Journal of Nonlinear Modeling and Analysis
Volume 7, Number 2, April 2025, 666–679
https://journal.global-sci.org/jnma
DOI:10.12150/jnma.2025.666

Improved Contact Tracing SIR Model for Randomly Mixed Populations*

Meili Li[1], Boxiang Yu[1], Junling Ma[2] and Manting Wang[2,†]

Abstract Contact tracing allows for more efficient quarantine and isolation, and is thus a key control measure in combating infectious diseases. Mathematical models that accurately describe the contact tracing process are important tools for studying the effectiveness of contact tracing. Recently, we developed a novel contact tracing SIR model based on pair dynamics, which uses pairs (two-individual) interactions to approximate triple (three-individual) interactions to close the model. However, the pair approximation used in the model is only a crude estimate. We extend this model to improve the approximation. Specifically, the new model tracks infectious individuals who have or have not infected others, as they play different roles in triples. We conduct a theoretical analysis to calculate the control reproduction number. The results of the new model are compared with those of the original model by numerical analysis. We find that the two models yield a similar epidemic final size. However, the original model yields a larger control reproduction number and thus underestimates the effect of contact tracing. This discrepancy increases as contact tracing is strengthened.

Keywords Compartmental disease model, control reproduction number, pair dynamics

MSC(2010) 34A34, 92D25, 92D30.

1. Introduction

Since the beginning of the COVID-19 pandemic, governments worldwide have widely adopted physical containment measures in an effort to control the spread [6]. These policies have significantly reduced disease transmission [12]. Contact tracing is an important epidemic prevention measure to reduce the spread of infectious diseases by blocking the chain of infection. Traditional contact tracing involves identifying the close contacts of an infected individual. These contacts are considered at risk of infection and are advised to take actions to reduce transmission, such as self-isolation [1]. Its effectiveness depends on several factors: timely detection

†the corresponding author.
Email address: stylml@dhu.edu.cn (M. Li), 2212272@mail.dhu.edu.cn (B. Yu), junlingm@uvic.ca (J. Ma), mantingwang@uvic.ca (M. Wang)
[1]School of Mathematics and Statistics, Donghua University, Shanghai, 201620, China
[2]Department of Mathematics and Statistics, University of Victoria, Victoria, BC, V8N 4V3, Canada
*The authors were supported by the National Natural Science Foundation of China (Grant No.12271088) (ML), Natural Science Foundation of Shanghai (Grant No.21ZR1401000) (ML) and a discovery grants of Natural Science and Engineering Research Council Canada (JM).

and isolation of indicator cases [9], timely and comprehensive identification of contacts [5], and quarantine compliance [7,11]. Technology can address some of these limitations by automating the processing of test results or symptom reports and utilizing mobile phones to identify and notify contacts at risk of infection [4].

Due to its importance, the mathematical modelling of contact tracing has been studied using multiple approaches. A contact network model for contact tracing was developed by [8,16], as close contacts can be naturally modeled by contact networks. However, network models require information about the contact network, which is often difficult to obtain. During the COVID-19 pandemic, much effort has been dedicated to collecting this information [15]. However, human contacts may be spatially, culturally, and economically specific, limiting the generalizability of these contact network studies to other regions. A traditional compartmental model that incorporates quarantine and isolation was first proposed by [3] to study the 2003 SARS pandemic. This model assumes that a fraction of the cases are contact traced. It has been adapted to study the COVID-19 pandemic [13]. However, it is difficult for these models to precisely describe contact tracing because traditional compartmental models assume that the population is randomly mixed and do not trace individual contacts. Branching process models for contact tracing have recently been developed [14] to study the COVID-19 pandemic. These models can precisely describe the contact tracing process for a single patient, thus yielding a precise control reproduction number. However, it is difficult to use these models to study disease dynamics without involving large-scale agent-based simulations.

Bednarski et al. [2] established a novel compartmental SIR contact tracing model for a randomly mixed population, thus avoiding the need for contact network information. This model tracks contacts using pairs, which are formed by disease transmission. The authors recognized that disease transmissions form a tree of transmissions. They borrowed the edge dynamics idea from network models to study contact tracing as a dynamic process on this tree.

Like the network models in [10,18], the dynamics of pairs depend on interactions between the individuals in the pair and other individuals, i.e., triple interactions, which in turn depend on four-individual interactions and so on. To truncate this infinite chain of dependences and simplify the model, triple interactions are approximated by pairs, using the triple closure method introduced in [10]. However, this is only a crude approximation, as it ignores the fact that not all infectious individuals play the same role in triple interactions. The specific approximation and its problems are explained in more detail in Subsection 2.1.

The goal of this paper is to improve this model by more accurately approximating triples. To do this, in Section 2, we extend the Bednarski et al. model to track infectious individuals who have or have not infected others. We calculate the control reproduction number and compare the simulation results of our new SIR model with the previous contact tracing SIR model [2] in Section 3.

2. Modelling contact tracing

2.1. The simple SIR contact tracing model

The SIR contact tracing model [2] divides the population into susceptible (S), infectious (I), diagnosed (T), contact tracing initiated (X), and recovered without being diagnosed (R) compartments. The infection process dynamically generates a

tree of transmission, where the nodes are infected individuals (including infectious, recovered, and diagnosed individuals who may or may not have triggered contact tracing). The edges represent transmissions, and the direction denotes who-infected-who, namely, from the infector to the infectee. An edge is labeled as $[U \leftarrow V]$, where U and V are the infection states of the infectee and the infector, respectively. Each new infection forms a $[I \leftarrow I]$ pair. The pair changes its state as the state of either the infector or the infectee changes due to disease progression or contact tracing. For example, the $[I \leftarrow I]$ pair becomes $[I \leftarrow T]$ when the infector is diagnosed.

The contact tracing process is initiated at a diagnosed (T) node on the tree, changing the state of the node to X (so that contact tracing will not be re-initiated for the same node), and it follows the $[I \leftarrow T]$ and $[T \leftarrow I]$ pairs to the I neighbors on the tree in both directions.

