi^2(P_1, P_2)$  and  $\chi^2(P_2, P_1)$ :

$$X(P_1, P_2) = \frac{1}{2} [\chi^2(P_1, P_2) + \chi^2(P_2, P_1)] = \frac{1}{2} (p_1 - p_2)^2 \left[ \frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{q_1} + \frac{1}{q_2} \right]$$

Written as a function of the two ratios  $R_{12}=p_1/p_2$  and  $\check{R}_{21}=q_2/q_1$ , the symmetrized chi-squared measure is given by

$$X(R_{12}, \check{R}_{21}) = \frac{1}{2} (R_{12} - 1)(\check{R}_{21} - 1) \left( \frac{1}{R_{12}} + \frac{1}{\check{R}_{21}} \right)$$

Triangle discrimination measure: Another way to symmetrize chi-squared divergence is to use the arithmetic average  $P^*=(P_1+P_2)/2$  as the reference in constructing the measure. Thus, the probabilities  $P^*=(p^*,q^*)$  are seen as the cell probabilities that would be expected for population subgroup  $j$  in a  $1 \times 2$  contingency table where the probabilities  $P_j=(p_j,q_j)$  have been “observed.” The resulting measure is the triangle discrimination measure:

$$\Delta(P_1, P_2) = \frac{1}{2} [\chi^2(P_1, P^*) + \chi^2(P_2, P^*)] = \frac{1}{2} (p_1 - p_2)^2 \left[ \frac{1}{p_1 + p_2} + \frac{1}{q_1 + q_2} \right]$$

Note that the measure  $\Delta$  is the square of the standardized mean difference with pooled variance  $D'$ :

$$\Delta(P_1, P_2) = [D_2(p_1, p_2)]^2$$

Written as a function of the two ratios  $R_{12}=p_1/p_2$  and  $\check{R}_{21}=q_2/q_1$ , the triangle discrimination measure is given by

$$\Delta(R_{12}, \check{R}_{21}) = \left( \frac{R_{12} - 1}{R_{12} + 1} \right) \left( \frac{\check{R}_{21} - 1}{\check{R}_{21} + 1} \right)$$

*Measures based on Kullback-Leibler divergence.* Kullback-Leibler (KL) divergence evaluates the weighted average log ratios between the cell probabilities from the two probability distributions  $P_1$  and  $P_2$ . Either of the two distributions can be used to determine the population subgroup weights. Here:

$$KL(P_1, P_2) = p_1 \ln \left[ \frac{p_1}{p_2} \right] + q_1 \ln \left[ \frac{q_1}{q_2} \right]$$

$$KL(P_2, P_1) = p_2 \ln \left[ \frac{p_2}{p_1} \right] + q_2 \ln \left[ \frac{q_2}{q_1} \right]$$

That the two KL divergences are nonnegative for any pair of probability distributions  $P_1$  and  $P_2$  is well established; it follows from Jensen’s inequality and the convexity of the function  $-\ln(x)$  (6, 7, 8).

There are various ways of combining the two KL divergences  $KL(P_1, P_2)$  and  $KL(P_2, P_1)$  to obtain a doubly symmetric measure of inequality. Two such measures are presented next.

Jeffreys divergence: The Jeffreys divergence is the arithmetic average of  $KL(P_1, P_2)$  and  $KL(P_2, P_1)$  (7, 4, 5):

$$J(P_1, P_2) = \frac{1}{2} [KL(P_1, P_2) + KL(P_2, P_1)] = \frac{1}{2} \left\{ (p_1 - p_2) \ln \left[ \frac{p_1}{p_2} \right] + (q_1 - q_2) \ln \left[ \frac{q_1}{q_2} \right] \right\}$$

Simplifying this expression:

$$J(P_1, P_2) = \frac{1}{2} (p_1 - p_2) \times \ln(\text{OR})$$

Written as a function of the two ratios  $R_{12}=p_1/p_2$  and  $\check{R}_{21}=q_2/q_1$ , Jeffreys divergence measure is given by

$$J(R_{12}, \check{R}_{21}) = \frac{1}{2} \frac{(R_{12} - 1)(\check{R}_{21} - 1)}{R_{12}\check{R}_{21} - 1} \times (\ln R_{12} + \ln \check{R}_{21})$$

Jensen-Shannon divergence: Just as with chi-squared divergence, KL divergence can be symmetrized by looking at the divergence of  $P_1$  and  $P_2$  from  $P_*= (P_1+P_2)/2$  instead. The Jensen-Shannon divergence (4, 5) is obtained:

$$S_1(P_1, P_2) = \frac{1}{2} [KL(P_1, P_*) + KL(P_2, P_*)] = \left[ \frac{p_1 \ln p_1 + p_2 \ln p_2}{2} - p_* \ln p_* \right] + \left[ \frac{q_1 \ln q_1 + q_2 \ln q_2}{2} - q_* \ln q_* \right]$$

In this paper, we use a rescaled version  $S_2$  of Jensen-Shannon divergence, with maximum value equal 1:

$$S_2(P_1, P_2) = \frac{1}{\ln 2} \times S_1(P_1, P_2)$$

Jensen-Shannon divergence also can be written as a function of the two ratios  $R_{12}=p_1/p_2$  and  $\check{R}_{21}=q_2/q_1$ . To see this, define the function  $s(r) = \frac{1}{2 \ln 2} \{r \ln r - (r + 1) \ln[(r + 1)/2]\}$ . Then  $S_2 = p_2 s(R_{12}) + q_2 s(1/\check{R}_{21})$ , where, as previously,  $p_2 = (\check{R}_{21} - 1)/(R_{12}\check{R}_{21} - 1)$  and  $q_2 = 1 - p_2 = (R_{12} - 1)\check{R}_{21}/(R_{12}\check{R}_{21} - 1)$ .

