

# On the use of chance-adjusted agreement statistic to measure the assortative transmission of infectious diseases

Keisuke Ejima · Kazuyuki Aihara · Hiroshi Nishiura

Received: 24 May 2012 / Revised: 4 October 2012 / Accepted: 4 October 2012 /  
Published online: 16 April 2013  
© SBMAC - Sociedade Brasileira de Matemática Aplicada e Computacional 2013

**Abstract** There have been only a small number of statistical measures to assess the assortativeness. The present study discusses the applicability of two chance-adjusted agreement statistics, kappa and AC1 as measures of the assortative transmission of infectious diseases. First, we show that the so-called assortativity coefficient corresponds to the proportion of contacts that are spent for within-group mixing in the preferential mixing formulation of heterogeneous transmission, and also that the assortative coefficient is identical to the Cohen's kappa statistic. Second, we demonstrate that the kappa statistic is vulnerable to the paradoxes in measuring infectious disease transmission, because the assortative transmission involves not only contact heterogeneity but also other intrinsic and extrinsic factors including relative susceptibility and infectiousness. AC1 can be a useful measure due to its paradox resistant nature, and we discuss the relevance of preferential mixing formulation to the computation of AC1.

**Keywords** Epidemiology · Network · Mathematical model · Cluster · Inter-observer variation

---

Communicated by Antonio Silva Neto.

---

K. Ejima · K. Aihara  
Department of Mathematical Informatics, Graduate School of Information Science and Technology,  
The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan

K. Ejima · H. Nishiura  
School of Public Health, The University of Hong Kong, Level 6, Core F,  
Cyberport 3, Pokfulam, Hong Kong

K. Aihara  
Institute of Industrial Science, The University of Tokyo, 4-6-1 Komaba,  
Meguro-ku, Tokyo 153-8505, Japan

H. Nishiura (✉)  
PRESTO, Japan Science and Technology Agency, Saitama, Japan  
e-mail: nishiura@hku.hk

**Mathematics Subject Classification (2000)** Primary 92B05; Secondary 62P10

## 1 Introduction

Assortativity is defined as a preference for a network's node to attach to others that have similar characteristics or in some way different characteristics (Newman 2002). Adding the assortativity to a mathematical model often helps us to closely capture and approximate the dynamics in real world, which has been in particular demonstrated in the transmission dynamics of infectious diseases (Jacquez et al. 1998; Nishiura et al. 2010). Provided that an infectious disease (e.g. pandemic influenza A (H1N1-2009)) is frequently transmitted within a group of individuals that share similar characteristics (e.g. school children), the counter measures of the epidemic should ideally focus on those specific groups or their neighbors to effectively curb the epidemic (e.g. school closure) (Nishiura et al. 2011; Lam et al. 2011).

The assortativity is not only applicable to individual-based datasets but can also be incorporated into approximate modeling framework when we employ a population-based dynamic model, i.e. even when we use a model with discrete type space, the assortativity can be analytically devised into the model in order to approximately capture the realistic transmission dynamics (Jacquez et al. 1998; Nold 1996). For instance, an epidemic model with the so-called "preferential mixing" assumption can be written by a set of ordinary differential equations (Kiss et al. 2009), and the assortativity is eventually quantified as one of model parameters based on an epidemic data (Fraser et al. 2009). Of course, not only by fitting the mathematical model to the epidemiological data but also by conducting a field survey of socially defined contact in a population, one can compute and quantify the assortative mixing of the heterogeneous transmission model (Wallinga et al. 2006; Mossong et al. 2008; Del Valle et al. 2007).

Despite these theoretically useful characteristics, there have been only a few statistical measures to quantify the assortativeness. The most straightforward measure of assortativeness may be the correlation between the degrees of linked pairs of nodes (Newman 2003), but the correlation coefficient only captures the extent of linear association, rather than the propensity of assortative mixing. Farrington et al. (2009) therefore, proposed the use of mean-squared deviation from assortativeness as an index of absolute disassortativeness. However, the proposed measure has remained as applicable to assess assortativeness in a population with continuous type space. Although discrete type space (e.g. mixing within and between age-groups rather than individual network with continuous age) is more relevant to analyzing widely available empirical data in practical setting (e.g. epidemiological surveillance data classified by discrete age groups), the measure of assortativity for the discrete data has yet to be discussed more than the original description by Newman (2002).