Note that the state of a node in a pair may be affected by other nodes outside the pair. For example, the infector (I) of the $[T \leftarrow I]$ pair may be contact traced and become T if its infector is I. That is, the dynamics of such pairs depend on the triple-node interactions $[\underline{T \leftarrow I} \leftarrow T]$ (the underline represents the original $[T \leftarrow I]$ pair). The dynamics of such triples depend on four-node interactions, etc. To reduce the complexity of the model, the number of triples is approximated by the pairs. Using network modeling terminology [10], this approximation is called triple closure. For example, the number of $[\underline{T \leftarrow I} \leftarrow T]$ triples is approximated by multiplying the number of $[T \leftarrow I]$ pairs by the fraction of T infectors of the I node, namely,

$$[\underline{T \leftarrow I} \leftarrow T] \approx [T \leftarrow I] \frac{[I \leftarrow T]}{I} . \tag{2.1}$$

The model parameters include θ (the rate of contact tracing of diagnosed individuals), p (the probability of an infectious contact being traced, also called the coverage probability), β (the transmission rate), γ (the recovery rate) and τ (the voluntary testing rate after showing symptoms).

Note that the I node in the $[\underline{T \leftarrow I} \leftarrow T]$ triple has already infected another node (the leftmost infectee T). However, not all I nodes have infected others, such as newly infected nodes that have not yet contacted others. These I nodes cannot be part of such a triple. In the triple closure approximation (2.1), the fraction is calculated using all I nodes, including those who cannot appear in such a triple. This causes the triple closure (2.1) to be inaccurate. This problem affects the approximation of other triples, such as $[\underline{I \leftarrow I} \leftarrow T]$. In this section, we develop an improved contact tracing SIR model that makes the triple closure more accurate.

2.2. The formulation of the improved model

To improve the accuracy of the triple closure (2.1), we divide the infectious compartment (I) into two compartments: those who have not infected other individuals (I_0) and those who have caused further spread (I_+). The population dynamics of the contact tracing SIR model in [2] becomes

$$S' = -\beta S \frac{I_0 + I_+}{N}, \tag{2.2a}$$

$$I_0{}' = \beta S \frac{I_+}{N} - (\gamma + \tau) I_0 - \theta p([I_0 \leftarrow T]) , \tag{2.2b}$$

$$I_+{}' = \beta S \frac{I_0}{N} - (\gamma + \tau) I_+ - \theta p([I_+ \leftarrow T] + [T \leftarrow I_+]) , \tag{2.2c}$$

$$T' = \tau(I_0 + I_+) + \theta p([I_0{\leftarrow}T] + [I_+{\leftarrow}T] + [T{\leftarrow}I_+]) - \theta T\,, \tag{2.2d}$$
$$X' = \theta T\,, \tag{2.2e}$$
$$R' = \gamma(I_0 + I_+)\,. \tag{2.2f}$$

Note that the contact tracing process follows the pairs in these equations, and pairs such as $[T{\leftarrow}I_0]$ do not exist because the I_0 node has not yet infected others and thus cannot be an infector.

To complete the model, we need to derive the dynamics of the pairs $[I_0{\leftarrow}T]$, $[I_+{\leftarrow}T]$, and $[T{\leftarrow}I_+]$. Figure 1 shows the flowchart of pair dynamics.

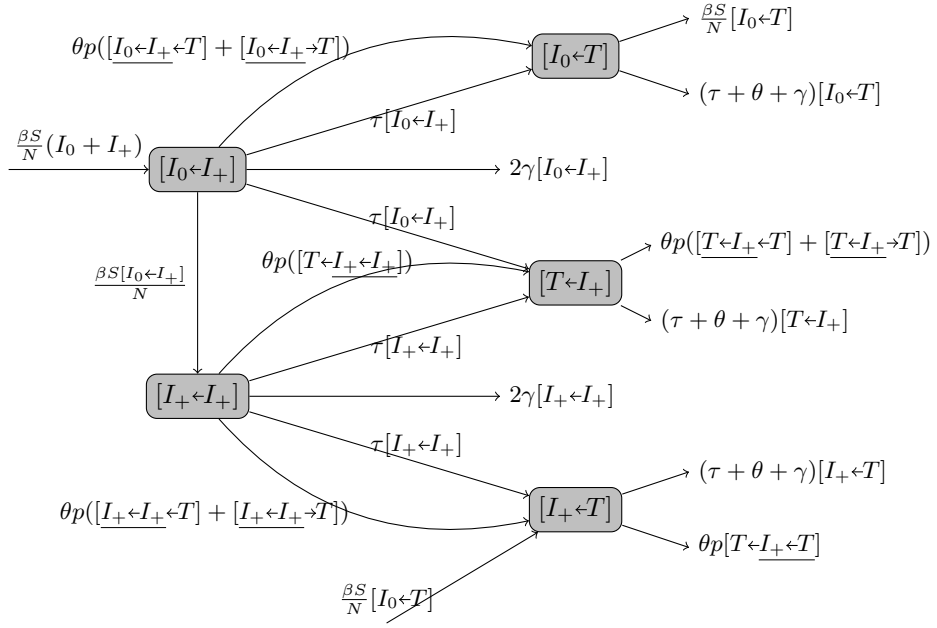

Figure 1. Flowchart of Pair Dynamics for SIR model. The underlines in triples represent the original pairs.

