Research Paper

The characteristics of COVID-19 transmission from case to high-risk contact, a statistical analysis from contact tracing data

Chayanon Phucharoen\textsuperscript{a,b,}\textsuperscript{*}, Nichapat Sangkaew\textsuperscript{a,b}, Kristina Stosic\textsuperscript{a,}\textsuperscript{3}

\textsuperscript{a} Faculty of Hospitality and Tourism, Prince of Songkla University (Phuket Campus), Phuket, Thailand
\textsuperscript{b} COVID-19 Joint Force Operation Centre (Phuket Province), Phuket, Thailand

ARTICLE INFO

Article History:
Received 17 May 2020
Revised 24 August 2020
Accepted 25 August 2020
Available online 21 September 2020

Keywords:
COVID-19
Household transmission
Quarantine
Probability of infection

ABSTRACT

Background: As of 24th of August 2020, the number of global COVID-19 confirmed cases is nearly 24 million. In the same period, the number of recorded infections in Thailand has remained at approximately 3300. This paper explores the specifics of COVID-19 or SARS-CoV-2 transmissions in Phuket, Thailand’s second most visited tourist destination.

Methods: High-risk contacts recorded by Phuket Provincial Public Health Office were analysed using the Probit model to investigate the risk factors for transmission from confirmed COVID-19 cases to their high-risk contacts. The analysis was further focused on the impact of quarantine measures in state provided facilities on contacts’ probability of infection.

Findings: 15.6% of 1108 high-risk contacts were found to be infected, and they accounted for 80% of 214 confirmed cases in Phuket till 29th April 2020. Moreover, 10.68% of all high-risk contacts were confirmed to be infected before the quarantine, and 4.55% after the policy was enforced. In addition, a contact who lived within the same household with a confirmed case was 25% more exposed to infection when compared to a contact who did not share a household.

Interpretation: Results confirmed that the quarantine policy, which mandated individual isolation in the state provided facilities for all high-risk contacts, diminished contact’s chance of infection from the confirmed cases, especially in the epicenter districts. Our findings confirmed that sharing accommodation with an infected case, and exposure to a case with several documented secondary transmission, generally increased the SARS-CoV-2 infection probability. Finally, some confirmed cases do exhibit a higher risk of spreading SARS-CoV-2 to their contacts compared to a typical confirmed case. Further studies of high reproduction groups of infected patients are recommended.

Funding: No funding was received for this research.

© 2020 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

1. Introduction

Thailand’s second most visited tourist destination [19], Phuket island, had its first recorded cases of COVID-19 infection in late January 2020 [21]. Despite the travel ban, posted on March 26th, 2020 [22], the number of infected cases kept increasing within the island till the end of April. The highest increase in the new infections was recorded between the 23rd of March 2020 and 10th of April 2020, with a 56% of total infections occurring during that period (as of April 29th, 2020).

Several public orders and policies were launched by the Phuket Provincial Public Health Office to control the infection spread on the island. One of the key measures was a mandatory quarantine in state provided facilities for all persons who were in high-risk contact (HRC) with an infected individual, within 24 h of an infection confirmation. The difference between the state mandated quarantine and the home quarantine was that the state mandated quarantine requires all HRC to stay in a repurposed hotel room for 14 days since an exposure to confirmed case. This measure was fully implemented after the 9th of April, since the state provided facilities were not available prior to that date. Some of the HRCs in two epicenter districts had undertaken home quarantine prior to the 9th of April. By the end of April, the number of confirmed infections was 217, and not 677, as previously predicted by Thailand’s Public Health Authorities [20]. However, a disproportionate number of infections amongst Phuket’s districts, and a concentration of cases in two fairly different population districts, Patong and Bangtao, raised questions about the risk factors of transmission.
Research in context

Evidence before this study

Exponential global spread of SAR-CoV-2, or COVID-19 virus started at the end of January 2020, and WHO had reported that Thailand was one of the first countries with the confirmed cases of the disease. Surveillance data from Phuket Provincial Public Health Office contact tracing report (from January till 9th May 2020) clearly revealed that most of the infected cases in Phuket were classified as transmissions amongst the local residents, although Phuket was the main travel destination in Thailand for both Chinese and European travellers during the first quarter of 2020. The empirical findings in China recently revealed the severity of household secondary attack of COVID-19. Moreover, it was found in the US that younger contacts are less vulnerable to COVID-19 infection when compared to the adult contacts. In addition, there is evidence in Iran and Malaysia that male contacts of the confirmed cases are in the group of higher infection risks.