In this study, we aim to discuss the applicability of two known chance-adjusted agreement statistics, kappa and AC1 to measure the assortativeness of infectious disease transmission. In particular, we aim to show that AC1 statistic can address known paradoxes of kappa, and thus, perhaps allows us to assess the assortativeness of transmission more appropriately than kappa. We first review the existing measures of assortativity in the next section, which is subsequently followed by a description of our motivations and computation of AC1.

## 2 Existing measures of assortativeness

In the following, we denote the contact rate between host groups  $i$  and  $j$  by  $c_{ij}$ . Let the sum of all the elements of the contact matrix  $\{c_{ij}\}$  be  $C$ , we denote the normalized contact rate

by  $e_{ij}$ . The sums over a single row and single column of the normalized contact matrix are, respectively, denoted by  $a_i$  and  $b_j$ , namely,

$$a_i = \sum_j e_{ij}, \tag{1}$$

$$b_j = \sum_i e_{ij}. \tag{2}$$

The assortativity coefficient,  $r$ , proposed by Newman (2003), is written as:

$$r = \frac{\sum_i e_{ii} - \sum_i a_i b_i}{1 - \sum_i a_i b_i}, \tag{3}$$

where the trace of matrix  $\{e_{ij}\}$  gives the observed fraction of within-group contacts, while the product of marginal sums is interpreted as the fraction of within-group contacts that occur by chance. The assortativity coefficient  $r$  typically takes the value from 0 to 1 with  $r = 1$  indicating perfect assortative mixing, while  $r = 0$  means random mixing. The measure is based on cross-classification of existing contacts. As the probability of within-group contact is calculated as the product of marginal sums of all columns and rows, it should be noted that the probability of assortative mixing is evaluated as if all observed contacts may result in within-group contact by chance.

Prior to the coefficient  $r$ , there was an earlier measure in epidemiology, proposed by Gupta et al. (1989). The earlier measure intended to quantify the impact of mixing patterns of sexual contacts on the spread of HIV epidemic. The  $Q$  statistic, a measure of the degree of within-group mixing, was proposed as:

$$Q = \frac{1}{m - 1} \sum_i \frac{e_{ii} - a_i b_i}{a_i}, \tag{4}$$

where  $m$  is the number of node types. The measure captures assortativeness, varying between  $-1/(m - 1)$  (minimally disassortative) and 1 (maximally assortative).  $Q$  is regarded as an ad hoc measure of assortativeness, because the quantity is interpreted as the proportion of contacts that occur along the main diagonal of the contact matrix. However,  $Q$  was later shown to be vulnerable to grouping of hosts used to define the diagonal of the contact matrix and to be sensitive to different sub-population sizes between different types of host (Newman 2003). Accordingly, we focus on the assortative coefficient  $r$  in the following discussion.

An interesting property of  $r$  in (3) is that the measure is consistent with the classical preferential mixing formulation in an approximate modeling approach. Let  $p$  be the proportion of contact that is spent for within-group mixing among the total contacts. The contact rate  $c_{ij}$  is then modeled as a simple mixture of an assortative mixing component and a proportionate mixing component:

$$c_{ij} \propto \begin{cases} (1 - p)n_i, & \text{if } i \neq j, \\ p + (1 - p)n_i, & \text{if } i = j, \end{cases} \tag{5}$$

where  $n_i$  is the relative population size of host  $i$ . In order to calculate assortativity coefficient, we normalize  $c_{ij}$  as

$$e_{ij} \propto \begin{cases} \frac{(1-p)n_i}{m}, & \text{if } i \neq j, \\ \frac{p+(1-p)n_i}{m}, & \text{if } i = j, \end{cases} \tag{6}$$

in which it is certain that  $e_{ij}$  sums up to 1. It is not difficult to find that the parameter  $p$  exactly corresponds to  $r$ , because we have

$$a_i \propto \frac{p + \sum_j (1-p)n_j}{m} = \frac{p + (1-p)mn_i}{m}, \quad (7)$$

where  $m$  is again the number of host types, and

$$b_j \propto \frac{p + \sum_i (1-p)n_i}{m} = \frac{1}{m}, \quad (8)$$

leading the proportion of contact  $p$  to be identical to  $r$ . This indicates that the interpretation of the assortativity coefficient in relation to its underlying contact mechanism can be as simple as that shown in the mixture model (5) in which only the proportionate mixing component is expected to explain between-group contact frequency. To be strict, the mixture model (5) is unlikely to hold in practice, and thus, rather than using the Kronecker delta-type assumption in (5), the use of distribution to describe the influence of preferential mixing has been proposed elsewhere (Glasser et al. 2012).