A pair forms when a susceptible is infected by an infectious node, becoming an AII pair. These pairs form at a rate $\beta S(I_0 + I_+)/N$. They leave the state when either node recovers (at a total rate $2\gamma[I_0{\leftarrow}I_+]$), or when either node is voluntarily tested (at a total rate $2\tau[I_0{\leftarrow}I_+]$). The infector I_+ may be contact traced from its infector in a triple $[\underline{I_0{\leftarrow}I_+}{\leftarrow}T]$, or from another diagnosed intectee in the triple $[\underline{I_0{\leftarrow}I_+}{\rightarrow}T]$. Contact tracing occurs on each triple at a rate θp (because the patient is captured by contact tracing with a probability p). In addition, such a pair becomes $[I_+{\leftarrow}I_+]$ when the infectee I_0 infects another node and becomes I_+, which happens at a rate $\beta S[I_0{\leftarrow}I_+]/N$. Thus, the dynamics of the $[I_0{\leftarrow}I_+]$ pairs can be written as

$$\begin{aligned}[I_0{\leftarrow}I_+]' &= \beta\frac{S}{N}(I_0 + I_+) - \beta\frac{S[I_0{\leftarrow}I_+]}{N} - 2(\gamma+\tau)[I_0{\leftarrow}I_+] \\ &\quad -\theta p([\underline{I_0{\leftarrow}I_+}{\rightarrow}T] + [\underline{I_0{\leftarrow}I_+}{\leftarrow}T]) \\ &= \beta\frac{S}{N}(I_0 + I_+) - \beta\frac{S[I_0{\leftarrow}I_+]}{N} - 2(\gamma+\tau)[I_0{\leftarrow}I_+]\end{aligned}$$

$$-\theta p[I_0{\leftarrow}I_+]\left(\frac{[T{\leftarrow}I_+]}{I_+}+\frac{[I_+{\leftarrow}T]}{I_+}\right).$$

For an $[I_+{\leftarrow}I_+]$ pair, either node may recover or be diagnosed. The infectee may be contact traced from one of their secondary infections, which happens in a triple interaction $[T{\leftarrow}\underline{I_+{\leftarrow}I_+}]$. Similarly, the infector may be traced from their infector in triples $[\underline{I_+{\leftarrow}I_+}{\leftarrow}T]$, or from another of their secondary infections in $[\underline{I_+{\leftarrow}I_+}{\rightarrow}T]$. Thus,

$$\begin{aligned}[I_+{\leftarrow}I_+]' &= \beta\frac{S[I_0{\leftarrow}I_+]}{N}-2(\gamma+\tau)[I_+{\leftarrow}I_+]-\theta p([T{\leftarrow}\underline{I_+{\leftarrow}I_+}]+[\underline{I_+{\leftarrow}I_+}{\leftarrow}T]\\ &\quad+[\underline{I_+{\leftarrow}I_+}{\rightarrow}T])\\ &= \beta\frac{S[I_0{\leftarrow}I_+]}{N}-2(\gamma+\tau)[I_+{\leftarrow}I_+]-\theta p[I_+{\leftarrow}I_+]\left(2\frac{[T{\leftarrow}I_+]}{I_+}+\frac{[I_+{\leftarrow}T]}{I_+}\right).\end{aligned}$$

The $[I_+{\leftarrow}T]$ pairs come from $[I_+{\leftarrow}I_+]$ pairs when the infector is diagnosed, or from $[I_0{\leftarrow}T]$ pairs when the infectee I_0 infects others. The pair leaves the state when the infectee I_+ recovers (at a total rate $\gamma[I_+{\leftarrow}I_+]$), or when the infectee is diagnosed due to voluntary testing or contact tracing. Thus,

$$\begin{aligned}[I_+{\leftarrow}T]' &= \tau[I_+{\leftarrow}I_+]+\beta\frac{S[I_0{\leftarrow}T]}{N}+\theta p([\underline{I_+{\leftarrow}I_+}{\rightarrow}T]+[\underline{I_+{\leftarrow}I_+}{\leftarrow}T])\\ &\quad-(\gamma+\tau+\theta)[I_+{\leftarrow}T]-\theta p[T{\leftarrow}\underline{I_+{\leftarrow}T}]\\ &= \tau[I_+{\leftarrow}I_+]+\beta\frac{S[I_0{\leftarrow}T]}{N}+\theta p[I_+{\leftarrow}I_+]\left(\frac{[T{\leftarrow}I_+]}{I_+}+\frac{[I_+{\leftarrow}T]}{I_+}\right)\\ &\quad-(\gamma+\tau+\theta)[I_+{\leftarrow}T]-\theta p[I_+{\leftarrow}T]\frac{[T{\leftarrow}I_+]}{I_+}.\end{aligned}$$

Similarly, the $[I_0{\leftarrow}T]$ pairs come from $[I_0{\leftarrow}I_+]$ when the infector I_+ is diagnosed either by voluntary testing or contact tracing. The pairs leave the state when the infectee I_0 recovers, is diagnosed, or infects others. Thus,

$$\begin{aligned}[I_0{\leftarrow}T]' &= \tau[I_0{\leftarrow}I_+]+\theta p([\underline{I_0{\leftarrow}I_+}{\rightarrow}T]+[\underline{I_0{\leftarrow}I_+}{\leftarrow}T])-(\gamma+\tau+\theta)[I_0{\leftarrow}T]-\beta\frac{S[I_0{\leftarrow}T]}{N}\\ &= \tau[I_0{\leftarrow}I_+]+\theta p[I_0{\leftarrow}I_+]\left(\frac{[T{\leftarrow}I_+]}{I_+}+\frac{[I_+{\leftarrow}T]}{I_+}\right)-(\gamma+\tau+\theta)[I_0{\leftarrow}T]-\beta\frac{S[I_0{\leftarrow}T]}{N}.\end{aligned}$$

Finally, a $[T{\leftarrow}I_+]$ pair comes from an $[I_+{\leftarrow}I_+]$ pair or an $[I_0{\leftarrow}I_+]$ pair when the infectee is diagnosed either by voluntary testing or contact tracing. The pair leaves the state when the infector I_+ recovers or is diagnosed. Thus,

$$\begin{aligned}[T{\leftarrow}I_+]' =&\tau[I_0{\leftarrow}I_+]+\tau[I_+{\leftarrow}I_+]+\theta p[T{\leftarrow}\underline{I_+{\leftarrow}I_+}]-(\gamma+\tau+\theta)[T{\leftarrow}I_+]\\ &-\theta p([\underline{T{\leftarrow}I_+}{\leftarrow}T]+[\underline{T{\leftarrow}I_+}{\rightarrow}T])\\ =&\tau([I_0{\leftarrow}I_+]+[I_+{\leftarrow}I_+])+\theta p[I_+{\leftarrow}I_+]\frac{[T{\leftarrow}I_+]}{I_+}-(\gamma+\tau+\theta)[T{\leftarrow}I_+]\\ &-\theta p[T{\leftarrow}I_+]\left(\frac{[I_+{\leftarrow}T]}{I_+}+\frac{[T{\leftarrow}I_+]}{I_+}\right).\end{aligned}$$