Added value of this study

As COVID-19 is still considered a concurrent pandemic, most of the previous findings employed descriptive analysis. This paper explores the risk factors for SARS-CoV-2 transmission in Phuket through regression analysis, which enables the analyses of the effects of each risk factor for transmission, while keeping the other factors unchanged. Moreover, this study investigates whether quarantine of High-Risk Contacts in state provided facilities could mitigate the number of COVID-19 confirmed cases in Phuket. The analyses confirmed that the main risk factor for COVID-19 transmission from confirmed cases to contacts was sharing a household, as the infection probability of contacts could increase by 25% from this factor alone, and this marginal rate of infection went up to 94% in the disease epicentres. Moreover, results also suggest that a contact exposed to a confirmed case with a high record of documented secondary transmissions had higher infection probability than a contact exposed to a typical confirmed case. The value of these findings is that they can help focus the future studies on factors that seem to carry the highest risks for transmission.

Implications of all the available evidence

The results of this study confirmed that a quarantine measures for all High-Risk Contact in state provided facilities statistically diminished contact’s chance of infection and could therefore guide the future states’ regulations for protection from the COVID-19 widespread transmissions. As the severity of the household transmissions was also confirmed, prompt identification of infected cases should be further added to the state infection spread protective measures. Moreover, a closer look at domestic behaviours could possibly shed the light on the transmission mechanisms within a household in further studies; e.g. behavioural studies and pathology. Finally, a study of confirmed cases with high numbers of documented secondary transmissions is recommended from the results of this investigation, in order to minimise the effects of this confirmed strong transmission factor.

Phuket Provincial Public Health Office’s rationale for HRCs quarantine in state provided facilities was supported by recent discoveries in China by Li et al. [16], who found that household secondary infections were as high as 16.3%. Moreover, many areas of China also recorded most secondary COVID-19 infections within families [23]. Furthermore, Chen et al. [5] and Cauchemez et al. [4] also found that the virus spreads more easily among people living in the same household [5,12]. In the US, there was a clear report of family cluster transmission in Westchester, and evidence of transmission of virus among married couples in California and Illinois [9].

Apart from literature identified factor of high household transmissions, Urton [24] had mathematically simulated the social distancing measures and their outcomes, and found that a friend visiting another friend would reconnect 71% of previously isolated household. Therefore, this factor was included in this paper’s analysis of risks for transmission.

Researchers also noted differences in number of infections according to gender. In Switzerland, Spain, and Italy, a larger portion of COVID-19 reported cases were female, while in Iran they were male. Moreover, Biswas et al. [2] found that in Iran and Malaysia, male contacts had 23% higher risk of infection, when compared to female contacts who were exposed to confirmed cases. However, Medicalnewstoday [18] denoted that it was relatively hard to conclude the infection rate across genders through an observation of infected proportions alone. Based on these findings, the gender variable is one of the studied parameters in this paper.

Another variable examined in this study is infected cases, and their high-risk contacts age. For example, Bi et al. [1] revealed an individual’s infection risk was not statistically different across age groups. However, an investigation by Li et al. [16], as well as Khan’s study [15] found that the secondary infection rate in children occurred in lower numbers than adults. With regression testing of this study’s data, it could be verified whether the age risk factors of transmission are statistically valid, given that other factors remain constant.

Through in-depth investigation of the infected cases by Phuket Provincial Public Health Office [21], and subsequent statistical analyses of the contact tracing data for all high-risk contacts, the links between the confirmed cases and their high-risk contacts, as well as the effects of the state provided quarantine for all high-risk contacts (HRCs) on Phuket COVID-19 situation were observed. As Liu et al. [17], and Colburn [6] noted, the understanding of the COVID-19 risk factors of transmission could not only assist in design of public health policies, but could also help the authorities in outlining re-entry strategy for tourism-led economy. Therefore, the aim of this paper is to verify the risk factors of transmission between the confirmed cases and their high-risk contacts, as well as to investigate the effectiveness of the state imposed quarantine of high-risk contacts in decreasing the number of Phuket COVID-19 infections.