### 3 Vulnerability of kappa to assortative transmission

When mathematical models are applied to describe infectious disease epidemics, two different types of matrix should be explicitly distinguished. One is the contact matrix  $\{c_{ij}\}$  describing the contact rates per unit time within and between groups of host. As described in model (5), the mixture type assumption may be employed to parameterize  $\{c_{ij}\}$  in the simplest manner. For clarity, hereafter we refer to the assortativeness of  $\{c_{ij}\}$  as “contact assortativity”.

On the other hand, there is a different matrix  $\mathbf{K} = \{k_{ij}\}$ , which is more relevant to the transmission dynamics, gives the average number of secondary cases in host  $i$  generated by a single primary case of host  $j$  throughout its entire course of infectiousness in a fully susceptible population. The matrix is referred to as the next-generation matrix (Diekmann et al. 2010), mapping the distribution of secondary cases based on that of primary cases, describing the heterogeneous patterns of transmission in a single generation of transmission event. Each element  $k_{ij}$  is dimensionless. Other than the contact frequency, the frequency of infectious disease transmission is regulated by susceptibility of exposed individuals, infectiousness of primary cases and other factors (including biological and non-biological ones), and the next-generation matrix captures these features as well as the contact heterogeneity. Using the above-mentioned mixture type of contact, let  $\alpha_i$  and  $\beta_j$  represent age-specific susceptibility and infectiousness of hosts of type  $i$  and  $j$ , respectively,  $\{k_{ij}\}$  may be parameterized as

$$k_{ij} \propto \begin{cases} \alpha_i \beta_j (1-p)n_i, & \text{if } i \neq j, \\ \alpha_i \beta_j p + \alpha_i \beta_j (1-p)n_i, & \text{if } i = j, \end{cases} \quad (9)$$

as was used in practical applications elsewhere (Nishiura et al. 2011; Lam et al. 2011; Fraser et al. 2009). Hereafter, we refer to the assortativeness of  $\{k_{ij}\}$  as “transmission assortativity”.

Here, the distinction of two different types of assortativity, i.e. contact and transmission, is made, because the transmission is not only characterized by contact but also by all other intrinsic and extrinsic factors including  $\alpha_i$  and  $\beta_j$  in model (9). For example, when children are far more susceptible to influenza than adults (which is believed as the case based on empirical evidence (Nishiura and Oshitani 2011)), the transmission assortativity would be the result of contact assortativity (with high frequency of child-to-child contacts) weighted by high relative susceptibility among children due to model (9). In such an instance, the transmission

**Table 1** The age-dependent next generation matrix and the corresponding agreement

|                  | A <sup>b</sup> |       | B <sup>b</sup> |       | C <sup>b</sup> |       |
|------------------|----------------|-------|----------------|-------|----------------|-------|
|                  | Child          | Adult | Child          | Adult | Child          | Adult |
| Child            | 1.34           | 0.32  | 1.61           | 0.32  | 1.34           | 0.02  |
| Adult            | 0.33           | 0.83  | 0.33           | 0.56  | 0.64           | 0.83  |
| obs <sup>a</sup> | 0.767          |       | 0.767          |       | 0.767          |       |
| Kappa            | 0.517          |       | 0.459          |       | 0.540          |       |
| AC1              | 0.548          |       | 0.590          |       | 0.548          |       |

<sup>a</sup> *obs* observed agreement

<sup>b</sup> Examples of the next-generation matrix. In matrix A, the parameters were fixed at  $n_c = 0.32$ ,  $\alpha_c = 2.06$ ,  $\alpha_a = \beta_c = \beta_a = 1$ , and  $\theta = 0.50$  (Lam et al. 2011). The basic reproduction number is given by the dominant eigenvalue and calculated as 1.5. Matrices B and C have identical observed agreement values with A, but matrix B increased the frequency of child-to-child transmission by 1.2 times as that in A and matrix C increased the frequency of child-to-adult transmission by 1.9 times as that in A

assortativity requires a particular attention in appropriately quantifying the propensity of within-group contacts that are made by chance.

