Short Communication

# Variation Index: A New Alternative for Measuring Income Inequality

**Hening Huang**[1]

1. Independent researcher

**This technical note proposes a new index for measuring income inequality (or the inequality of a distribution), named as the "variation index (VI)". The proposed VI is dimensionless and bounded between 0 and 1, with 0 indicating perfect equality and 1 indicating extreme inequality. Several examples are provided to demonstrate the effectiveness of the proposed VI and to compare it with existing inequality indices, including the coefficient of variation (CV) and the Gini coefficient.**

## 1. Introduction

Inequality refers to the uneven distribution of resources or outcomes within a population. It is defined relative to a normative ideal of perfect equality. For example, income inequality means that individuals' incomes differ from one another compared to the ideal case where everyone has exactly the same income (a degenerate distribution with zero variance). Measures of inequality are used to assess how much the observed distribution deviates from that ideal of perfect equality.

The Gini coefficient is the most commonly used measure of income inequality. De Maio[1] reviewed the Gini coefficient as well as several alternative methods, including the coefficient of variation (CV), the generalized entropy (GE) index, the Kakwani progressivity index, and the Robin Hood index. More recently, Kim et al.[2] pointed out that conventional income inequality indices may misassess the degree of inequality due to three problems. They proposed a new inequality index called the $L_2$ index, based on a relative unequally distributed (RUD) income-based framework developed by Park et al.[3]. Sitthiyot and Holasut[4] proposed a composite index that comprises three indicators: the Gini index, the income share held by the top 10%, and the income share held by the bottom 10%.

The technical note proposes a new alternative for measuring income inequality (or the inequality of a distribution), named as the "variation index (VI)". In the following sections, Section 2 gives the definition

 doi.org/10.32388/KUE8YA 

of the proposed variation index (VI). Section 3 presents examples including nine hypothetical income distributions and eight continuous distributions. Sections 4 provides discussion and conclusion.

## 2. The proposed variation index (VI)

Consider a distribution of income *X* with population mean $\mu$ and variance $\sigma^2$. The baseline for measuring income inequality is the case of perfect equality, that is, when every individual has exactly the same income. In statistical terms, this corresponds to an income distribution with zero variance, which is a degenerate distribution denoted by $X_E : \{\mu,\ \mu,\ \ldots\ \mu\}$. This degenerate distribution serves as a baseline for measuring the inequality of the given income distribution *X*.

The difference between *X* and $X_E$ is written as

$$\Delta = X - X_E \ = X - \mu. \tag{1}$$

Thus, *X* can be written as

$$X = \mu + \Delta. \tag{2}$$

Taking squares on both sides of Eq. (1) yields

$$X^2 = \mu^2 + 2\mu\Delta + \Delta^2. \tag{3}$$

Then taking the expectation on both sides of Eq. (2) yields

$$E(X^2) = E(\mu^2) + 2\mu E(\Delta) + E(\Delta^2). \tag{4}$$

The expectation of $\mu^2$ is the same as $\mu^2$, the expectation of $\Delta$ is zero because $E(X) = \mu$, and the expectation of $\Delta^2$ is the same as the variance of $X$. Therefore, Eq. (4) can be rewritten as

$$E(X^2) = \mu^2 + Var(X) = \mu^2 + \sigma^2. \tag{5}$$

In statistics, $\sqrt{E(X^2)}$ is called the root mean square (RMS), which is an overall (a kind of "average") magnitude of $X$.

**Definition.** The proposed variation index (VI) is defined as the ratio between the standard deviation: $\sigma = \sqrt{Var(X)}$ and the root mean square (RMS): $\sqrt{E(X^2)} = \sqrt{\mu^2 + \sigma^2}$. That is,

$$VI = \sqrt{\frac{Var(X)}{E(X^2)}} = \frac{\sigma}{\sqrt{\mu^2 + \sigma^2}} \ . \tag{6}$$

Thus, the proposed VI is a normalization of the standard deviation $\sigma$ by the RMS. Note that $\sigma$ is the average deviation of *X* from its mean $\mu$ or the baseline $X_E$. Therefore, the proposed VI quantitatively


doi.org/10.32388/KUE8YA

measures the degree of variation of $X$ relative to its overall magnitude. This direct connection to "variation" is why we call it the variation index.