2. Methods

The following diagram (Fig. 1) reveals the investigation flow-chart which was undertaken by Phuket Provincial Public Health office. Once the person is confirmed by the Polymerase Chain Reaction (PCR) test as a COVID-19 confirmed case, an immediate in-depth telephone interview with the patient is conducted. This is done to obtain the extensive list of people with whom the patient might have been in physical contact with during the past two weeks, and for periods longer than five minutes without wearing a facial mask. With that information, a team from Phuket’s Department of Disease Control classified the identified contacts as either high-risk contacts (HRC), or low-risk contact (LRC), and further documented all contacts’ demographic information, and type of relationship with the confirmed case(s).

It should be noted that only a few contacts in Phuket were identified as low-risk contacts (LRC). This is because Phuket Public Health Authority classified all contacts who spent longer than five minutes exposed to a confirmed case without a facial mask, or being within physical distance of one meter, as the high-risk contacts (HRC). LRCs
were persons who were exposed to confirmed cases either with 2 meter physical distance, less than 2 meter physical distance with a facial mask, or exposed to the confirmed case for no longer than 5 minutes. Most of the recorded LRCs were convenience store cashiers, food vendors and people who interacted with a confirmed case for a short period of time. HRCs’ information were collected in a spreadsheet, with slight corrections for spelling and date anomalies.

As it can be noted in Fig. 2, most infections were recorded in two districts. The first one is the major tourist entertainment district of Phuket, called Patong, with a total of 86 confirmed cases. The second district is a residential area with a majority Muslim population, called Bangtao, and 61 confirmed cases. These two districts were marked with the incidence rate of 0.57% and 0.34%, per number of residents in each district respectively. Other popular tourist and residential
C. Phucharoen et al. / EClinicalMedicine 27 (2020) 100543

The analyses of data were conducted through an application of a basic t-statistic, in order to test for mean difference on each variable of interest, as per Eq. (1). This was done between the group of infected HRCs and non-infected HRCs, as well as through use of mean difference on infection probability before and after quarantine measures. However, the key testing platform was the next explained Probit regression, which allowed evaluation of the effects of each factor, while holding other factor constant. Since the dependent variable in this study is a dichotomous variable, the two choices for this study were Logit and Probit models. Hahn et al. [11] pointed out that the Probit equation is preferable if the testing model does not contain independent variables with extreme value. Moreover, since the focus of this study is finding the marginal effect from each risk factor of transmission, both Logit and Probit models would relatively yield similar marginal effect [3, 7]. However, since the exogenous variables in this study were mainly dummy and numerical variables that do not contain extreme values, Probit model was selected for the analysis. The result tables in the next section report marginal effect and p-value of each variable, while variable coefficients are reported in the appendix tables.

$$\text{Detected}_i = \beta_x + \beta_y \text{Friend}_i + \beta_z \text{Household}_i + \beta_y \text{Sexdifferential}_i + \beta_x \text{Num. Case Exposure}_i + \beta_y \text{Case}_i \text{Age} + \beta_x \text{Case}_i \text{Spreading ratio}_i + \beta_y C \gamma_i + \epsilon_i \quad \text{Equation 1}$$

The endogenous variable in this model (Eq. (1)) is whether the high-risk contact i was detected with COVID-19. 1 if detected, 0 not-detected. If contacts were reported positive from the PCR tests, either during or before they completed 14 days of quarantine, they would be regarded as 1 detected (infected). The subscript i represents case x who contact i was exposed to. There were 172 confirmed cases, identified from the 1108 high-risk contact reported during the observation period of this study. Beside the self-explanatory name of variables in Eq. (1), Table 1 explains the measurements of each exogenous variable, its expected sign.

Through the current pandemic, there was a relatively limited amount of empiric literature about the risk factors for transmission between the confirmed cases and their contacts. Most of the findings were devoted to understanding of the SARS-CoV-2’s epidemiology, pinpointing the COVID-19 infection determining factors. These literatures identified factors were hence included in the above regression. In addition, variables included in our model were selected based on evidence already reported in the literature section of the introduction part. For example, contact’s age, admitted condition of an infected case, and a number of days of illness before confirmed case was admitted to a hospital, were added in the vector of $C \gamma_i$, and the results with these added factors are observed in the extended model of this paper.