Here, we consider the chance-adjusted agreement measure. Although not explicitly mentioned by Newman (2003), the assortativity coefficient (3) is mathematically identical to the so-called Cohen’s kappa statistic (Cohen 1960), which is known as the most commonly used chance-adjusted agreement measure for multiple ratings. In the case of infectious disease transmission, there are only two raters, i.e. contactor and contactee, with discrete grouping of choices such as age-groups. In other words, as long as the matrix captures the transmission between a pair of individuals (i.e. one susceptible and one infectious host) over a single generation, the agreement statistic can be restricted to the case of two raters. Although Cohen’s kappa is more robust measure than simple calculation of observed agreement, it is also known that there are situations in which the kappa yields unexpected results. The phenomenon is referred to as the paradoxes of kappa (Feinstein and Cicchetti 1990a,b), and this is directly relevant to considering the transmission assortativity.

The paradoxes can be illustrated by considering the next-generation matrix adapted from Lam et al. (2011), which employed the mixture-type assumption for contact and also described the transmission dynamics of pandemic influenza (H1N1-2009) within and between populations of children and adults using model (9) (Table 1). We consider three different matrices, A, B and C. As for the baseline matrix A, we follow the parameterization of model (9), assuming that  $n_c = 0.32$ ,  $\alpha_c = 2.06$ ,  $\alpha_a = \beta_c = \beta_a = 1$ , and  $p = 0.50$  (Lam et al. 2011), where subscripts *c* and *a* stand for children and adults, respectively. The basic reproduction number, the average number of secondary cases generated by a single primary case in a fully susceptible population, is calculated as the dominant eigenvalue of **K**, and in this example set at 1.5. Within-group transmission, which is measured by the observed agreement, is seen in 76.7 % of all secondary transmissions, while the chance-adjusted measure, kappa is calculated as 0.517. In Matrix B, the frequency of child-to-child transmission is magnified by 1.2 times as compared to matrix A, and the increment of the secondary transmissions among children is reduced from adult-to-adult transmission (so that the total of within-group secondary transmissions is kept as identical to matrix A). Other two elements, between-group transmission frequencies are unaltered from matrix A. Of course, the observed agreement of matrix B remains the same as matrix A, because the sum of diagonal elements is unaltered. However, kappa is calculated as 0.459. Namely, by magnifying the within-group transmission in a

specific single host type (i.e. children), the chance-adjusted agreement statistic was reduced without any sensible reason. This is referred to as the kappa’s paradox I.

In Matrix **C**, the frequency of child-to-adult transmission is magnified by 1.9 times to that of matrix **A**. The sum of anti-diagonal elements is kept identical to matrices **A** and **B**, and the diagonal elements are unaltered from matrix **A**. Again, the observed agreement of matrix **C** is calculated as 76.7 %, identical to those from matrices **A** and **B**. However, kappa is calculated as 0.540. By introducing the bias in between-group transmission, the kappa statistic was elevated. This increase in kappa owing to the bias in non-diagonal elements is referred to as the kappa’s paradox II.

These paradoxes may be unlikely to matter a lot for contact assortativity, while the introduction of host-specific characteristics such as  $\alpha_i$  and  $\beta_j$  in model (9) to describing the transmission assortativity can easily lead to observing the paradoxes (see below). In other words, the assortativity coefficient (3) (which is mathematically identical to kappa statistic) could be vulnerable as a measure of transmission assortativity, especially when the assortativity of transmission introduces the sources of paradoxes I and II to the contact matrix.

### 4 Comparison between kappa and AC1

As a paradox-resistant measure of agreement, [Gwet \(2010\)](#) has proposed the so-called AC1 statistic in which AC stands for “agreement coefficient”. Let  $\gamma$  be the coefficient of transmission assortativity given by the AC1, and is written as:

$$\gamma = \frac{\sum_i e_{ii} - p_e}{1 - p_e}, \tag{10}$$

where  $p_e$  is the chance agreement probability. The right-hand side of (10) is conceptually the same as Cohen’s kappa in which  $p_e$  was calculated as a summation of the product of two marginal sums. In the case of AC1 statistic, it is calculated as