Therefore, the improved model can be written as:

$$S' = -\beta S\frac{I_0+I_+}{N}, \tag{2.3a}$$

$$I_0{}' = \beta S\frac{I_+}{N} - (\gamma+\tau)I_0 - \theta p([I_0{\leftarrow}T]), \tag{2.3b}$$

$$[I_0{\leftarrow}I_+]' = \beta\frac{S}{N}(I_0+I_+) - \beta\frac{S[I_0{\leftarrow}I_+]}{N} - 2(\gamma+\tau)[I_0{\leftarrow}I_+] - \theta p[I_0{\leftarrow}I_+]\frac{[T{\leftarrow}I_+]+[I_+{\leftarrow}T]}{I_+}, \tag{2.3c}$$

$$[I_0{\leftarrow}T]' = \tau[I_0{\leftarrow}I_+] + \theta p[I_0{\leftarrow}I_+]\frac{[T{\leftarrow}I_+]+[I_+{\leftarrow}T]}{I_+} - (\gamma+\tau+\theta)[I_0{\leftarrow}T] - \beta\frac{S[I_0{\leftarrow}T]}{N}, \tag{2.3d}$$

$$I_+{}' = \beta S\frac{I_0}{N} - (\gamma+\tau)I_+ - \theta p([I_+{\leftarrow}T]+[T{\leftarrow}I_+]), \tag{2.3e}$$

$$[I_+{\leftarrow}I_+]' = \beta\frac{S[I_0{\leftarrow}I_+]}{N} - 2(\gamma+\tau)[I_+{\leftarrow}I_+] - \theta p[I_+{\leftarrow}I_+]\frac{2[T{\leftarrow}I_+]+[I_+{\leftarrow}T]}{I_+}, \tag{2.3f}$$

$$[I_+{\leftarrow}T]' = \tau[I_+{\leftarrow}I_+] + \beta\frac{S[I_0{\leftarrow}T]}{N} + \theta p[I_+{\leftarrow}I_+]\frac{[I_+{\leftarrow}T]+[T{\leftarrow}I_+]}{I_+} - (\gamma+\tau+\theta)[I_+{\leftarrow}T] - \theta p[T{\leftarrow}I_+]\frac{[I_+{\leftarrow}T]}{I_+}, \tag{2.3g}$$

$$[T{\leftarrow}I_+]' = \tau([I_0{\leftarrow}I_+]+[I_+{\leftarrow}I_+]) + \theta p[I_+{\leftarrow}I_+]\frac{[T{\leftarrow}I_+]}{I_+} - (\gamma+\tau+\theta)[T{\leftarrow}I_+] - \theta p[T{\leftarrow}I_+]\frac{[I_+{\leftarrow}T]+[T{\leftarrow}I_+]}{I_+}, \tag{2.3h}$$

$$T' = \tau(I_0+I_+) + \theta p([I_0{\leftarrow}T]+[I_+{\leftarrow}T]+[T{\leftarrow}I_+]) - \theta T, \tag{2.3i}$$

$$X' = \theta T, \tag{2.3j}$$

$$R' = \gamma(I_0+I_+). \tag{2.3k}$$

Note that, when no infection has occurred (i.e. $I_+ = 0$, e.g., when the disease is not present), the model is not well-posed because the fractions in this model are not defined. In this case, the triples that give rise to these fractions do not exist, and so these fractions can be set to 0 to eliminate the triples. Thus, if $I_+ = 0$, the system becomes

$$S' = -\beta S\frac{I_0+I_+}{N}, \tag{2.4a}$$

$$I_0{}' = \beta S\frac{I_+}{N} - (\gamma+\tau)I_0 - \theta p([I_0{\leftarrow}T]), \tag{2.4b}$$

$$[I_0{\leftarrow}I_+]' = \beta\frac{S}{N}(I_0+I_+) - \beta\frac{S[I_0{\leftarrow}I_+]}{N} - 2(\gamma+\tau)[I_0{\leftarrow}I_+], \tag{2.4c}$$

$$[I_0{\leftarrow}T]' = \tau[I_0{\leftarrow}I_+] - (\gamma+\tau+\theta)[I_0{\leftarrow}T] - \beta\frac{S[I_0{\leftarrow}T]}{N}, \tag{2.4d}$$

$$I_+{}' = \beta S\frac{I_0}{N} - (\gamma+\tau)I_+ - \theta p([I_+{\leftarrow}T]+[T{\leftarrow}I_+]), \tag{2.4e}$$

$$[I_+{\leftarrow}I_+]' = \beta\frac{S[I_0{\leftarrow}I_+]}{N} - 2(\gamma+\tau)[I_+{\leftarrow}I_+], \tag{2.4f}$$

$$[I_+{\leftarrow}T]' = \tau[I_+{\leftarrow}I_+] + \beta\frac{S[I_0{\leftarrow}T]}{N} - (\gamma+\tau+\theta)[I_+{\leftarrow}T], \tag{2.4g}$$

$$[T{\leftarrow}I_+]' = \tau([I_0{\leftarrow}I_+]+[I_+{\leftarrow}I_+]) - (\gamma+\tau+\theta)[T{\leftarrow}I_+], \tag{2.4h}$$

$$T' = \tau(I_0 + I_+) + \theta p([I_0 \leftarrow T] + [I_+ \leftarrow T] + [T \leftarrow I_+]) - \theta T, \tag{2.4i}$$
$$X' = \theta T\,, \tag{2.4j}$$
$$R' = \gamma(I_0 + I_+)\,. \tag{2.4k}$$