Table 1
Explanation of key variables used in Equation 1

| Variable                           | Measurement                                                                 | Expected sign | Exogenous Variable |
|------------------------------------|----------------------------------------------------------------------------|---------------|--------------------|
| Case-Contact Relationship Detected | Contact, was found as COVID-19 confirmed case                               | 1: Yes, 0: No | *                  |
| Friend                             | Contact, spent time and at least had one meal with a confirmed case during the past 14 days (Excluding household members of a confirmed case) | 1: Yes, 0: No | *                  |
| Household                          | Contact, shared a household with a confirmed case during the past 14 days (Household member of a confirmed case) | 1: Yes, 0: No | *                  |
| Sex differential                   | Whether contact, and a confirmed case are the same gender                   | 1: Yes, 0: No | *                  |
| Num. Case Exposure                  | The number of confirmed cases, to which contact was exposed to in the last 14 days | Numerical     | *                  |
| Age                                | Infected case’s age                                                        | Numerical     | *                  |
| Case’s spreading ratio             | Infected case’s spreading ratio (Number of contacts which had been infected by the case, divided by a total number of all contacts who were exposed to the case) | Ratio         | *                  |
| Additional control variables       |                                                                            |               |                    |
| Contact’s age                      | Contact’s age                                                              | Numerical     | *                  |
| Case’s condition                   | Admitted condition of an infected case                                     | 1: Patient with pneumonia, 0: Patient without pneumonia                   | *                  |
| Case’s Sickday                     | Case’s illness days before admitted to hospital                             | Numerical     | *                  |
The additional factors were included to enhance the understanding of the epidemiology, at the cost of leaving out a number of recorded cases from the analyses because of the missing data. The number of observations in the analysis therefore declined from 1033 to 873, as there were 160 high-risk contacts without the reported age data, due to a non-disclosure of personal information choice by contacts. In addition, there were 45 confirmed cases classified as asymptomatic for which we had no information, as their number of sick days in a hospital was simply recorded 0. Hence, we decided to clearly identify in the result table that this extended model is only applicable to high-risk contacts who were exposed to symptomatic confirmed cases. Thus, the number of high-risk contact observations for this extended model excluded 243 HRCs who were exposed to 45 asymptomatic confirmed cases. It was later found that the output generated from each model was robustly reported across both baseline and the extended models. Result from the extended models are illustrated in the appendix Table 6 As the second objective of this paper was to investigate the impact of high-risk contact quarantine policy, the main analysis focused on each risk factor before, and after the high-risk contact quarantine policy implementation. Moreover, the infection probability of each HRC from the Eq. (1) was further statistically compared with t-statistic and through plotted line in Kernel Density estimation and Normal distribution, depicted in Fig. 5, to compare the effects of the state quarantine policy per se. The paper is reported following the STROBE statement [25].

### Ethics approval and consent

For this study, Institutional Review Board or any formal approval from ethical committee is not required as this study use pre-existing and de-identified data. Data is anonymized and irreversibly de-identified to protect patient and all contacts. It should be noted that identifying details of the observations in this study completely remain as anonymity. No trials were conducted on either humans or animals for this study.

### Table 2

Descriptive statistics and related testing on each exogenous variable

| Variables | Hangout and having meal together | Living in the same roof |
|-----------|---------------------------------|-------------------------|
|           | Detected (1) | Non-infection (0) | Detected (1) | Non-infection (0) |
| Number of observations (1,108) | 172 | 936 | 172 | 936 |
| (Yes) Friend=1 | 27 | 360 | - | - |
| (No) Not Friend=0 | 145 | 576 | - | - |
| (Yes) Household =1 | - | - | 82 | 89 |
| (No) Not household =0 | - | - | 90 | 847 |
| Correlation coefficient (Probability) | -0.1887(0.0000) | 0.3876(0.0000) |

### Table 3

T test and multiple comparisons of detected and non-detected

| Variables | Number of case exposure Detected | Case’s age Detected | Case spreading ratio Detected | Non-infection | Non-infection | Non-infection |
|-----------|---------------------------------|---------------------|-------------------------------|---------------|---------------|---------------|
| Number of observations | 172 | 936 | 172 | 936 | 172 | 936 |
| Minimum value | 1 | 1 | 1 | 0.25 | 0 | 0 |
| Maximum value | 7 | 7 | 97 | 78 | 100 | 83 |
| Mean value | 2.23 | 1.42 | 36.61 | 33.6 | 42% | 11% |
| Number of observations | 1106 | 937 | 1089 |
| t Stat | 41.0016 | 63.9774 | -0.4650 |
| P value | 0.0000 | 0.0000 | 0.6420 |
2.2. Role of funding

No funding was received in study design, data collection, data analysis, data interpretation, writing of the article, or the decision to submit for publication. All authors had full access to all anonymized data in the study and were responsible for the decision to submit the Article for publication.