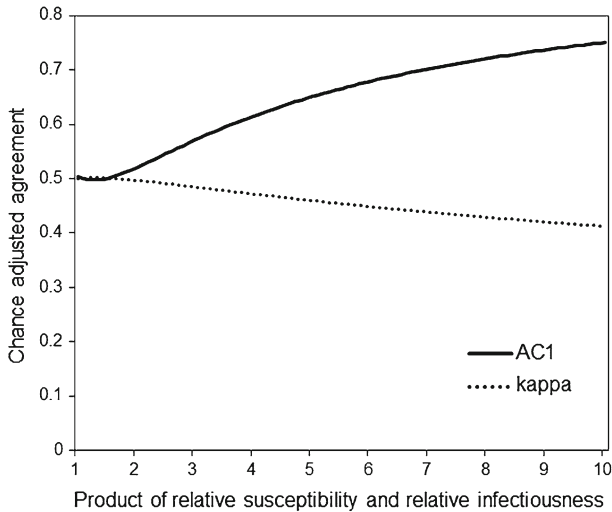
$$p_e = \frac{1}{m - 1} \sum_{k=1}^m \pi_k(1 - \pi_k), \tag{11}$$

where  $\pi_k$  is the average of marginal sum over row  $k$  and column  $k$ , i.e.,

$$\pi_k = \frac{\sum_j k_{kj} + \sum_i k_{ik}}{2 \sum_i \sum_j k_{ij}}. \tag{12}$$

The chance agreement for AC1 is calculated as shown in (11), because AC1 considers the chance agreement as the product of (i) the probability that two raters agree given that the subject being rated was assigned a non-deterministic score (i.e. the probability of simple chance agreement is  $1/m$ ) and (ii) the propensity that a rater will assign a non-deterministic score, which is estimated by the ratio:  $\sum_{k=1}^m \pi_k(1 - \pi_k)/(1 - 1/m)$ .

Kappa statistic regards the chance agreement probability as if all observed ratings may yield an agreement by chance. However, [Gwet 2010](#) pointed out that this may lead to unpredictable results with agreement data that actually have a rather small propensity for chance agreement. This may be in many instances the case for the transmission of infectious diseases. The AC1 statistic considers the chance agreement as proportional to the portion of ratings and conditional on the random rating. By appropriately accounting for the propensity of chance agreement, AC1 successfully reduces the chance agreement to the right magnitude. Further



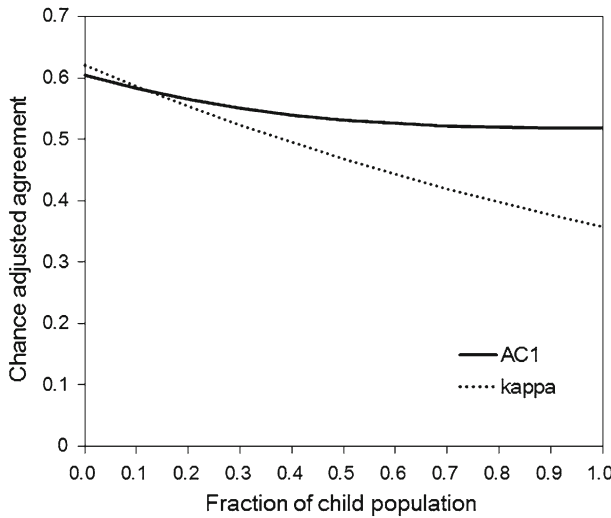
**Fig. 1** Comparison between kappa and AC1 statistics by the product of relative susceptibility and relative infectiousness among children

details, theoretical properties and examples related to the AC1 statistic are given elsewhere (Gwet 2008, 2010).

Table 1 shows the estimates of AC1 corresponding to each of the matrices **A**, **B** and **A**. Using the baseline matrix **A**, AC1 is estimated at 0.548. When the within-child transmission is increased (matrix **B**), AC1 is calculated as 0.590. When the bias of between-group transmission is introduced (matrix **C**), AC1 remains to be 0.548. The variation of AC1 was within 10 %, and perhaps more importantly, AC1 was not underestimated even when the matrix which induces paradox I is analyzed. As there is no perfect chance-adjusted agreement, AC1 is also not the perfect measure (i.e. not entirely free from conceptual error), but this statistic is regarded as far less vulnerable to known paradoxes of kappa statistic and can be strictly interpreted as the conditional probability that two randomly selected raters agree given that there is no agreement by chance (Gwet 2008). As long as the measure of assortativity employs the chance-adjusted agreement coefficient, the biggest concern of the transmission assortativity is the possibility to appropriately account for chance agreement, which indicates that AC1 suits to measure the transmission assortativity.<sup>1</sup>

As a numerical comparison between kappa and AC1, Fig. 1 shows the sensitivities of these measures to the product of relative susceptibility and infectiousness in the formulation (9). As  $\alpha$  and  $\beta$  among children are elevated, observed and actual within-group transmission would increase among the total of secondary transmissions. kappa is greatly influenced by paradoxes, especially paradox I due to a representation of child-to-child transmission. kappa