If the population mean and variance are unknown, for a given dataset: $\{x_1, x_2, ..., x_i, ... x_n\}$, the proposed VI can be estimated as

$$\widehat{VI} = \sqrt{\frac{s^2}{\overline{x}^2 + s^2}} = \frac{s}{\sqrt{\overline{x}^2 + s^2}}. \tag{7}$$

where $\overline{x}$ is the sample mean and $s^2$ is the sample variance.

# 3. Examples

## *3.1. Nine hypothetical income distributions*

Kim et al.[2] examined nine hypothetical income distributions denoted by $X_1$, $X_2$, ... and $X_8$. They calculated the CV, the Gini coefficient ($G$), and their proposed $L_2$ index. Table 1 shows the nine hypothetical distributions and their results, along with the VI proposed in this study. Note that all the income distributions except $X_8$ have $\mu$=3.

 doi.org/10.32388/KUE8YA 

| $X$ | CV | $G$ | $L_2$ | VI | VI/CV | VI/$G$ | VI/$L_2$ |
|---|---|---|---|---|---|---|---|
| $X_1$: {1,2,2,5,5} | 0.5578 | 0.2933 | 0.4408 | 0.4871 | 0.8733 | 1.6608 | 1.1051 |
| $X_2$: {0,3,3,4,5} | 0.5578 | 0.2933 | 0.6035 | 0.4871 | 0.8733 | 1.6608 | 0.8072 |
| $X_3$: {2,2,2,4,5} | 0.4216 | 0.2133 | 0.2206 | 0.3885 | 0.9214 | 1.8214 | 1.7612 |
| $X_4$: {1,3,3,3,5} | 0.4216 | 0.2133 | 0.4230 | 0.3885 | 0.9214 | 1.8214 | 0.9185 |
| $X_5$: {1,3,3,4,4} | 0.3651 | 0.1867 | 0.4967 | 0.3430 | 0.9393 | 1.8372 | 0.6906 |
| $X_6$: {2,3,3,3,4} | 0.2108 | 0.1067 | 0.2451 | 0.2063 | 0.9785 | 1.9333 | 0.8416 |
| $X_7$: {-1,4,4,4,4} | 0.6667 | -- | 0.9029 | 0.5547 | 0.8321 | 2.0799 | 0.6144 |
| $X_8$: {0,4,4,4,4} | 0.5333 | 0.2000 | 0.7498 | 0.4706 | 0.8824 | 2.3529 | 0.6276 |
| $X_9$: {1,2,3,4,5} | 0.4714 | 0.2667 | 0.4200 | 0.4264 | 0.9045 | 1.5988 | 1.0152 |

**Table 1.** Nine hypothetical income distributions and the corresponding CVs, Gini coefficients ($G$), $L_2$, and VI

According to the CV, $G$, and VI values, the income inequality of distributions $X_1$ and $X_2$ is the same, and likewise, the income inequality of $X_3$ and $X_4$ is the same. This is expected because two income distributions with the same variation should yield the same CV, $G$, and VI values. In contrast, the $L_2$ index is more sensitive to changes in distribution[2]. Note that $X_7$ contains a negative income value; as a result, the Gini coefficient for $X_7$ is undefined. If the negative value in $X_7$ were deleted, the resulting distribution would reflect perfect equality, and if the negative value is replaced with zero, then $X_7$ would become distribution $X_8$.

Table 1 also shows the ratios VI/CV, VI/$G$, and VI/$L_2$ for all nine distributions. In all cases, the ratio VI/CV is less than 1. This is because the VI is the standard deviation normalized by the RMS, while the CV is the standard deviation normalized by the mean, and the RMS is always greater than the mean. On the other hand, the ratio VI/$G$ is greater than 1 for all distributions, reflecting that the VI and the Gini coefficient operate on different scales. In contrast, the ratio VI/$L_2$ varies: it exceeds 1 for six distributions and falls


doi.org/10.32388/KUE8YA

below 1 for three, indicating that the VI can be greater or less than the $L_2$ index depending on the shape of the distribution.