*Hellinger distance.* Hellinger distance between  $P_1$  and  $P_2$  is a doubly symmetric measure of inequality (4, 5):

$$H(P_1, P_2) = \sqrt{\left[ (\sqrt{p_1} - \sqrt{p_2})^2 + (\sqrt{q_1} - \sqrt{q_2})^2 \right] / 2} = \sqrt{1 - \sqrt{p_1 p_2} - \sqrt{q_1 q_2}}$$

Written as a function of the two ratios  $R_{12}=p_1/p_2$  and  $\check{R}_{21}=q_2/q_1$ , Hellinger distance is given by

$$H(R_{12}, \check{R}_{21}) = \sqrt{1 - \frac{(R_{12} - 1)\sqrt{1/\check{R}_{21}}}{R_{12}\check{R}_{21} - 1} - \frac{(\check{R}_{21} - 1)\sqrt{R_{12}}}{R_{12}\check{R}_{21} - 1}}$$

Behavior of selected inequality measures for  $|p_1 - p_2|$  near 0

Let  $p_1 = p_2 + \delta$  for  $\delta > 0$ , and let  $q_2 = 1 - p_2$ . We investigate the rate of change in  $d(p_1, p_2)$  as  $\delta \rightarrow 0^+$ , i.e., the ratio:

$$\frac{d(p_2 + \delta, p_2) - d(p_2, p_2)}{\delta - 0} = \frac{d(p_2 + \delta, p_2)}{\delta}$$

*Absolute difference.* The absolute difference  $a(p_2 + \delta, p_2) = \delta$ , independently of the value of  $p_2$ . Indeed,

$$a(p_2 + \delta, p_2) = |p_2 + \delta - p_2| = \delta.$$

*Standardized absolute difference.* As  $\delta \rightarrow 0^+$ , the magnitude of  $D_1$  will depend on  $p_2$  as follows:

$$D_1(p_2 + \delta, p_2) \approx \delta \times \frac{1}{\sqrt{p_2 q_2}}$$

To show this, note that as  $\delta \rightarrow 0^+$ , the ratio  $D_1(p_2 + \delta, p_2)/\delta$  approaches the one-sided partial derivative of

$D_1(p_1, p_2)$  with respect to its first argument, which is  $\frac{1}{\sqrt{p_1 q_1}}$ .

*Standardized absolute difference with pooled variance.* As  $\delta \rightarrow 0^+$ ,  $D_2$  will depend on  $p_2$  as follows:

$$D_2(p_2 + \delta, p_2) \approx \delta \times \frac{1}{2 \times \sqrt{p_2 q_2}}$$

To show this, note that as  $\delta \rightarrow 0^+$ , the ratio  $D_2(p_2 + \delta, p_2)/\delta$  approaches the one-sided partial derivative of

$D_2(p_1, p_2)$  with respect to its first argument, which is  $\frac{1}{2\sqrt{p_1 q_1}}$ .

*Rescaled absolute arcsine difference.* As  $\delta \rightarrow 0^+$ ,  $h$  will depend on  $p_2$  as follows:

$$h(p_2 + \delta, p_2) \approx \delta \times \frac{1}{\pi \times \sqrt{p_2 q_2}}$$

To see this, note that the derivative of  $\arcsin(\sqrt{p_1})$  with respect to  $p_1$  is  $\frac{1}{2\sqrt{p_1(1-p_1)}}$ . Thus, as  $\delta \rightarrow 0^+$ , the ratio

$h(p_2 + \delta, p_2)/\delta$  approaches the one-sided partial derivative of  $h(p_1, p_2)$  with respect to its first argument, which is given by  $\frac{1}{\pi\sqrt{p_1q_1}}$ .

*Absolute logit difference.* As  $\delta \rightarrow 0^+$ ,  $\ell$  will depend on  $p_2$  as follows:

$$\ell(p_2 + \delta, p_2) \approx \delta \times \frac{1}{p_2q_2}$$

Here, the derivative of  $\text{logit}(p_1)$  with respect to  $p_1$  is  $\frac{1}{p_1(1-p_1)}$ . Thus, as  $\delta \rightarrow 0^+$ , the ratio  $\ell(p_2 + \delta, p_2)/\delta$

approaches the one-sided partial derivative of  $\ell(p_1, p_2)$  with respect to its first argument, which is given by  $\frac{1}{p_1q_1}$ .

*Absolute probit difference.* As  $\delta \rightarrow 0^+$ ,  $b$  will depend on  $p_2$  as follows:

$$b(p_2 + \delta, p_2) \approx \delta \times \sqrt{2\pi} \times \exp\{ [\Phi^{-1}(p_2)]^2/2 \}$$

Here, the derivative of  $\text{probit}(p_1)$  with respect to  $p_1$  is  $\frac{1}{\Phi'(\Phi^{-1}(p_1))}$ , where  $\Phi(z)$  is the cumulative distribution

function for the standard normal distribution,  $\Phi'(z)$  is its derivative, and  $\Phi^{-1}$  its inverse function. Thus, as  $\delta \rightarrow 0^+$ ,

the ratio  $b(p_2 + \delta, p_2)/\delta$  approaches the one-sided partial derivative of  $b(p_1, p_2)$  with respect to its first argument, which is  $\sqrt{2\pi} \times \exp\{ [\Phi^{-1}(p_2)]^2/2 \}$ .

*Hellinger distance.* With a slight abuse of notation, write  $H(p_1, p_2)$  for  $H(P_1, P_2)$ . As  $\delta \rightarrow 0^+$ ,  $H$  will depend on  $p_2$ :

$$H(p_2 + \delta, p_2) \approx \delta \times \frac{1}{(2\sqrt{2}) \times \sqrt{p_2q_2}}$$

To show this, note that as  $\delta \rightarrow 0^+$ , the ratio  $H(p_2 + \delta, p_2)/\delta$  approaches the one-sided partial derivative of

$H(p_1, p_2)$  with respect to its first argument, which is seen to be  $\frac{1}{2\sqrt{2p_1q_1}}$ .



For the remaining information-theoretic divergence measures, just as with Hellinger distance, we write  $d(p_1, p_2)$  for  $d(P_1, P_2)$ . Also, due to the fact that the derivatives exist and equal zero at  $p_1 = p_2$  for these measures, we investigate the rate of change of  $d(p_2 + \delta, p_2)$  as  $\delta \rightarrow 0$ , relative to  $\delta^2$  instead of just  $\delta$ :

$$\frac{d(p_2 + \delta, p_2)}{\delta^2}$$

Thus, we look at the second-order Taylor linearization of  $d(p_2 + \delta, p_2)$  for  $\delta > 0$  near zero.