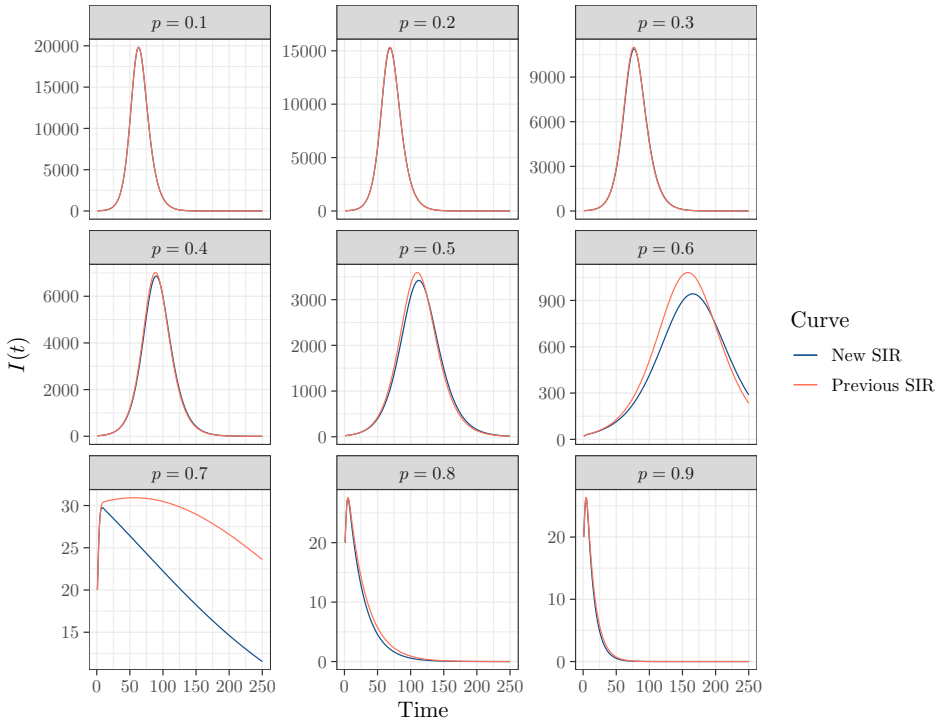

Figure 2. The comparison of $I(t)$ numerically solved from our contact tracing model with the previous SIR model. The parameter values are $\beta = 0.4$, $\theta = 1$, $\gamma = 0.1$, $\tau = 0.15$, $N = 300000$, and the coverage probability $p = 0.1, 0.2, ..., 0.9$. The initial conditions are $S(0) = N$, $I_0(0) = 20$ and $I_+(0) = 0$ for the improved model while $I(0) = 20$ for the original model, and 0 for all other variables.

The comparison of $I(t)$ numerically solved from our improved SIR contact tracing model and the original SIR model in [2] is shown in Figure 2, for several values of the coverage probability p. This figure shows that, when the value of p is small, the difference between the two models is very small, as contact tracing is rare in this case. As p gradually increases and the impact of contact tracing becomes larger, the difference between the two models gradually becomes apparent. However, when p is large, contact tracing becomes very effective, and thus the final size is small. This means that the number of contact traced individuals is also small, resulting in a small difference between the two models.

3. Model analysis

In this section, we calculate the control reproduction number $\mathcal{R}_C$ of our new SIR model and analyze the dependence of $\mathcal{R}_C$ and the final epidemic size on the model parameters through numerical simulation.

At the disease-free equilibrium where $S = N$ and all other variables are 0, the model (2.3) becomes (2.4). However, we cannot study the stability of (2.4) to determine the disease threshold. This is because the threshold determines the behavior when the system is close to, but not at, the disease-free equilibrium. In this situation, (2.3) cannot be reduced to (2.4). We introduce the following variable change to avoid this problem. Let

$$x = \frac{[I_0{\leftarrow}I_+]}{I_+}, \quad y = \frac{[I_0{\leftarrow}T]}{I_+}, \quad z = \frac{I_0}{I_+}, \quad u = \frac{[I_+{\leftarrow}I_+]}{I_+}, \quad v = \frac{[I_+{\leftarrow}T]}{I_+}, \quad w = \frac{[T{\leftarrow}I_+]}{I_+}.$$

Then system (2.3) can be rewritten as

$$S' = -\beta S\frac{I_0 + I_+}{N}, \tag{3.1a}$$

$$I_0{}' = \beta S\frac{I_+}{N} - (\gamma + \tau)I_0 - \theta p\frac{y}{z}I_0, \tag{3.1b}$$

$$I_+{}' = \beta S\frac{I_0}{N} - (\gamma + \tau)I_+ - \theta p(v + w)I_+, \tag{3.1c}$$

$$x' = \frac{\beta S}{N}(1 + z - x - xz) - (\gamma + \tau)x, \tag{3.1d}$$

$$y' = -\beta\frac{S}{N}(y + yz) - \theta y + \tau x + \theta p(v + w)(x + y), \tag{3.1e}$$

$$z' = \beta\frac{S}{N}(1 - z^2) + \theta pz(v + w) - \theta py, \tag{3.1f}$$

$$u' = \beta\frac{S}{N}(x - uz) - (\gamma + \tau)u - \theta pwu, \tag{3.1g}$$

$$v' = \beta\frac{S}{N}(y - vz) + \theta p(uv + uw + v^2) - \theta v + \tau u, \tag{3.1h}$$

$$w' = -\beta\frac{S}{N}wz + \tau(x + u) + \theta pwu - \theta w. \tag{3.1i}$$

3.1. Well-posedness

We begin by establishing a subset of the state space in the first quadrant, denoted as Ω, which is biologically meaningful. First, all variable values must be nonnegative since they represent counts. Secondly, $[I_0{\leftarrow}I_+] + [I_0{\leftarrow}T] \le I_0$ and $[I_+{\leftarrow}I_+] + [I_+{\leftarrow}T] \le I_+$. Intuitively, the first inequality holds because each patient has only one infector. Therefore, each I_0 corresponds to a unique pair where I_0 is the infectee, and the infector could be I_+, T, X, or R. In other words, the pairs $[I_0{\leftarrow}I_+] + [I_0{\leftarrow}T]$ account for only a fraction of I_0. Similarly, the second inequality must also be satisfied. Thus, the biologically meaningful states must satisfy

$$0 \le S, I_0, I_+, T, X, R, [I_0{\leftarrow}I_+], [I_0{\leftarrow}T], [I_+{\leftarrow}I_+], [I_+{\leftarrow}T], [T{\leftarrow}I_+], \tag{3.2a}$$

$$I_0 \ge [I_0{\leftarrow}I_+] + [I_0{\leftarrow}T], \tag{3.2b}$$

$$I_+ \ge [I_+{\leftarrow}I_+] + [I_+{\leftarrow}T]. \tag{3.2c}$$

Under the new variables, the inequalities (3.2) that define the set of biologically meaningful states Ω become

$$0 \le S, I_0, I_+, x, y, z, u, v, w, \tag{3.3a}$$

$$z \geq x + y, \tag{3.3b}$$
$$1 \geq u + v. \tag{3.3c}$$

This set is denoted as $\tilde{\Omega}$. The following theorem guarantees that the system (3.1) is well-posed.