### *3.2. Eight continuous distributions*

Consider the following eight continuous distributions: Dirac delta function, uniform A, uniform B, normal, exponential, Pareto, Gamma, and Weibull. Table 2 summarizes the formulas for the PDF, Gini coefficient ($G$), and VI for these distributions. The formulas for the VI are derived in this study, while the formulas for the Gini coefficient are sourced from Wikipedia.

| Distribution | PDF $p(x)$ | G | VI |
|---|---|---|---|
| Dirac delta function | $\delta(x - x_0), \quad x_0 > 0$ | 0 | 0 |
| Uniform A | $\begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0, & otherwise \end{cases}$ | $\dfrac{(b-a)}{3(b+a)}$ | $\dfrac{(b-a)}{\sqrt{3(a+b)^2 + (b-a)^2}}$ |
| Uniform B | $\begin{cases} \frac{1}{2a}, & -a \leq x \leq a \\ 0, & otherwise \end{cases}$ | -- | 1 |
| Normal | $\dfrac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$ | -- | $\dfrac{\sigma}{\sqrt{\mu^2 + \sigma^2}}$ |
| Exponential | $\lambda \exp(-\lambda x), \quad x > 0$ | $\dfrac{1}{2}$ | $\dfrac{1}{\sqrt{2}}$ |
| Pareto | $\begin{cases} \frac{\alpha x_m^\alpha}{x^{\alpha+1}} & x \geq x_m \\ 0 & x < x_m \end{cases}$ | $\begin{cases} 1, & 0 < \alpha < 1 \\ \frac{1}{2\alpha-1}, & \alpha \geq 1 \end{cases}$ | $\begin{cases} 1, & \alpha \leq 2 \\ \frac{1}{\alpha-1}, & \alpha > 2 \end{cases}$ |
| Gamma | $\dfrac{1}{\Gamma(\alpha)\theta^\alpha} x^{\alpha-1} \exp\left(-\frac{x}{\theta}\right)$ | $\dfrac{\Gamma\left(\frac{2\alpha+1}{2}\right)}{\alpha\Gamma(\alpha)\sqrt{\pi}}$ | $\dfrac{1}{\sqrt{1+\alpha}}$ |
| Weibull | $\dfrac{k}{\lambda^k} x^{k-1} \exp\left(-\frac{x^k}{\lambda^k}\right)$ | $1 - 2^{-\frac{1}{k}}$ | $\sqrt{1 - \dfrac{\left(\Gamma\left(1+\frac{1}{k}\right)\right)^2}{\Gamma\left(1+\frac{2}{k}\right)}}$ |

**Table 2.** Formulas for the PDF, Gini coefficient ($G$), and VI for nine continuous distributions

As can be seen from Table 2, both the Gini coefficient and the VI equal 0 for the Dirac delta function, which is expected because a degenerate (single-point) distribution exhibits perfect equality with no variation. This result for the VI can also be obtained by taking the limit as $\sigma \to 0$ in the VI formula for the normal distribution.

It is noteworthy that the Gini coefficient does not exist for certain cases such as Uniform B or the normal distribution when these distributions include negative values. This is because the Gini coefficient does


doi.org/10.32388/KUE8YA

not account for negativity in incomes. In contrast, the proposed VI does not have this issue.

Furthermore, the VI for Uniform A is 1 because it is centered at 0, and similarly, the VI for the normal distribution with $\mu$=0 is also 1. This shows that when a distribution is symmetrical around zero, it has the greatest inequality according to the VI.

Figure 1 shows a comparison of three inequality measures: the CV, Geni coefficient, and VI for the Gamma distribution as a function of the shape parameter $\alpha$. The CV for the Gamma distribution is given by $1/\sqrt{\alpha}$ . Figure 2 shows a similar comparison for the Weibull distribution. The CV for the Weibull distribution is given by $\sqrt{\Gamma\left(1+\frac{2}{k}\right)/\left(\Gamma\left(1+\frac{1}{k}\right)\right)^2-1}$.