*Symmetrized chi-squared measure.* As  $\delta \rightarrow 0$ , the magnitude of  $X$  will depend on  $p_2$  as follows:

$$X(p_2 + \delta, p_2) \approx \delta^2 \times \frac{1}{p_2 q_2}$$

Here, as  $\delta$  approaches zero, the ratio  $X(p_2 + \delta, p_2)/\delta^2$  approaches one-half times the second partial derivative of  $X(p_1, p_2)$  with respect to its first argument, i.e.,  $\frac{1}{2} \frac{\partial^2}{\partial p_1^2} = \frac{1}{p_1 q_1}$ .

*Triangle discrimination measure.* As  $\delta \rightarrow 0$ ,  $\Delta$  will depend on  $p_2$  as follows:

$$\Delta(p_2 + \delta, p_2) \approx \delta^2 \times \frac{1}{4 \times p_2 q_2}$$

Here, as  $\delta$  approaches zero, the ratio  $\Delta(p_2 + \delta, p_2)/\delta^2$  approaches one-half times the second partial derivative of  $\Delta(p_1, p_2)$  with respect to its first argument, i.e.,  $\frac{1}{4} \frac{\partial^2}{\partial p_1^2}$ .

*Jeffreys divergence.* As  $\delta \rightarrow 0$ ,  $J$  will depend on  $p_2$  as follows:

$$J(p_2 + \delta, p_2) \approx \delta^2 \times \frac{1}{2 \times p_2 q_2}$$

Here, as  $\delta$  approaches zero, the ratio  $J(p_2 + \delta, p_2)/\delta^2$  approaches one-half times the second partial derivative of  $J(p_1, p_2)$  with respect to its first argument, i.e.,  $\frac{1}{2} \frac{\partial^2}{\partial p_1^2}$ .

*Jensen-Shannon divergence.* As  $\delta \rightarrow 0$ ,  $S_1$  will depend on  $p_2$  as follows:

$$S_1(p_2 + \delta, p_2) \approx \delta^2 \times \frac{1}{8 \times p_2 q_2}$$

Here, as  $\delta$  approaches zero, the ratio  $S_1(p_2 + \delta, p_2)/\delta^2$  approaches one-half times the second partial derivative of  $S_1(p_1, p_2)$  with respect to its first argument, i.e.,  $\frac{1}{8 p_1 q_1}$ .

Local relationships among selected effect size-based inequality measures for  $p_1 \approx p_2$

The effect size-based measures  $D_1$ ,  $D_2$ , and  $h$  in Table 1, as well as Hellinger distance  $H$ , are nearly proportional to  $\pm \delta \times \frac{1}{\sqrt{p_2 q_2}}$  for  $|p_1 - p_2| = \delta$  near zero. Additionally,

- If  $p_2$  falls within the interval  $\frac{1}{2} \pm \frac{1}{2\sqrt{2}}$ , i.e., approximately  $0.1464 \leq p_2 \leq 0.8536$ , then

$$h(p_1, p_2) \leq H(p_1, p_2) \leq a(p_1, p_2)$$

- If  $p_2$  falls between  $\frac{1}{2} - \sqrt{\frac{1}{4} - \frac{1}{\pi^2}}$  and  $\frac{1}{2} - \frac{1}{2\sqrt{2}}$ , i.e., approximately  $0.1144 \leq p_2 \leq 0.1464$ , or if  $p_2$  falls between  $\frac{1}{2} + \frac{1}{2\sqrt{2}}$  and  $\frac{1}{2} + \sqrt{\frac{1}{4} - \frac{1}{\pi^2}}$ , i.e., approximately  $0.8536 \leq p_2 \leq 0.8856$ , then

$$h(p_1, p_2) \leq a(p_1, p_2) \leq H(p_1, p_2)$$

- If  $p_2$  is outside of the interval  $\frac{1}{2} \pm \sqrt{\frac{1}{4} - \frac{1}{\pi^2}}$ , i.e., approximately  $p_2 \leq 0.1144$  or  $p_2 \geq 0.8856$ , then

$$a(p_1, p_2) \leq h(p_1, p_2) \leq H(p_1, p_2)$$

Local relationships among selected information theory-based inequality measures for  $p_1 \approx p_2$

The information theory-based measures listed in Table 1, including squared Hellinger distance, are nearly proportional to  $\delta^2 \times \frac{1}{p_2 q_2}$  for  $|p_1 - p_2| = \delta$  near zero. Additionally:

- If  $p_2$  falls within the interval  $\frac{1}{2} \pm \sqrt{\frac{1}{4} - \frac{1}{8 \ln 2}}$ , i.e., approximately  $0.2361 \leq p_2 \leq 0.7639$ , then

$$h(p_1, p_2)^2 \leq H(p_1, p_2)^2 \leq S_2(p_1, p_2) \leq a(p_1, p_2)^2$$

- If  $p_2$  falls between  $\frac{1}{2} - \frac{1}{2\sqrt{2}}$  and  $\frac{1}{2} - \sqrt{\frac{1}{4} - \frac{1}{8 \ln 2}}$ , i.e., approximately  $0.1464 \leq p_2 \leq 0.2361$ , or if  $p_2$  falls between  $\frac{1}{2} + \sqrt{\frac{1}{4} - \frac{1}{8 \ln 2}}$  and  $\frac{1}{2} + \frac{1}{2\sqrt{2}}$ , i.e., approximately  $0.7639 \leq p_2 \leq 0.8536$ , then

$$h(p_1, p_2)^2 \leq H(p_1, p_2)^2 \leq a(p_1, p_2)^2 \leq S_2(p_1, p_2)$$

- If  $p_2$  falls between  $\frac{1}{2} - \sqrt{\frac{1}{4} - \frac{1}{\pi^2}}$  and  $\frac{1}{2} - \frac{1}{2\sqrt{2}}$ , i.e., approximately  $0.1144 \leq p_2 \leq 0.1464$ , or if  $p_2$  falls between  $\frac{1}{2} + \frac{1}{2\sqrt{2}}$  and  $\frac{1}{2} + \sqrt{\frac{1}{4} - \frac{1}{\pi^2}}$ , i.e., approximately  $0.8536 \leq p_2 \leq 0.8856$ , then