Theorem 3.1. *The set of biologically meaningful states $\tilde{\Omega}$ is a positive invariant set of the system* (3.1).

Proof. We will show that $\tilde{\Omega}$ is positively invariant. To see this, first note that $S'|_{S=0} = 0$, that is, the hyperplane $\{S = 0\}$ is invariant. In addition, for any solution starting in $\tilde{\Omega}$, ${I_0}'|_{I_0=0} \geq 0$, which also holds for all other variables except for z, as shown at the end of the proof. First, we define

$$\xi = z - x - y.$$

Hence, from (3.1d-f)

$$\xi' = -\beta\frac{S}{N}(1+z)\xi + \theta p(v+w)\xi + \theta y(1-p) + \gamma x.$$

Therefore, for a solution starting in $\tilde{\Omega}$,

$$\xi'|_{\xi=0} = \theta y(1-p) + \gamma x \geq 0,$$

which means the solutions enter $\tilde{\Omega}$ on the boundary $\{\xi = 0\}$.

Similarly, we can show that $u + v \leq 1$. Let

$$\eta = 1 - u - v.$$

From (3.1g-h)

$$\begin{aligned}\eta' &= 1 - \beta\frac{S}{N}(x + y - uz - vz) + \gamma u - \theta p(uv + v^2) + \theta v \\ &\geq 1 - \beta\frac{S}{N}z(1-u-v) + \gamma u + \theta pv(1-u-v).\end{aligned}$$

Hence,

$$\eta'|_{\eta=0} = 1 + \gamma u \geq 0.$$

This implies that solutions enter $\tilde{\Omega}$ on the boundary $\eta = 0$.

Finally, note that

$$\begin{aligned}z' &= \beta\frac{S}{N}(1-z^2) + \theta pz(v+w) - \theta py \\ &\geq \beta\frac{S}{N}(1-z^2) + \theta pz(v+w) - \theta pz.\end{aligned}$$

For a solution starting in $\tilde{\Omega}$,

$$z'|_{z=0} = \beta\frac{S}{N} \geq 0,$$

which means that the solutions enter $\tilde{\Omega}$ on the boundary $\{z = 0\}$. □

Since system (3.1) is equivalent to system (2.3), and the set $\tilde{\Omega}$ is equivalent to Ω, we can establish the following parallel result for (2.3):

Corollary 3.1. *The set Ω is the positive invariant set of the system* (2.3).

3.2. Disease-free equilibrium and control reproduction number

At the disease-free equilibrium point, (x, y, z, u, v, w) in system (3.1) is independent of S and I, and their dynamics are governed by

$$x' = \beta(1 + z - x - xz) - (\gamma + \tau)x, \tag{3.4a}$$
$$y' = -\beta(y + yz) - \theta y + \tau x + \theta p(v + w)(x + y), \tag{3.4b}$$
$$z' = \beta(1 - z^2) + \theta pz(v + w) - \theta py, \tag{3.4c}$$
$$u' = \beta(x - uz) - (\gamma + \tau)u - \theta pwu, \tag{3.4d}$$
$$v' = \beta(y - vz) + \theta p(uv + uw + v^2) - \theta v + \tau u, \tag{3.4e}$$
$$w' = -\beta wz + \tau(x + u) + \theta pwu - \theta w. \tag{3.4f}$$

Conjecture 3.1. *System* (3.4) *has only one biologically relevant positive equilibrium point, satisfying (3.3b) and (3.3c).*

Despite the difficulty in analytically solving the equilibrium of (3.4), we use numerical analysis to find the positive equilibria via Newton's method. We randomly generate 100 sets of positive initial guesses for Newton's method. These initial guesses yield two positive equilibria, but only one equilibrium point lies in the biologically meaningful set $\tilde{\Omega}$ given in (3.3). This strongly suggests that there is a unique biologically relevant positive equilibrium. This corresponds to the disease-free equilibrium of (3.4), denoted as $E_0 = (N, 0, 0, x^*, y^*, z^*, u^*, v^*, w^*)$.

Theorem 3.2. *If the conjecture 3.1 holds, then the control reproduction number of system* (3.1) *is given by*

$$\mathcal{R}_C = \frac{\beta}{\sqrt{(\gamma + \tau + \theta py^*/z^*)[\gamma + \tau + \theta p(v^* + w^*)]}}. \tag{3.5}$$

Proof. After decoupling the equations for I_0 and I_+ from the linearization of (3.4) about the disease-free equilibrium E_0, we use the next generation matrix method [17] to calculate the control reproduction number. The Jacobian matrix of the linearized system for I_0 and I_+ is

$$J = \begin{pmatrix} -(\gamma + \tau) - \theta py^*/z^* & \beta \\ \beta & -(\gamma + \tau) - \theta p(v^* + w^*) \end{pmatrix}.$$

Rewrite the system as

$$\frac{d}{dt}\begin{pmatrix} I_0 \\ I_+ \end{pmatrix} = (F - V)\begin{pmatrix} I_0 \\ I_+ \end{pmatrix},$$

where the new infection matrix F and the transition matrix V are

$$F = \begin{pmatrix} 0 & \beta \\ \beta & 0 \end{pmatrix}, V = \begin{pmatrix} (\gamma + \tau) + \theta py^*/z^* & 0 \\ 0 & (\gamma + \tau) + \theta p(v^* + w^*) \end{pmatrix}.$$