3. Results

Most of the infected cases were classified as transmissions amongst the local residents, although Phuket was the main travel destination in Thailand for both Chinese and European travellers during the first quarter of this year. From the list of confirmed cases, it was found that less than 10% of all confirmed cases in Phuket were foreign imported cases [21]. Infections through local transmissions, presented as gray bars in the Fig. 3, sharply increased after the initial imported transmissions subdued.

(Foreign imported case is a confirmed case who had travelled/came from abroad, Secondary transmission from foreign imported case is a confirmed case who was exposed to/contracted the disease from a foreign imported case, Local to local transmission is a confirmed case who got exposed to/contracted the disease from a local confirmed case who had no record of traveling abroad)

Tables 2 and 3 report both descriptive statistics and related preliminary testing. The discussion of the results is presented in the sub sections following the tables.

3.1. Household and friends

For risk of transmission within a household, a simple correlation report from Table 2 shows a positive correlation between sharing a household and contact’s infection. Results from the baseline model (Table 4, Variable: Household,) also reports that a contact who had

![Figure 4. Comparison of the infection spreading rates from an exposure to a confirmed case with several documented secondary transmissions, and exposure to a confirmed case with a single secondary transmission.](image-url)
lived within the same household with a confirmed case was more exposed to infection than a contact who had not lived in the same household with a confirmed case. Specifically, the marginal probability of infection from the confirmed case, when sharing a household, was reported to be 25%. In the extended model, which includes additional variables, the reported marginal probability of infection caused by a household factor was very similar. Regarding SARS-CoV-2 transmissions amongst friends in Phuket, it was found that friends-status between confirmed cases and contacts had no statistical linkage to an infection probability of a contact.

4. Number of exposures to confirmed cases and exposure to a super spreader

The results from baseline model illustrated in Table 4 show that if a contact was exposed to one additional confirmed case (\(Num\text{-}Case\text{Exposure}_{i}\)) his or her probability of infection would increase by 3.13%, with all other factors being constant. Moreover, it was found that the marginal effect of a variable, representing an exposure to a confirmed case with high reproduction capabilities, was reported as 0.57%. The exposure to a confirmed case, such as case 169 would yield a marginal probability of infection 22.8% higher than an exposure to a confirmed case 123, as illustrated on the right side of Fig. 4.

5. Contact's age, case's admitted condition, and days before hospitalization

It was found that contact's infection probability was negatively linked with the age of a contact. The reported coefficient also enabled us a finding that if an interaction occurs between a confirmed case between confirmed cases and contacts had no statistical linkage to an infection probability of a contact. Regarding SARS-CoV-2 transmissions amongst friends in Phuket, it was found that friends-status between confirmed cases and contacts had no statistical linkage to an infection probability of a contact.
and contacts of different ages, the contacts that are older in interaction (by 10 years) have 1.4% more probability of contracting the infection than the younger contacts involved in the interaction. Results from extended models, Table 9 reported in the appendix, reveal that a condition of a confirmed case when admitted to a hospital was not found to have a statistical linkage with contact’s infection probability. In addition, the duration between confirmed case’s illness start date, and a hospitalization date was not found to have a positive link with contact’s infection probability.

6. The role of high-risk contacts state quarantine in epicenter districts

The question in this subtopic was whether the State Quarantine (SQ) for all high-risk contacts was effective in mitigating COVID-19 infection spread in Phuket. As most of the confirmed cases erupted in the two districts, the impact of SQ policy on infection probability was assessed in these epicenter districts. It was found that 10.68% of all high-risk contacts were confirmed to be infected before the quarantine, and 4.55% after the policy has been enforced. In appendix Table 7, the simulated infection probability from the baseline model was categorized to before and after SQ policy. It was found that the infection probability per se of a contact was 36% before quarantine, and 29% after the quarantine, while the result from t-test indicated that this difference is statistically significant at 95% confidence interval. The results from Table 5 explain how all HRCs quarantine policy alleviated the transmittable magnitude of SARS-CoV-2 in Phuket epicenter districts.

From Table 5, it was found that before the state quarantine policy for all HRCs was implemented, interaction between a confirmed case and friends could statistically yield higher probability of COVID-19 infection in two highest affected districts. However, the marginal effect of this variable was no longer statistically significant after SQ had been implemented. It was also found that the contact’s marginal chance of infection caused by confirmed cases with high record of secondary transmission declined from 4.5% to 0.8%. In addition, risk factor caused by living in the same household with a confirmed case significantly declined from 93% to 51% post SQ enforcement.