<sup>1</sup> The chance-adjusted agreements of secondary transmissions, derived from the next-generation matrix, are shown as a function of the product of relative susceptibility and relative infectiousness among children ( $\alpha_c$  and  $\beta_c$ ). The solid line shows the AC1, whereas the Cohen's kappa is drawn in dashed line. Other parameters for the next-generation matrix were fixed at  $n_c = 0.32$ ,  $\alpha_a = 1$ ,  $\beta_a = 1$  and  $p = 0.5$  among which  $p$  refers to the proportion of contacts that are spent for within-group mixing (Lam et al. 2011). It should be noted that the elements of the next-generation matrix are the frequencies of between- and within-group contacts weighted by relative susceptibility and relative infectiousness (and thus, the kappa value is different from 0.50) and here we examine the impact of kappa's paradox I (i.e. domination of child-to-child transmission) on the resulting chance-adjusted agreement coefficients.



**Fig. 2** Comparison between kappa and AC1 statistics by the proportion of children in a population

even decreases with the increase in the product of  $\alpha$  and  $\beta$  among children. Nevertheless, the increasing feature of within-group transmission is captured by AC1 in Fig. 1, avoiding underestimation of chance-adjusted agreement due to paradox I. Similarly, Fig. 2 examines the sensitivity of chance-adjusted agreement coefficients to the proportion of children in the population. As the fraction of child population size increases, the chance agreement increases, and thus, kappa and AC1 decrease. However, as the child-to-child transmission increases with an increase in the fraction of children, the kappa experiences greater decline than AC1 does due to paradox I.

AC1 statistic is regarded as more valid measure than kappa to evaluate the transmission assortativity, and its usefulness in practice may extend to the contact assortativity, especially in the case we observe clusters of contact only among specific types of host (e.g. clustering only among school-age children). However, kappa (or the classical assortativity coefficient) may be preferred for measuring the contact assortativity, because kappa has been known to mechanistically correspond to  $p$ , i.e. the proportion of contacts that are spent for within-group mixing, in the simplest form of preferential mixing assumptions (5). It is thus important to explore the relationship between  $p$  and computation of AC1 in the simplest model (5) of the contact assortativity.<sup>2</sup>

In the case of  $m$  different types of host, AC1 is written as

$$\gamma = \frac{\sum_i e_{ii} - \frac{1}{m-1} \sum_{k=1}^m \pi_k (1 - \pi_k)}{1 - \frac{1}{m-1} \sum_{k=1}^m \pi_k (1 - \pi_k)}, \tag{13}$$

<sup>2</sup> The chance-adjusted agreement of secondary transmissions, derived from the next-generation matrix, are shown as a function of the fraction of children in the population ( $n_c$ ). The solid line shows the AC1, whereas the Cohen’s kappa is drawn in dashed line. Other parameters for the next-generation matrix were fixed at  $\alpha_c = 1, \alpha_a = 1, \beta_c = 2.06, \beta_a = 1$  and  $p = 0.50$  among which  $p$  refers to the proportion of contacts that are spent for within-group mixing (Lam et al. 2011). It should be noted that the elements of the next-generation matrix are the frequencies of between- and within-group contacts weighted by relative susceptibility and relative infectiousness (and thus, the kappa value is different from 0.50) and here we examine the sensitivity of chance-adjusted agreement coefficients to differing proportion of the population.



where

$$\sum_i e_{ii} \propto mp + (1 - p), \quad (14)$$

and

$$\pi_k(1 - \pi_k) \propto [p + m(1 - p)n_k + 1] \left( 1 - \frac{p + m(1 - p)n_k + 1}{2} \right). \quad (15)$$

Although we cannot come up with further insightful analytical findings, one can notice that there are several special cases. If there is only a single type of host ( $m = 1$ ) constituting a population, both  $\gamma$  and  $p$  are not practically relevant measures, but they agree to be 1. If all the contacts are spent for within-group mixing ( $p = 1$ ), the corresponding AC1 statistic  $\gamma$  would also be 1.

When each subpopulation is equally distributed so that  $n_i = 1/m$  for any  $i$ , this would greatly simplify the chance agreement (15). The trace of the contact matrix is given by (14), and the chance agreement would be zero. Since the sum of all the elements of contact matrix is  $m$ , the AC1 is calculated as

$$\gamma = \frac{\sum_i e_{ii}}{m} = p + \frac{1 - p}{m}. \quad (16)$$

Two important messages from Eq. (16) are that (i) it indicates that  $\gamma$  is greater than  $p$  as long as the population is equally distributed. This may be regarded as consistent with the numerical results in Figs. 1 and 2. (ii) When there are so many different types of host (so that  $m \rightarrow \infty$ ), the difference between  $\gamma$  and  $p$  would be diminished and the two are approximated.