The next generation matrix is obtained as

$$FV^{-1} = \begin{pmatrix} 0 & \frac{\beta}{(\gamma+\tau)+\theta p(v^*+w^*)} \\ \frac{\beta}{(\gamma+\tau)+\theta p y^*/z^*} & 0 \end{pmatrix}.$$

The characteristic polynomial of the next generation matrix is

$$|\lambda E - FV^{-1}| = \lambda^2 - \frac{\beta^2}{(\gamma+\tau+\theta p y^*/z^*)[\gamma+\tau+\theta p(v^*+w^*)]}.$$

This gives the control reproduction number as the largest eigenvalue

$$\mathcal{R}_C = \frac{\beta}{\sqrt{(\gamma+\tau+\theta p y^*/z^*)[\gamma+\tau+\theta p(v^*+w^*)]}}.$$

□

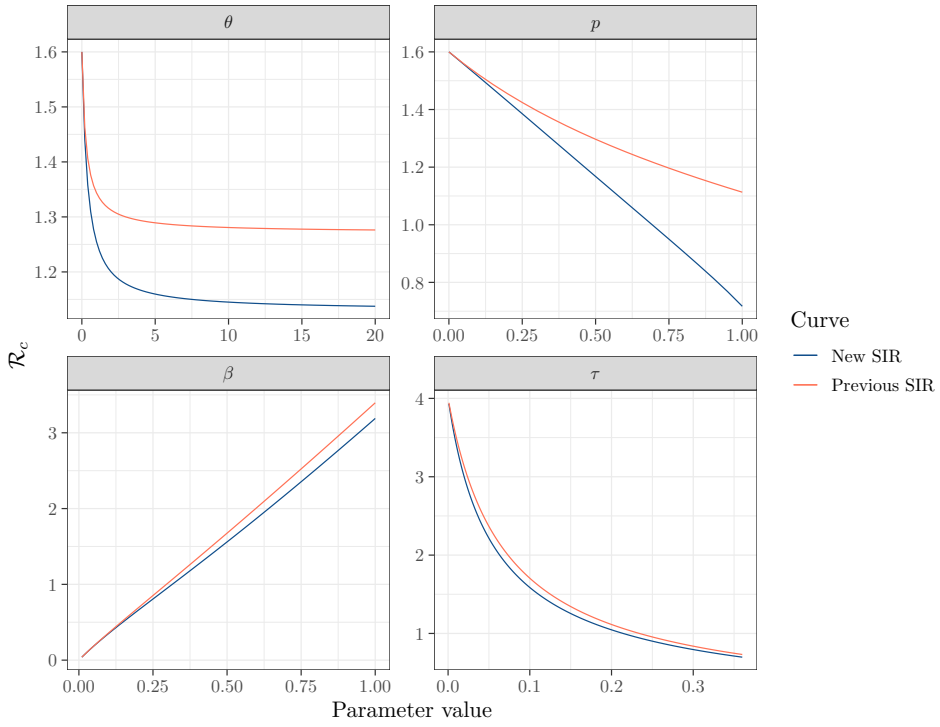

Figure 3. The dependence of the control reproduction number $\mathcal{R}_C$ on parameters θ, p, β and τ. We set $\beta = 0.4$, $\theta = 1$, $p = 0.4$, $\gamma = 0.1$ and $\tau = 0.15$. Here, we vary the corresponding parameter values and fix the others.

3.3. Dependency of $\mathcal{R}_C$ on model parameters

In this subsection, we study the dependence of $\mathcal{R}_C$ and the final epidemic size on model parameters through numerical simulations.

Figure 3 shows that for both models, $\mathcal{R}_C$ is an increasing function of β and a decreasing function of θ, p, τ. The figure shows that there is no significant difference

in $\mathcal{R}_C$ between the two models by changing the values of β and τ. However, the original model (2.3) significantly underestimates the effect of contact tracing on $\mathcal{R}_C$ for large coverage probability p or large contact tracing rate θ.

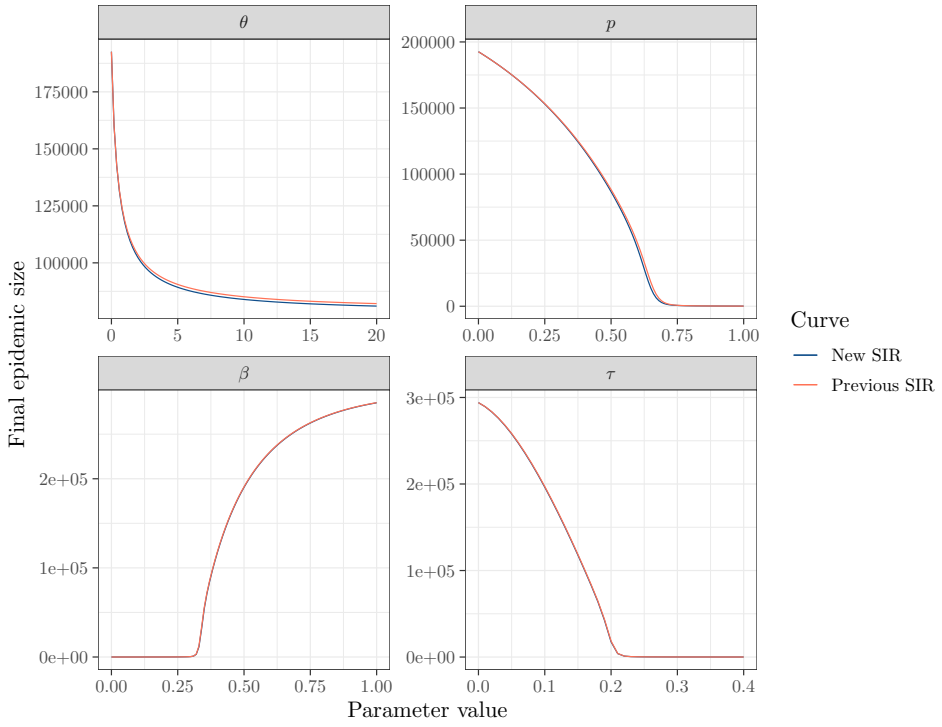

Figure 4. The dependence of the final epidemic size on parameters. We use the same parameter values as Figure 3.