Moreover, a comparison of estimated probability of infection from baseline model categorized as before and after the state quarantine for all HRC through Kernel density estimation and normal distribution function in Fig. 5 also confirm the reduction of HRC’s infection probability after state quarantine policy was implemented.

7. Discussion

This study found strong empirical evidence for SARS-CoV-2 household transmissions. These findings conform with the recent findings Jing et al. [13] and Li et al. [16]’s results from China, in which the household secondary attack was relatively imminent. Similar results were found in California, Illinois and Westchester, recording a very high chance of transmission between husbands and wives [9]. In addition, Chen et al. [5] and Cauchemez et al. [4] also found that the virus spread more easily amongst people living within the same household [5,12]. A plausible link between these international results, and findings in Phuket could be derived from the recent behavioural studies of Thai people during the COVID-19, by Faculty of Economics, Chulalongkorn University [10], which found that it was still relatively common for households in the rural area of Thailand to share plates during meals.

The findings of this analysis confirmed the higher infection rates upon exposure to a confirmed case with several documented secondary transmissions in Phuket. Even in the extended models, in which more control variables were taken into consideration, the effect from this variable did not dilute. The higher risk of infection from certain cases could be similar to the well-known super spreader Patient-31

| Group Statistics | Independent Samples Test | Independent Samples Test | 95% Confidence Interval of the Difference (Lower) | Mean Difference | Std. Error Difference | 95% Confidence Interval of the Difference (Upper) | Std. Error Difference | Mean Difference | Std. Error Difference |
|------------------|--------------------------|--------------------------|-----------------------------------------------|----------------|----------------------|-----------------------------------------------|----------------------|----------------|----------------------|
| Quarantine       |                          |                          |                                               |                |                       |                                               |                      |                |                       |
| Detected After Quarantine | 241 | 0.29 | 0.279 | 0.018 | 0.090 | 0.007 | 0.016 | 0.039 | 0.007 | 0.013 | 0.007 |
| Detected Before Quarantine | 91  | 0.36 | 0.328 | 0.001 | 1.000 | 0.075 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
in Daegu, South-Korea. An infectious disease expert McGraw [14] stated that the combination of pathogens, biology, surrounding environment and most importantly, the behaviour of confirmed cases define how some cases inherently have a high secondary attack rate. These results should alarm the authorities to promptly isolate all contacts who have been exposed to the confirmed cases who already infected other contacts.

Finally, it was found that contact’s infection probability was negatively linked with the age of a contact as suggested by Li et al. [16] and Liu et al. [17]. This finding further supports the general speculation which recommends the higher protective measures for the senior population, as they are more vulnerable to the COVID-19 in terms of infection and hospitalization.

As of August 2020, COVID-19 is still considered a concurrent pandemic, and many researchers across the globe are continuously investigating the newly emerged SARS-CoV-2 disease. Since this study was conducted in April 2020, the limitations of the analysis are the construction of the testing models, as it was based on the current understanding of the disease. Therefore, the list of the exogenous variables is not complete to include all factors which could potentially affect high risk contact’s infection probability. In addition, the surveillance data could only reflect whether contact had lived in the same

| Table 8 | Results from Probit regression of baseline model and model with contact age (Reported Coefficient) |
| Variables | Baseline model | Model with contact age |
| C | -2.462*** | -2.5125*** |
| (0.0000) | (0.0000) |
| Friend | 0.0675 | -0.0186 |
| (0.6470) | (0.9040) |
| Household | 1.0159*** | 0.9593*** |
| (0.0000) | (0.0000) |
| Num*CaseExposure | 0.1885*** | 0.1567*** |
| (0.0018) | (0.0114) |
| Sexdifferential | 0.0858 | 0.0610 |
| (0.3493) | (0.5213) |
| Case**s_age | 0.0035 | 0.0025 |
| (0.3394) | (0.4986) |
| Contact**age | 0.0342*** | 0.0333*** |
| (0.0000) | (0.0000) |
| R square | 0.3891 | 0.3698 |

Note: Number in parenthesis is the p value of coefficient, figures reported with * are significant with 0.10 significance level, ** are significant with 0.05 significance level, figures reported with *** are significant at 0.01 significance level.