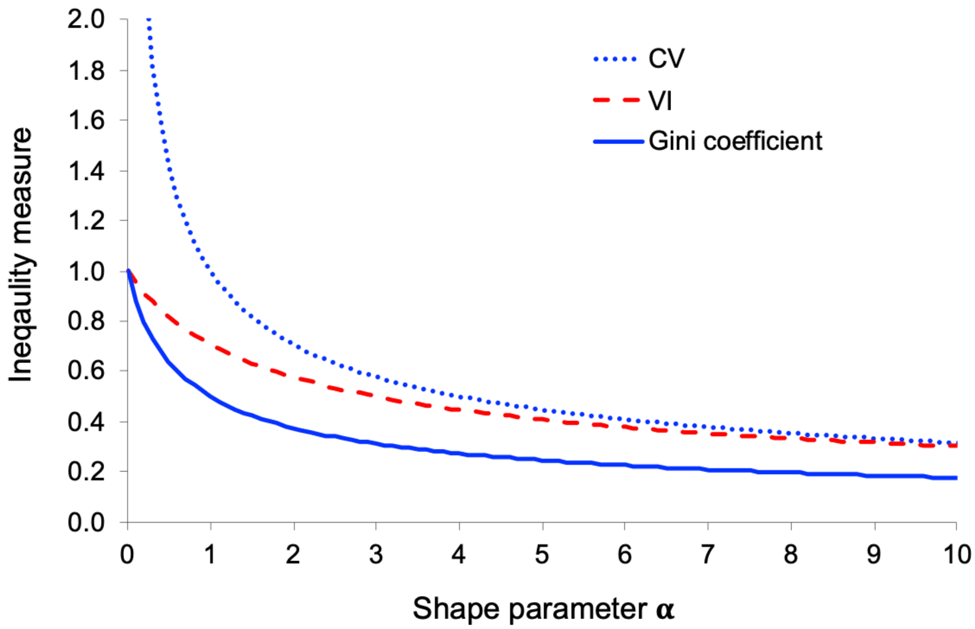


**Figure 1.** Comparison of the three inequality measures for the Gamma distribution as a function of the shape parameter $\alpha$


doi.org/10.32388/KUE8YA

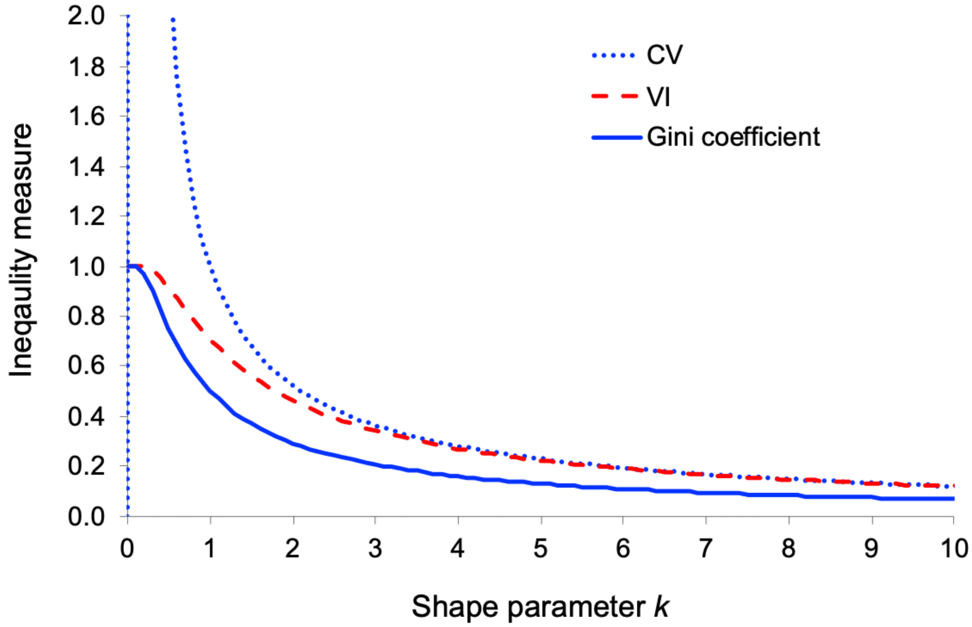


**Figure 2.** Comparison of the three inequality measures for the Weibull distribution as a function of the shape parameter $k$

As can be seen from Figure 1 or Figure 2, the CV does not have an upbound and tends to infinity when the shape parameter ($\alpha$ or $k$) approaches 0. In contrast, both the Gini coefficient and the VI remain bounded by 1 as the shape parameter approaches 0. Moreover, as the shape parameter becomes large, the CV converges to the VI. It is noteworthy that the Gini coefficient is always smaller than the CV or VI, except in the limit case where $\alpha = 0$ or $k = 0$, for which $G$ = VI = 1.