## 5 Discussion

The present study discussed the use of chance-adjusted agreement coefficients to measure the assortativity of contact and transmission of infectious diseases. We have demonstrated that  $p$  in the preferential mixing in infectious disease modeling has excellently corresponded to the Newman's assortativity coefficient (or Cohen's kappa). Subsequently, we have explicitly distinguished the transmission assortativity from contact assortativity, because the former captures not only the contact heterogeneity but also many other intrinsic and extrinsic factors characterizing the frequency of within- and between-group transmission. The distinction between the contact and the transmission was made, because kappa statistic is vulnerable to the paradoxes which are likely to be the case to assess the transmission assortativity. In such an instance, AC1 statistic, a relatively new chance-adjusted agreement coefficient, computed in similar way to kappa and not very computationally intensive measure, was shown to be paradox resistant. However, AC1 was shown to be less interpretable than kappa, and does not easily correspond to the mechanistically interpretable mixture model to describe the preferential mixing.

There is no doubt that each of the currently available agreement coefficients involves a variety of technical problems, and none has been regarded as perfect measure. In fact, it is well known that Cohen's kappa does not adjust for both chance agreement and misclassification errors. Although AC1 was shown to be paradox resistant, the statistic is not entirely free from the paradoxes, and moreover, our application has shown that it does not lead to useful mixing assumption in parameterizing the kinetics (i.e. mechanistic features) of transmission due to a difficulty in eliminating the relative population size in the chance agreement (15). In the future,

it is likely that multiple measures will be required to assess different aspects of the assortative network. In the context of assortativity, the strength may be measured by chance-adjusted agreement or correlation, and the propensity of contact (e.g. the distance between two different types of host) should also be measured by absolute disassortativeness (Farrington et al. 2009). The direction of the contact would also be an important issue in appropriately capturing the transmission dynamics on an explicit network (Meyers et al. 2006). The relevance of these topological aspects to mathematical formulation of the approximate heterogeneous transmission dynamics has yet to be explored (Ejima et al. 2012a,b).

As quantified in social contact surveys (Wallinga et al. 2006; Mossong et al. 2008; Del Valle et al. 2007), the actual heterogeneous mixing has been shown not to be well captured by classical model such as classical preferential mixing in model (5). As seen in an effort to capture the age-dependent heterogeneity using a contact surface (Farrington and Whitaker 2005), the model to be applied to empirically observed data needs to capture more realistic features than the mechanistic mixture model (9) does. As seen in an attempt by Glasser et al. (2012), more mathematical formulations would be required to express the assortative mixing as a measurable quantity so that we can implement the statistical estimation. However, it is also true that one of the simplest models to be employed and fitted to the early outbreak data with a discrete group structure would be the one-parameter preferential mixing model (Fraser et al. 2009). For this reason, we believe that this study has satisfied an essential need to emphasize the importance of measuring transmission assortativity using paradox-resistant change-adjusted agreement measure.

**Acknowledgments** HN received funding support from the Japan Science and Technology Agency (JST) PRESTO program. KE received scholarship support from the Japan Society for Promotion of Science (JSPS). KA received funding support from the Aihara Project, the FIRST program from JSPS, initiated by CSTP. This work also received financial support from the Harvard Center for Communicable Disease Dynamics from the National Institute of General Medical Sciences (grant no. U54 GM088558) and the Area of Excellence Scheme of the Hong Kong University Grants Committee (grant no. AoE/M-12/06). The funding bodies were not involved in the collection, analysis and interpretation of data, the writing of the manuscript or the decision to submit for publication.