| Table 9 | Results from Probit regression of extended model with admitted case condition and number of sick days to Hospital |
| Variables | Model with (applied only symptomatic cases) |
| Admitted case condition | Marginal Effect | Number of sick days to hospital | Marginal Effect |
| C | -2.4879*** | -2.3712*** |
| (0.0000) | (0.0000) |
| Friend | 0.0021 | 0.0081 |
| (0.9903) | (0.9903) |
| Household | 0.9734*** | 1.0145*** |
| (0.0000) | (0.0000) |
| Num*CaseExposure | 0.2138*** | 0.223*** |
| (0.0025) | (0.0002) |
| Sexdifferential | 0.0792 | 0.0009 |
| (0.4532) | (0.4796) |
| Case**s_age | 0.0001 | 0.0001 |
| (0.9845) | (0.8522) |
| Contact**age | 0.0113 | 0.0172 |
| (0.7328) | (0.7328) |
| Case**Condition | 0.0005*** | 0.0005*** |
| (0.0000) | (0.0000) |
| Case**Sickday | 0.0000 | 0.0000 |
| (0.0000) | (0.0000) |
| R square | 0.3905 | 0.3925 |

Note: Number in parenthesis is the p value of coefficient, figures reported with * are significant with 0.05 significance level, figures reported with ** are significant at 0.01 significance level.

| Table 10 | Results from Probit regression of baseline model before and after strict state quarantine (Reported Coefficient) |
| Variables | Epicenter districts | Partial quarantine policy |
| C | -2.3717*** | -5.2657*** |
| (0.0000) | (0.0166) |
| Friend | -0.1904 | 4.3646** |
| (0.6631) | (0.0131) |
| Household | 1.4526*** | 4.3066*** |
| (0.0000) | (0.0088) |
| Num*CaseExposure | 0.0084 | -0.5158** |
| (0.3607) | (0.0234) |
| Sexdifferential | 0.4244** | 0.6283 |
| (0.0155) | (0.1032) |
| Case**s_age | 0.0045 | -0.0113 |
| (0.4850) | (0.5683) |
| Case**s_spreadingratio | 0.0255*** | 0.1314*** |
| (0.0002) | (0.0010) |
| R square | 0.3553 | 0.6233 |

Note: Number in parenthesis is the p value of coefficient, figures reported with * are significant with 0.05 significance level, figures reported with ** are significant at 0.01 significance level.
household with the confirmed case as dummy variable, but the data does not contain detailed information on sizes of households, house sharing facilities, or physical dimensions of houses in this analysis. Finally, the present study is based on contact tracing data in Phuket, which has specific geographical characteristics. It is an island in which authorities can control population’s mobility at an absolute scale, therefore it would be interesting to test the applied statistical model on the other areas’ contact tracing data, on either national or global scale, and verify the results of this study.

Based on the contact tracing reports, 80% of the observed infections in Phuket were previously recorded as high-risk contacts of the confirmed cases. With application of the Probit statistical model, it was found that the main risk factor for COVID-19 transmission from confirmed cases to contacts was a household transmission. Therefore, the results of this study in Phuket confirmed these findings and revealed the importance of the quarantine of all high-risk contacts in state provided facilities. It should be restated that the marginal rate of household infections in the epicenter districts declined from 94% to 51% after the quarantine policy was fully implemented, while the marginal rate of infection amongst friends went from 91% to statistically insignificant. The results also suggest that a contact with a confirmed case with several documented secondary transmission had higher infection probability than a contact who was exposed to a typical confirmed case. It would be beneficial to conduct a further scientific and behavioral research of this group of confirmed cases. This study recommends to the researchers of all relevant disciplines, including science, behavioural studies, and pathology, to investigate this group of infected patients without delay in order to shed more light on their transmission mechanisms.

Declaration of Interests

We declare no conflict of interest.

Author Contributions

Phucharoen, C. initiated the research question of this study, contributed to the design of statistical method, interpreted results, and wrote discussion sections. Sangkaew, N. and Phucharoen, C. performed the statistical analysis, wrote introduction. Stosic, K. and Sangkaew, N. verified the flow of writing, calibrated all figures and tables. Stosic, K. proofread English of the manuscript.

Data sharing

The authors confirm that all statistical data supporting the findings of this study are available within the article and its supplementary materials. Anonymized and de-identified data that support the statistical analysis of this study are available from the corresponding author, Phucharoen, C. upon request.

Funding

No funding was received for this research.