Figure 3 shows the plots of the ratio VI/CV against the shape parameter for both the Gamma and Weibull distributions. As the shape parameter approaches zero, the ratio VI/CV falls toward zero, reflecting the fact that CV $\to \infty$ in this limit. As the shape parameter becomes large, the ratio VI/CV converges to a constant value of approximately 0.96 for the Gamma distribution and 0.99 for the Weibull distribution.

Figure 4 shows the plots of the ratio VI/$G$ against the shape parameter for the two distributions. When the shape parameter is near zero, VI/$G \approx 1$, because both the VI and the Gini coefficient approach 1 in this limit. As the shape parameter becomes large, the ratio VI/$G$ converges to a constant value of approximately 1.72 for the Gamma distribution and 1.80 for the Weibull distribution.


doi.org/10.32388/KUE8YA

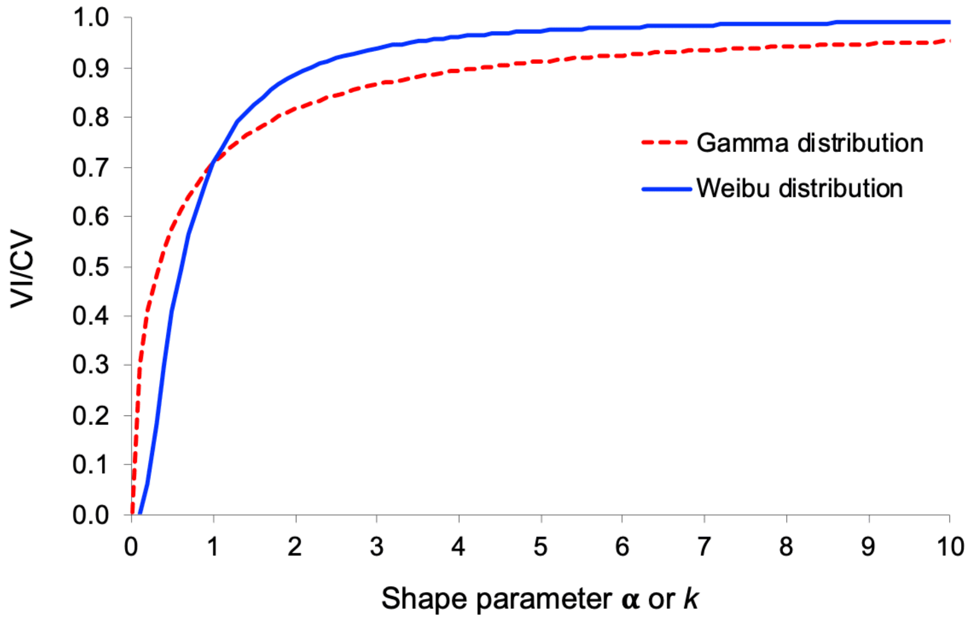


**Figure 3.** The ratio VI/CV for the Gamma and Weibull distributions as functions of the shape parameters

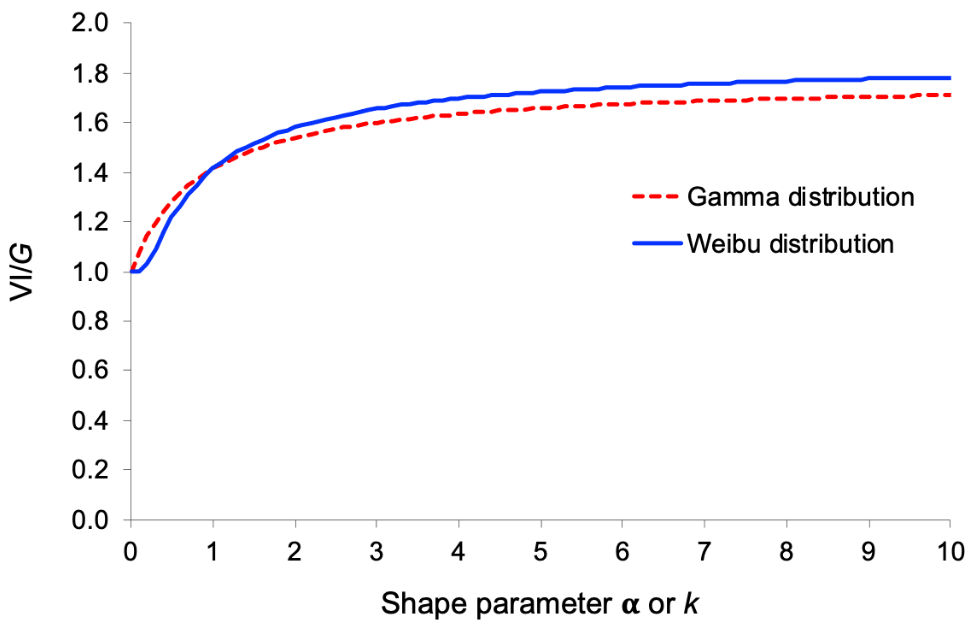


**Figure 4.** The ratio VI/$G$ for the Gamma and Weibull distributions as functions of the shape parameters

## 4. Discussion and conclusion

The proposed VI is both mathematically sound and intuitively interpretable. It is a dimensionless index with the desirable property of being bounded between 0 and 1, where 0 indicates perfect equality and 1 indicates extreme inequality (maximal variation relative to the RMS). Low VI values (close to 0) suggest an even distribution, while high VI values (close to 1) indicate a highly uneven distribution. Moreover, the VI remains meaningful even if the mean is zero or near zero. In contrast, the CV does not have an upbound and can yield misleading results when the mean is near zero. Consequently, the VI is a superior measure of inequality compared to the CV.

Like the $L_2$ index proposed by Kim et al.[2], the VI can be applied to income distributions that include negative incomes. In contrast, the Gini coefficient usually requires the assumption that income is non-negative.

Similar to the Gini coefficient, the key advantage of the VI is that it provides a single, easy-to-interpret statistic summarizing the inequality of an entire income distribution, with values bounded between 0 and 1. This boundedness facilitates comparisons across countries with differing population sizes. However, the VI shares a similar drawback with the Gini coefficient: two or more countries having the same VI value does not necessarily imply that they have the same level of income inequality.