$$h(p_1, p_2)^2 \leq a(p_1, p_2)^2 \leq H(p_1, p_2)^2 \leq S_2(p_1, p_2)$$

- If  $p_2$  is outside of the interval  $\frac{1}{2} \pm \sqrt{\frac{1}{4} - \frac{1}{\pi^2}}$ , i.e., approximately  $p_2 \leq 0.1144$  or  $p_2 \geq 0.8856$ ,

$$a(p_1, p_2)^2 \leq h(p_1, p_2)^2 \leq H(p_1, p_2)^2 \leq S_2(p_1, p_2)$$

Global relationships among selected measures

The six measures  $h$ ,  $H$ ,  $D_2$ ,  $a$ ,  $\Delta$  and  $S_2$  satisfy:

$$h \leq H \leq D_2$$

$$a \leq D_2 = \sqrt{\Delta} \quad [3]$$

$$\frac{1}{2} \Delta \leq (\ln 2) S_2 \leq H^2 \leq S_2 \leq \Delta \quad [4]$$

To show that  $a \leq D_2$ , as stated in expression 3, note that  $2\sqrt{p_*q_*} \leq 1$  for  $q_* = 1 - p_*$ . Thus

$$D_2 = \frac{|p_1 - p_2|}{2\sqrt{p_*q_*}} \geq |p_1 - p_2| = a$$

That  $a^2 \leq \Delta$  follows from  $a \leq D_2$  and  $D_2 = \sqrt{\Delta}$ .

To show that  $H \leq D_2$  (expression 2), it is sufficient to show  $H^2 \leq \Delta$ , since  $D_2 = \sqrt{\Delta}$ . Now,

$$\begin{aligned} 2\Delta &= \frac{(p_1 - p_2)^2}{p_1 + p_2} + \frac{(q_1 - q_2)^2}{q_1 + q_2} \\ &= (\sqrt{p_1} - \sqrt{p_2})^2 \times \frac{(\sqrt{p_1} + \sqrt{p_2})^2}{p_1 + p_2} + (\sqrt{q_1} - \sqrt{q_2})^2 \times \frac{(\sqrt{q_1} + \sqrt{q_2})^2}{q_1 + q_2} \\ &= (\sqrt{p_1} - \sqrt{p_2})^2 \times \left[ 1 + \frac{2\sqrt{p_1p_2}}{p_1 + p_2} \right] + (\sqrt{q_1} - \sqrt{q_2})^2 \times \left[ 1 + \frac{2\sqrt{q_1q_2}}{q_1 + q_2} \right] \end{aligned}$$

The terms in square brackets in this last equality remain  $\geq 1$ , thus

$$2\Delta \geq (\sqrt{p_1} - \sqrt{p_2})^2 + (\sqrt{q_1} - \sqrt{q_2})^2 = 2H^2$$

or  $\Delta \geq H^2$ , as claimed. In fact, the arithmetic-geometric mean inequality ensures  $\sqrt{xy} \leq (x + y)/2$  for any two positive scalars  $x$  and  $y$ . Thus, the terms in square brackets in the above equality are  $\leq 2$ , resulting in the inequality

$$2\Delta \leq 2(\sqrt{p_1} - \sqrt{p_2})^2 + 2(\sqrt{q_1} - \sqrt{q_2})^2 = 4H^2$$

i.e.,  $\frac{1}{2}\Delta \leq H^2$  (expression 4). To show that  $h \leq H$  (expression 2), note that, by definition of the arcsine

transformation and the fact that  $\sin^2(x) + \cos^2(x) = 1$ , we have  $p_1 = \sin^2[\arcsin(\sqrt{p_1})]$  and  $q_1 = 1 - p_1 =$

$\cos^2[\arcsin(\sqrt{p_1})]$ . Thus, we can write  $H^2$  as follows:

$$\begin{aligned}
H(p_1, p_2)^2 &= 1 - \sqrt{p_1 p_2} - \sqrt{q_1 q_2} \\
&= 1 - \sin[\arcsin(\sqrt{p_1})] \sin[\arcsin(\sqrt{p_2})] - \cos[\arcsin(\sqrt{p_1})] \cos[\arcsin(\sqrt{p_2})] \\
&= 1 - \cos[\arcsin(\sqrt{p_1}) - \arcsin(\sqrt{p_2})] \\
&= 1 - \cos\left[\frac{\pi}{2} \times h\right] \\
&= 2 \sin^2\left[\frac{\pi}{4} \times h\right]
\end{aligned}$$

The third equality above is a result of the trigonometric formula  $\cos(x - y) = \cos x \cos y + \sin x \sin y$ , the fourth equality is due to the symmetry of the cosine,  $\cos(-x) = \cos(x)$ , and the last equality is a result of the half angle formula  $\cos(x) = 1 - 2 \sin^2(x/2)$ . Thus,

$$H - h = \sqrt{2} \sin\left[\frac{\pi}{4} \times h\right] - h$$

It is straightforward to verify that the right-hand side remains  $\geq 0$  for  $0 \leq h \leq 1$ , i.e.,  $H - h \geq 0$ , as claimed.