Figure 4 shows the dependence of the final epidemic size $(S(\infty)-S(0))$ on model parameters. As contact tracing is strengthened (with a larger contact tracing rate θ or coverage probability p), the final size decreases. In addition, the final size is an increasing function of the transmission rate β and a decreasing function of the voluntary testing rate τ. The difference in the final size between the two models with the same parameter values is negligible, even for large values of θ and p.

4. Conclusion

We establish a more accurate contact tracing model to study the transmission dynamics of the disease based on [2] by considering a more realistic triple closure approximation. We extend this model to include two infectious compartments: I_0 (have not infected others) and I_+ (have infected other individuals). This prevents us from counting I_0 in some related triples and also makes our model more accurate after the triple approximation. By comparing the solutions of the two models numerically, we find the difference between the solutions of the two models is small when the tracing coverage probability p is small. This is because, for a small p,

the difference between these two models is based on a small number contact traced individuals. As p increases, the difference increases to maximum, then starts to decrease as p becomes even larger. This is because, for a large p, contact tracing may become very effective in controlling the epidemic, so the total number of contact traces becomes small again.

By analyzing the dependence of the control reproduction number $\mathcal{R}_C$ on model parameters, we found that the control reproduction number $\mathcal{R}_C$ of both models increases with the transmission rate β and decreases with the voluntary testing rate τ similarly. Additionally, $\mathcal{R}_C$ is a decreasing function of θ (contact tracing rate) and p (tracing coverage) as they increase. That is, increasing the strength of contact tracing reduces disease transmission, which is an intuitive result. However, $\mathcal{R}_C$ of the improved model decreases faster with the parameters than the original model, i.e., the Bednarski et al. model significantly underestimates the effect of contact tracing on disease transmission.

The final epidemic size for both models is similar and shows a similar dependence on model parameters. Therefore, when contact tracing measures are not strict or when the epidemic size is the quantity of interest, the Bednarski et al. model can still be used. On the other hand, the improved model should be used to study the effect of contact tracing, as it is more precise and provides a more accurate value of the controlled reproduction number.

The improve model ignores the latent period and asymptomatic infection. However, it can be extended to incorporate exposed and asymptomatic compartments, establishing a more realistic contact tracing model. To further improve this model, we need to accurately model the tree of transmission formed by the disease dynamics, and incorporate its degree distribution using a similar triple closure method as in [18]. However, it will be mathematically challenging because the resulting model may be infinite dimensional.

Acknowledgements

We sincerely thank the anonymous reviewers and editors for their thorough review and helpful suggestions.

References

[1] C. Adlhoch, A. Baka, O. Cenciarelli et al., *Contact tracing: Public health management of persons, including healthcare workers, having had contact with covid-19 cases in the european union*, European Centre for Disease Prevention and Control, 2020.

[2] S. Bednarski, L. L. Cowen, J. Ma et al., *A contact tracing sir model for randomly mixed populations*, Journal of Biological Dynamics, 2022, 16(1), 859–879.

[3] M. Bushman, R. Kahn, B. P. Taylor et al., *Population impact of sars-cov-2 variants with enhanced transmissibility and/or partial immune escape*, Cell, 2021, 184(26), 6229–6242.

[4] J. Chan, D. Foster, S. Gollakota et al., *Pact: Privacy sensitive protocols and mechanisms for mobile contact tracing*, arXiv preprint arXiv:2004.03544, 2020.

[5] C. Fraser, S. Riley, R. M. Anderson and N. M. Ferguson, *Factors that make an infectious disease outbreak controllable*, Proceedings of the National Academy of Sciences, 2004, 101(16), 6146–6151.

[6] T. Hale, N. Angrist, R. Goldszmidt et al., *A global panel database of pandemic policies (oxford covid-19 government response tracker)*, Nature Human Behaviour, 2021, 5(4), 529–538.

[7] J. Hellewell, S. Abbott, A. Gimma et al., *Feasibility of controlling covid-19 outbreaks by isolation of cases and contacts*, The Lancet Global Health, 2020, 8(4), e488–e496.

[8] R. Huerta and L. S. Tsimring, *Contact tracing and epidemics control in social networks*, Physical Review E, 2002, 66(5), 056115.

[9] S. Jamaludin, N. A. Azmir, A. F. M. Ayob and N. Zainal, *Covid-19 exit strategy: Transitioning towards a new normal*, Annals of Medicine and Surgery, 2020, 59, 165–170.

[10] M. J. Keeling, *The effects of local spatial structure on epidemiological invasions*, Proceedings of the Royal Society of London. Series B: Biological Sciences, 1999, 266(1421), 859–867.

[11] M. J. Keeling, T. D. Hollingsworth and J. M. Read, *Efficacy of contact tracing for the containment of the 2019 novel coronavirus (covid-19)*, J Epidemiol Community Health, 2020, 74(10), 861–866.

[12] M. U. Kraemer, C.-H. Yang, B. Gutierrez et al., *The effect of human mobility and control measures on the covid-19 epidemic in china*, Science, 2020, 368(6490), 493–497.

[13] A. P. Lemos-Paiao, C. J. Silva and D. F. Torres, *A new compartmental epidemiological model for covid-19 with a case study of portugal*, Ecological Complexity, 2020, 44, 100885.

[14] J. Levesque, D. W. Maybury and R. D. Shaw, *A model of covid-19 propagation based on a gamma subordinated negative binomial branching process*, Journal of Theoretical Biology, 2021, 512, 110536.

[15] M. Li, J. Cui, J. Zhang et al., *Transmission characteristic and dynamic analysis of covid-19 on contact network with tianjin city in china*, Physica A: Statistical Mechanics and its Applications, 2022, 608, 128246.

[16] J. C. Miller, *Percolation and epidemics in random clustered networks*, Physical Review E, 2009, 80(2), 020901.

[17] P. van den Driessche and J. Watmough, *Reproduction numbers and subthreshold endemic equilibria for compartmental models of disease transmission*, Mathematical Biosciences, 2002, 180(1-2), 29–48.

[18] E. Volz, *Sir dynamics in random networks with heterogeneous connectivity*, Journal of Mathematical Biology, 2008, 56, 293–310.