It is important to note that, although both the VI and the Gini coefficient are bounded between 0 and 1, their scales are different. The VI is the standard deviation normalized by the RMS; the standard deviation measures the average deviation of $X$ from its mean $\mu$, i.e. relative to the baseline degenerate distribution $X_E : \{\mu,\ \mu,\ \ldots\ \mu\}$. The Gini coefficient is half the mean absolute difference between all pairs of observations, normalized by the arithmetic mean (equivalent to the definition based on the Lorenz-curve[5]). Unlike the standard deviation, the mean absolute difference is not anchored to a single reference value. Because the VI quantifies the overall variation around a central baseline, while the Gini coefficient quantifies the overall pairwise differences, they have different scales and will respond differently to the same data. Practitioners should keep these scale differences in mind when comparing the VI and the Gini coefficient. In fact, every inequality index has its own scaling nuances, which must be considered for meaningful interpretation and comparison in practical applications.

In summary, the proposed variation index (VI) provides an alternative to traditional inequality indices such as the coefficient of variation (CV) and the Gini coefficient. The presented examples have demonstrated the effectiveness of the proposed VI in capturing and quantifying distribution inequality.

## Statements and Declarations

### *Conflicts of Interest*

The author declares no conflicts of interest.

## References


1. ^ *De Maio FG. (2007). Income inequality measures. Journal of Epidemiology & Community Health. 61(10):849–852. doi:10.1136/jech.2006.052969.*
2. a, b, c, d *Park J, Kim Y, Ju AJ. (2024). New approach to measuring income inequality. Heliyon. 10(4):e26482. doi:10.1016/j.heliyon.2024.e26482.*
3. ^ *Park J, Kim Y, Ju AJ. (2021). Measuring income inequality based on unequally distributed income. J Econ Interact Coord. 16:309–322. doi:10.1007/s11403-020-00295-1.*
4. ^ *Sitthiyot T, Holasut K. (2020). A simple method for measuring inequality. Palgrave Commun. 6:112. doi:10.1057/s41599-020-0484-6.*
5. ^ *Sen A. (1977). On Economic Inequality. 2nd ed. Oxford: Oxford University Press.*


### Declarations


**Funding:** No specific funding was received for this work.

**Potential competing interests:** No potential competing interests to declare.