## References

- Newman MEJ (2002) Assortative mixing in networks. *Phys Rev Lett* 89:208701
- Jacquez JA, Simon CP, Koopman J, Sattenspiel L, Perry T (1988) Modeling and analyzing HIV transmission: the effect of contact patterns. *Math Biosci* 92:119–199
- Nishiura H, Chowell G, Safan M, Castillo-Chavez C (2010) Pros and cons of estimating the reproduction number from early epidemic growth rate of influenza A (H1N1) 2009. *Theor Biol Med Model* 7:1
- Nishiura H, Cook AR, Cowling BJ (2011) Assortativity and the probability of epidemic extinction: a case study of Pandemic Influenza A (H1N1-2009). *Interdiscip Perspect Infect Dis* 2011:194507
- Lam EH, Cowling BJ, Cook AR, Wong JY, Lau MS, Nishiura H (2011) The feasibility of age-specific travel restrictions during influenza pandemics. *Theor Biol Med Model* 8:44
- Nold A (1980) Heterogeneity in disease-transmission modeling. *Math Biosci* 52:227–240
- Kiss IZ, Simon PL, Kao RR (2009) A contact-network-based formulation of a preferential mixing model. *Bull Math Biol* 4:888–905
- Fraser C, Donnelly CA, Cauchemez S, Hanage WP, Van Kerkhove MD, Hollingsworth TD, Griffin J, Baggaley RF, Jenkins HE, Lyons EJ, Jombart T, Hinsley WR, Grassly NC, Balloux F, Ghani AC, Ferguson NM, Rambaut A, Pybus OG, Lopez-Gatell H, Alpuche-Aranda CM, Chapela IB, Zavala EP, Guevara DM, Checchi F, Garcia E, Hugonnet S, Roth C, WHO Rapid Pandemic Assessment Collaboration (2009) Pandemic potential of a strain of influenza A (H1N1): early findings. *Science* 324:1557–1561
- Wallinga J, Teunis P, Kretzschmar M (2006) Using data on social contacts to estimate age-specific transmission parameters for respiratory-spread infectious agents. *Am J Epidemiol* 164:936–944

- Mossong J, Hens N, Jit M, Beutels P, Auranen K, Mikolajczyk R, Massari M, Salmaso S, Tomba GS, Wallinga J, Heijne J, Sadkowska-Todys M, Rosinska M, Edmunds WJ (2008) Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Med* 5:e74
- Del Valle SY, Hyman JM, Hethcote HW, Eubank SG (2007) Mixing patterns between age groups in social networks. *Soc Netw* 29:539–554
- Newman MEJ (2003) Mixing patterns in networks. *Phys Rev E* 67:026126
- Farrington CP, Whitaker HJ, Wallinga J, Manfredi P (2009) Measures of disassortativeness and their application to directly transmitted infections. *Biom J* 51:387–407
- Gupta S, Anderson RM, May RM (1989) Networks of sexual contacts: implications for the pattern of spread of HIV. *AIDS* 3:807–817
- Glasser J, Feng Z, Moylan A, Del Valle S, Castillo-Chavez C (2012) Mixing in age-structured population models of infectious diseases. *Math Biosci* 235:1–7
- Diekmann O, Heesterbeek JA, Roberts MG (2010) The construction of next-generation matrices for compartmental epidemic models. *J R Soc Interface* 7:873–885
- Nishiura H, Oshitani H (2011) Household transmission of influenza (H1N1-2009) in Japan: age-specificity and reduction of household transmission risk by zanamivir treatment. *J Int Med Res* 39:619–628
- Cohen J (1960) A coefficient of agreement for nominal scales. *Educ Psychol Measure* 20:37–46
- Feinstein AR, Cicchetti DV (1990) High agreement but low kappa: I. The problems of two paradoxes. *J Clin Epidemiol* 43:543–549
- Feinstein AR, Cicchetti DV (1990) High agreement but low kappa: II. Resolving the paradoxes. *J Clin Epidemiol* 43:551–558
- Gwet KL (2010) *Handbook of Inter-rater Reliability* (2nd edn). Advanced Analytics. Gaithersburg, MD, USA
- Gwet KL (2008) Computing inter-rater reliability and its variance in the presence of high agreement. *Br J Math Stat Psychol* 61:29–48
- Meyers LA, Newman ME, Pourbohloul B (2006) Predicting epidemics on directed contact networks. *J Theor Biol* 240:400–418
- Ejima K, Omori R, Cowling BJ, Aihara K, Nishiura H (2012a) The time required to estimate the case fatality ratio of influenza using only the tip of an iceberg: joint estimation of the virulence and the transmission potential. *Comput Math Methods Med* 2012:978901
- Ejima K, Omori R, Aihara K, Nishiura H (2012b) Real-time investigation of measles epidemics with estimate of vaccine efficacy. *Int J Biol Sci* 8:620–629
- Farrington CP, Whitaker HJ (2005) Contact surface models for infectious diseases: estimation from serologic survey data. *J Am Stat Soc* 100:370–379