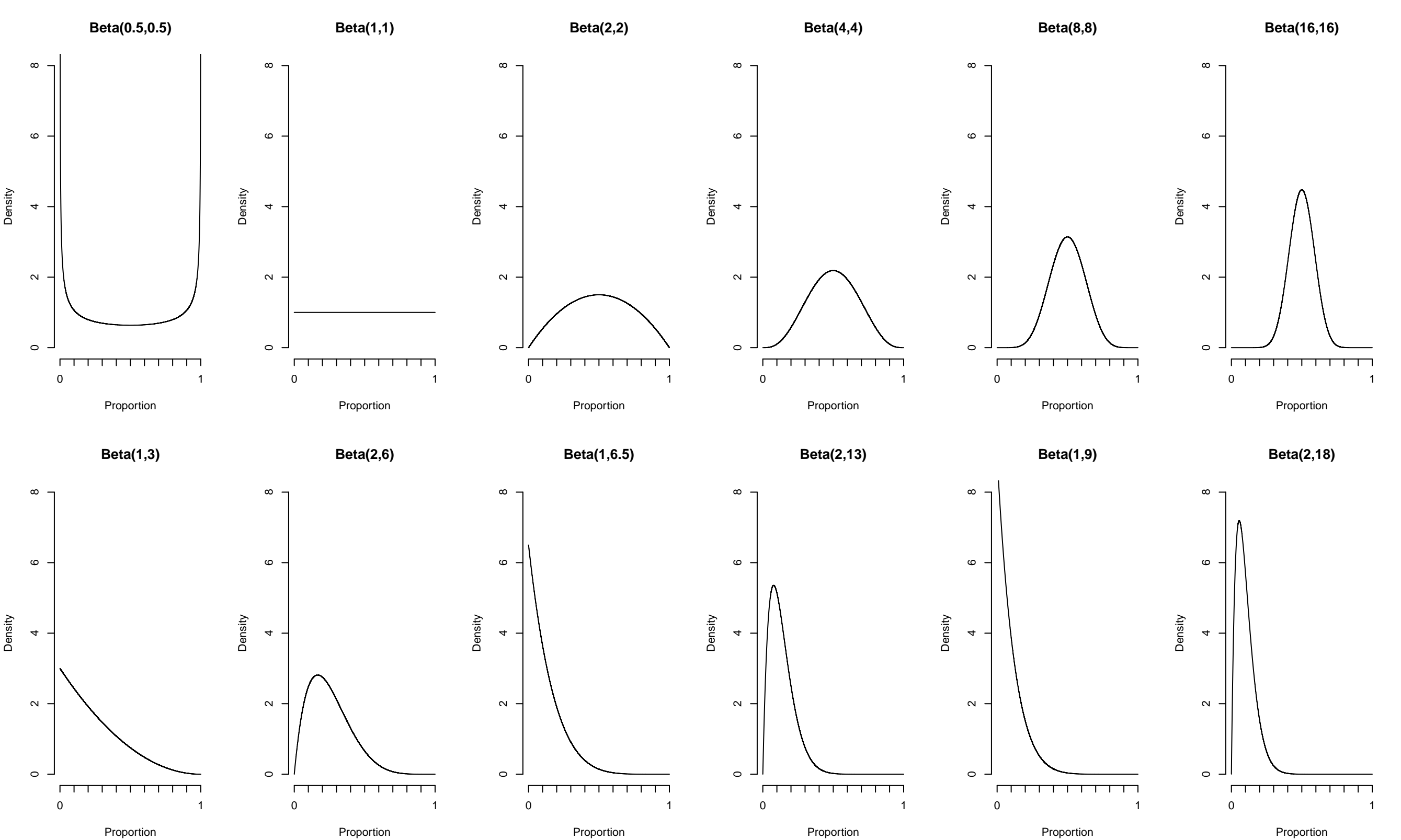
The full string of inequalities  $\frac{1}{2}\Delta \leq S_1 \leq H^2 \leq S_2 \leq \Delta$  (expression 4) can be shown using ratio bounds between different Csiszár's  $f$ -divergence measures; see (4, 5).

## References

1. Kelley K, Preacher KJ. On effect size. *Psychol Methods*. 2012;17(2):137–152.
2. Sánchez-Meca J, Marín-Martínez F, Chacón-Moscoso S. Effect-size indices for dichotomized outcomes in meta-analysis. *Psychol Methods*. 2003;8(4):448–467.
3. Cressie N, Read TRC. Multinomial goodness-of-fit tests. *J R Statist Soc B*. 1984;46(3):440–464.
4. Mitchell SA. Geometric comparison of popular mixture model distances. *Journal of Modern Mathematics Frontier*. 2012;1(4):1–21.
5. Taneja IJ. Refinement inequalities among symmetric divergence measures. *Australian Journal of Mathematical Analysis and Applications*. 2005;2(1):1–23.









6. Harper S, Lynch J, Meersman SC, et al. An overview of methods for monitoring social disparities in cancer with an example using trends in lung cancer incidence by area-socioeconomic position and race/ethnicity, 1992–2004. *Am J Epidemiol.* 2008;167(8):889–899.
7. Talih M. A reference-invariant health disparity index based on Rényi divergence. *Ann Appl Stat.* 2013;7(2):1217–1243.
8. Borrell LN, Talih M. A symmetrized Theil index measure of health disparities: An example using dental caries in U.S. children and adolescents. *Stat Med.* 2011;30(3):277–290.

# Web Figure 1. Probability Density Functions for Selected Beta( $\alpha,\beta$ ) Distributions





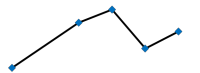
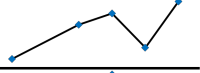

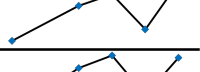

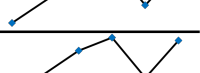
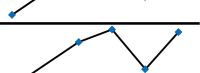
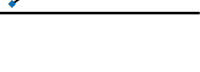
**Web Table 1. Proportion\* of Children Receiving 3+ Doses of Polio Vaccine by Age 19–35 Months, by Family Income: 2012–2017**

	2012	2013	2014	2015	2016	2017	
Percent poverty threshold							
<100	0.920	0.894	0.922	0.919	0.911	0.917	
600+	0.945	0.954	0.939	0.967	0.952	0.959	
<u>Selected inequality measures</u>							<u>Trend in inequality</u>
Ratio $p_1/p_2$ with $p_1 > p_2$	1.027	1.067	1.018	1.052	1.045	1.046	
Ratio $(1-p_2)/(1-p_1)$ with $p_1 > p_2$	1.455	2.304	1.279	2.455	1.854	2.024	
Absolute logit difference (or absolute log-odds ratio)	0.402	0.900	0.264	0.949	0.661	0.750	
Absolute difference	0.025	0.060	0.017	0.048	0.041	0.042	
Square root of triangle discrimination measure	0.050	0.113	0.033	0.104	0.081	0.087	
Rescaled absolute arcsine difference	0.032	0.074	0.021	0.067	0.052	0.056	
Hellinger distance	0.035	0.082	0.024	0.075	0.058	0.062	
Square root of rescaled Jensen-Shanon divergence	0.042	0.097	0.028	0.089	0.069	0.075	

\*Based on HP2020 objective IID-7.5

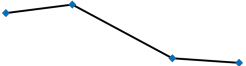
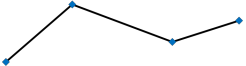
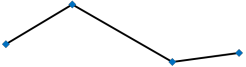
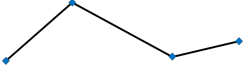
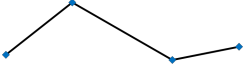
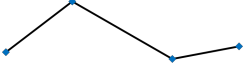
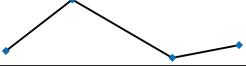
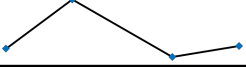
Data Source: National Immunization Survey (NIS), CDC/NCIRD and CDC/NCHS

Web Table 2. Proportion\* of Children Receiving 0 Doses of Recommended Vaccines by Age 19–35 Months, by Geographic Location: 2012–2017

	2012	2014	2015	2016	2017	
Geographic location						
Metropolitan	0.008	0.007	0.007	0.007	0.010	
Non-metropolitan	0.012	0.014	0.015	0.012	0.019	
Selected inequality measures						Trend in inequality
Ratio $p_1/p_2$ with $p_1 > p_2$	1.500	2.000	2.143	1.714	1.900	
Ratio $(1-p_2)/(1-p_1)$ with $p_1 > p_2$	1.004	1.007	1.008	1.005	1.009	
Absolute logit difference (or absolute log-odds ratio)	0.410	0.700	0.770	0.544	0.651	
Absolute difference	0.004	0.007	0.008	0.005	0.009	
Square root of triangle discrimination measure	0.020	0.034	0.038	0.026	0.038	
Rescaled absolute arcsine difference	0.013	0.022	0.025	0.017	0.024	
Hellinger distance	0.014	0.025	0.028	0.018	0.027	
Square root of rescaled Jensen-Shanon divergence	0.017	0.029	0.033	0.022	0.032	

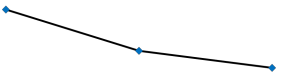
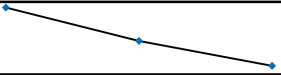
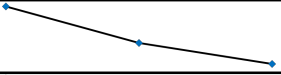
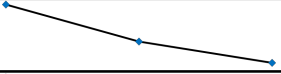
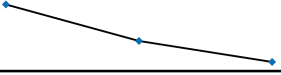
\*Based on HP2020 objective IID-9  
Data Source: National Immunization Survey (NIS), CDC/NCIRD and CDC/NCHS

**Web Table 3. Age-adjusted Proportion\* of Adults Aged 50–64 Receiving Colorectal Cancer Screening Based on the Most Recent Guidelines, by Health Insurance Status: 2008–2015**

	2008	2010	2013	2015	
Health insurance status					
Insured	0.525	0.604	0.578	0.610	
Uninsured	0.187	0.210	0.236	0.253	
Selected inequality measures					Trend in inequality
Ratio $p_1/p_2$ with $p_1 > p_2$	2.807	2.876	2.449	2.411	
Ratio $(1-p_2)/(1-p_1)$ with $p_1 > p_2$	1.712	1.995	1.810	1.915	
Absolute logit difference (or absolute log-odds ratio)	1.570	1.747	1.489	1.530	
Absolute difference	0.338	0.394	0.342	0.357	
Square root of triangle discrimination measure	0.353	0.401	0.348	0.360	
Rescaled absolute arcsine difference	0.231	0.264	0.227	0.235	
Hellinger distance	0.255	0.291	0.251	0.260	
Square root of rescaled Jensen-Shanon divergence	0.526	0.563	0.520	0.528	

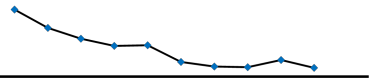
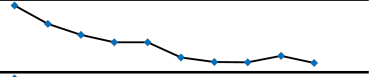
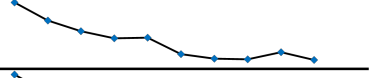
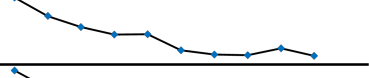
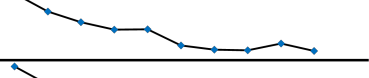
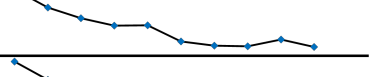
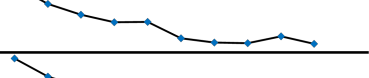
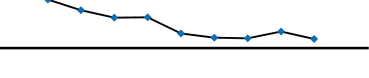
\*Based on HP2020 objective C-16  
Data Source: National Health Interview Survey (NHIS), CDC/NCHS

Web Table 4. Proportion\* of Children and Adolescents Aged 2–19 Years with Obesity, by Family Income: 2005–2008 to 2013–2016

	2005–2008	2009–2012	2013–2016	
Percent poverty threshold				
<100	0.199	0.209	0.210	
500+	0.098	0.116	0.123	
Selected inequality measures				Trend in inequality
Ratio $p_1/p_2$ with $p_1 > p_2$	2.031	1.802	1.707	
Ratio $(1-p_2)/(1-p_1)$ with $p_1 > p_2$	1.126	1.118	1.110	
Absolute logit difference (or absolute log-odds ratio)	0.827	0.700	0.639	
Absolute difference	0.101	0.093	0.087	
Square root of triangle discrimination measure	0.142	0.126	0.117	
Rescaled absolute arcsine difference	0.092	0.081	0.075	
Hellinger distance	0.102	0.090	0.083	
Square root of rescaled Jensen-Shanon divergence	0.122	0.108	0.100	

\*Based on HP2020 objective NWS-10.4  
Data Source: National Health and Nutrition Examination Survey (NHANES), CDC/NCHS

**Web Table 5. Proportion of Deaths Due to Heart Disease, for the Non-Hispanic White and Black Populations: 2007–2016**

	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	
Population subgroup											
Non-Hispanic white	0.259	0.254	0.250	0.246	0.241	0.239	0.238	0.237	0.237	0.235	
Non-Hispanic black	0.246	0.245	0.243	0.241	0.236	0.237	0.238	0.237	0.235	0.234	
<u>Selected inequality measures</u>											<u>Trend in inequality</u>
Ratio $p_1/p_2$ with $p_1 > p_2$	1.054	1.037	1.027	1.021	1.021	1.006	1.002	1.002	1.008	1.001	
Ratio $(1-p_2)/(1-p_1)$ with $p_1 > p_2$	1.018	1.012	1.009	1.007	1.007	1.002	1.001	1.001	1.003	1.000	
Absolute logit difference (or absolute log-odds ratio)	0.070	0.048	0.036	0.027	0.028	0.008	0.003	0.002	0.011	0.001	
Absolute difference	0.013	0.009	0.007	0.005	0.005	0.002	0.001	0.000	0.002	0.000	
Square root of triangle discrimination measure	0.015	0.010	0.008	0.006	0.006	0.002	0.001	0.000	0.002	0.000	
Rescaled absolute arcsine difference	0.010	0.007	0.005	0.004	0.004	0.001	0.000	0.000	0.001	0.000	
Hellinger distance	0.011	0.007	0.005	0.004	0.004	0.001	0.000	0.000	0.002	0.000	
Square root of rescaled Jensen-Shanon divergence	0.013	0.009	0.007	0.005	0.005	0.001	0.001	0.000	0.002	0.000	

Data Sources: National Vital Statistics System–Mortality (NVSS-M), CDC/NCHS; Bridged-race Population Estimates, CDC/NCHS and Census

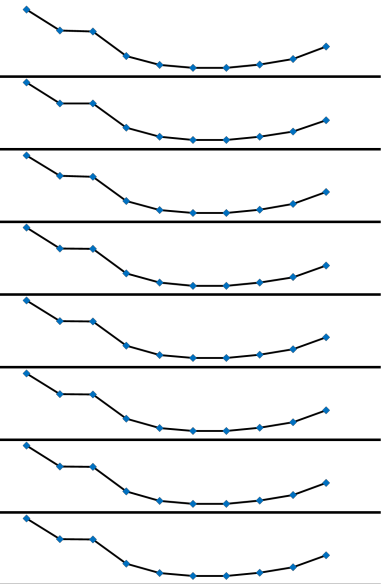
**Web Table 6. Proportion of Deaths Due to Malignant Neoplasms, for the Non-Hispanic White and Black Populations: 2007–2016**

	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Population subgroup										
Non-Hispanic white	0.235	0.231	0.235	0.234	0.229	0.229	0.225	0.226	0.220	0.219
Non-Hispanic black	0.222	0.222	0.226	0.230	0.231	0.229	0.225	0.224	0.217	0.213

Selected inequality measures

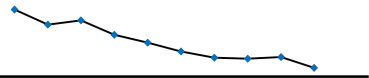
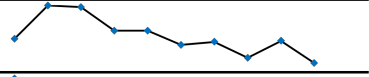
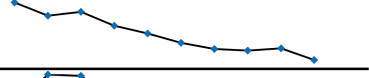
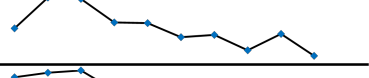
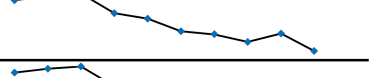
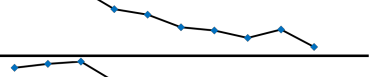
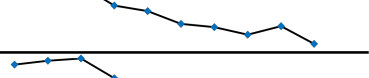
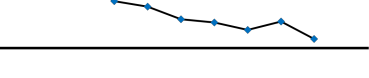
Ratio $p_1/p_2$ with $p_1 > p_2$	1.063	1.041	1.040	1.015	1.006	1.003	1.003	1.006	1.012	1.025
Ratio $(1-p_2)/(1-p_1)$ with $p_1 > p_2$	1.018	1.012	1.012	1.005	1.002	1.001	1.001	1.002	1.003	1.007
Absolute logit difference (or absolute log-odds ratio)	0.079	0.052	0.051	0.020	0.008	0.004	0.004	0.008	0.015	0.031
Absolute difference	0.014	0.009	0.009	0.003	0.001	0.001	0.001	0.001	0.003	0.005
Square root of triangle discrimination measure	0.017	0.011	0.011	0.004	0.002	0.001	0.001	0.002	0.003	0.006
Rescaled absolute arcsine difference	0.011	0.007	0.007	0.003	0.001	0.001	0.001	0.001	0.002	0.004
Hellinger distance	0.012	0.008	0.008	0.003	0.001	0.001	0.001	0.001	0.002	0.005
Square root of rescaled Jensen-Shanon divergence	0.014	0.009	0.009	0.004	0.001	0.001	0.001	0.001	0.003	0.005

Trend in inequality



Data Sources: National Vital Statistics System–Mortality (NVSS-M), CDC/NCHS; Bridged-race Population Estimates, CDC/NCHS and Census

**Web Table 7. Proportion of Deaths Due to Chronic Lower Respiratory Disease, for the Non-Hispanic White and Black Populations: 2007–2016**

	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	
Population subgroup											
Non-Hispanic white	0.059	0.064	0.063	0.062	0.063	0.063	0.064	0.063	0.064	0.063	
Non-Hispanic black	0.027	0.030	0.030	0.030	0.031	0.032	0.033	0.032	0.033	0.033	
Selected inequality measures											Trend in inequality
Ratio $p_1/p_2$ with $p_1 > p_2$	2.159	2.095	2.113	2.053	2.019	1.981	1.955	1.950	1.958	1.912	
Ratio $(1-p_2)/(1-p_1)$ with $p_1 > p_2$	1.034	1.036	1.035	1.034	1.034	1.033	1.033	1.033	1.034	1.032	
Absolute logit difference (or absolute log-odds ratio)	0.803	0.774	0.783	0.753	0.736	0.716	0.704	0.700	0.705	0.680	
Absolute difference	0.032	0.033	0.033	0.032	0.032	0.031	0.031	0.031	0.031	0.030	
Square root of triangle discrimination measure	0.078	0.079	0.079	0.076	0.075	0.073	0.073	0.072	0.073	0.071	
Rescaled absolute arcsine difference	0.050	0.051	0.051	0.049	0.049	0.047	0.047	0.046	0.047	0.045	
Hellinger distance	0.056	0.056	0.057	0.055	0.054	0.053	0.052	0.051	0.052	0.050	
Square root of rescaled Jensen-Shanon divergence	0.067	0.067	0.068	0.065	0.065	0.063	0.063	0.062	0.063	0.060	

Data Sources: National Vital Statistics System–Mortality (NVSS-M), CDC/NCHS; Bridged-race Population Estimates, CDC/NCHS and Census

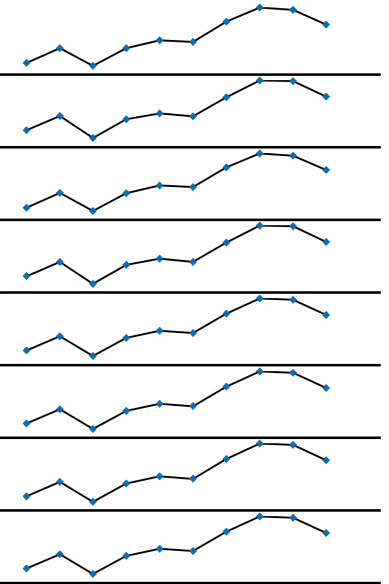
**Web Table 8. Proportion of Deaths Due to Cerebrovascular Disease, for the Non-Hispanic White and Black Populations: 2007–2016**

	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Population subgroup										
Non-Hispanic white	0.056	0.054	0.052	0.052	0.050	0.050	0.049	0.050	0.050	0.050
Non-Hispanic black	0.059	0.058	0.056	0.056	0.055	0.054	0.054	0.056	0.056	0.056

Selected inequality measures

Ratio $p_1/p_2$ with $p_1 > p_2$	1.065	1.080	1.062	1.080	1.088	1.087	1.108	1.122	1.120	1.105
Ratio $(1-p_2)/(1-p_1)$ with $p_1 > p_2$	1.004	1.005	1.003	1.004	1.005	1.005	1.006	1.006	1.006	1.006
Absolute logit difference (or absolute log-odds ratio)	0.067	0.082	0.064	0.082	0.090	0.088	0.108	0.122	0.120	0.105
Absolute difference	0.004	0.004	0.003	0.004	0.004	0.004	0.005	0.006	0.006	0.005
Square root of triangle discrimination measure	0.008	0.009	0.007	0.009	0.010	0.010	0.012	0.014	0.013	0.012
Rescaled absolute arcsine difference	0.005	0.006	0.005	0.006	0.006	0.006	0.008	0.009	0.009	0.007
Hellinger distance	0.006	0.007	0.005	0.007	0.007	0.007	0.008	0.010	0.009	0.008
Square root of rescaled Jensen-Shanon divergence	0.007	0.008	0.006	0.008	0.008	0.008	0.010	0.012	0.011	0.010

Trend in inequality



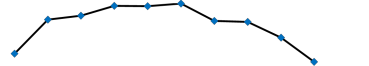
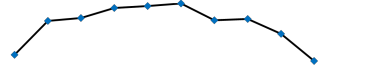
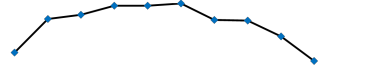
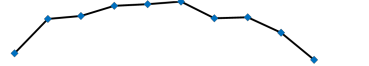
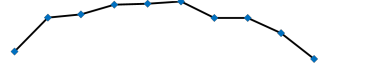
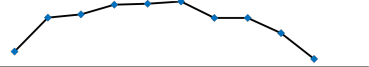

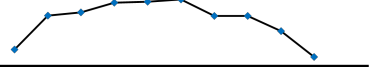
Data Sources: National Vital Statistics System–Mortality (NVSS-M), CDC/NCHS; Bridged-race Population Estimates, CDC/NCHS and Census



**Web Table 9. Proportion of Deaths Due to Unintentional Injury, for the Non-Hispanic White and Black Populations: 2007–2016**

	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Population subgroup										
Non-Hispanic white	0.049	0.048	0.047	0.048	0.049	0.049	0.049	0.051	0.053	0.057
Non-Hispanic black	0.047	0.043	0.042	0.042	0.043	0.043	0.044	0.045	0.049	0.055

Selected inequality measures

	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	Trend in inequality
Ratio $p_1/p_2$ with $p_1 > p_2$	1.047	1.118	1.126	1.146	1.146	1.151	1.116	1.113	1.080	1.030	
Ratio $(1-p_2)/(1-p_1)$ with $p_1 > p_2$	1.002	1.005	1.006	1.006	1.007	1.007	1.005	1.005	1.004	1.002	
Absolute logit difference (or absolute log-odds ratio)	0.049	0.116	0.125	0.143	0.143	0.148	0.115	0.113	0.081	0.031	
Absolute difference	0.002	0.005	0.005	0.006	0.006	0.006	0.005	0.005	0.004	0.002	
Square root of triangle discrimination measure	0.005	0.012	0.013	0.015	0.015	0.015	0.012	0.012	0.009	0.004	
Rescaled absolute arcsine difference	0.003	0.008	0.008	0.009	0.010	0.010	0.008	0.008	0.006	0.002	
Hellinger distance	0.004	0.009	0.009	0.010	0.011	0.011	0.009	0.009	0.006	0.003	
Square root of rescaled Jensen-Shanon divergence	0.004	0.010	0.011	0.013	0.013	0.013	0.010	0.010	0.008	0.003	

Data Sources: National Vital Statistics System–Mortality (NVSS-M), CDC/NCHS; Bridged-race Population Estimates, CDC/NCHS and Census