Acknowledgements

We would like to thank Phuket’s Police Major General Rungrute Thakurapunyasiri, who invited us as analysts in COVID-19 task force for Phuket province. The authors would also like to gratefully acknowledge Dr. Krit Sakulpet, M.D. (Medical expert of Preventive Medicine), Deputy Chief of Phuket Public Health, and his team for their dedication in contact tracing.

Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.eclinm.2020.100543.

Appendix

Tables 6–10

References

[1] Bi Q, Wu Y, Mei S, Ye C, Zou X, Zhang Z, Liu X, Wei L, Truelove SA, Zhang T, et al. Epidemiology and Transmission of COVID-19 in Shenzhen China: analysis of 391 cases and 1286 of their close contacts. MedRxiv 2020. doi: 10.1101/2020.03.03.20028423.

[2] Biswas M, Rahaman S, Biswas TK, Haque Z, Ibrahim B. Effects of sex, age and comorbidities on the risk of infection and death associated with COVID-19: a Meta-analysis of 74807 confirmed cases. SSRN Electron J 2020. doi: 10.2139/ssrn.3566146.

[3] Breen R, Karlon, Holm KB. Interpreting and understanding Logits, Probits, and other nonlinear probability models. Annu Rev Sociol 2018;44:39–54. doi: 10.1146/annurev-soc-073117-041429.

[4] Cauchemez S, Donnelly CA, Reed C, et al. Household transmission of 2009 pandemic influenza A (H1N1) virus in the United States. N Engl J Med 2009;361:2619–27. doi: 10.1056/NEJMoa0905498.

[5] Chen N, Zhou M, Dong X, et al. Epidemiological and clinical characteristics of 99 novel coronavirus pneumonia in Wuhan, China: a descriptive study. Lancet 2020;395:507–13. doi: 10.1016/S0140-6736(20)30211-7.

[6] Colbourn T. COVID-19: extending or relaxing distancing control measures. Lancet Public Health 2020. doi: 10.1016/S2468-2667(20)30072-4.

[7] Cornelißen T, Sonderhol K. Partial effects in Probit and Logit models with a triple dummy-variable interaction term. Stata J 2009;9:571–83.

[8] Department of Disease Control. Weekly disease. Thailand: Department of Disease Control; 2020. [cited 2020 Apr 22] Available from: https://ddc.moph.go.th/en.

[9] Edwards, E. Family clusters: a common pattern for how the coronavirus spreads. 2020. https://www.nbcnews.com/health/health-news/family-clusters-common-pattern-how-coronavirus-spreads-n1150646 (accessed June 17, 2020).

[10] Faculty of Economics, Chulalongkorn University. Behavioral insights of Thai household under covid-19. Thailand: Faculty of Economics, Chulalongkorn University; 2020. [cited 2020 May 20] Available from https://drive.google.com/ file/d/1s-tvSan0HmDWFbfs4t8vIeqWA/view.

[11] Hahn, E.D., Soyer, R. Probit and Logit models: differences in the multivariate realm [online]. 2005 [cited 2020 May 20] Available from: https://www.researchgate.net/publication/241144770_Probit_and_Logit_Models_Differences_in_the_Multivariate_Realm.

[12] Hui DS, Ahtar EI, Kim YJ, Memish ZA, Oh MD, Zumla A. Middle East respiratory syndrome coronavirus: risk factors and determinants of primary, household, and nosocomial transmission. Lancet Infect Dis 2018;18:2127–27. doi: 10.1016/S1473-3099(18)30127-0.

[13] Jing QL, Liu M, Yuan J, Zhang Z, Zhang A, Dean NE, et al. Household secondary attack rate of COVID-19 and associated determinants. MedRxiv 2020. doi: 10.1101/2020.04.11.20056010.

[14] McGraw, E. What is a super spreader? An infectious disease expert explains. United States: The Conversation Media; 2010-2020. [cited 2020 January 31] Available from: https://theconversation.com/what-is-a-super-spreader-an-infectious-disease-expert-explains-130756.

[15] Khan A. What COVID-19 patients young and old can teach us about the coronavirus. United States: Los Angeles Times; 2020. [cited 2020 May 10] Available from https://www.latimes.com/science/story/2020-05-08/what-younger-covid-19-patients-can-teach-us-about-the-coronavirus.

[16] Li W, Zhang Bo, Lu J, Liu S, Chang Z, Cao P, et al. The characteristics of household transmission patterns of COVID-19 outbreak – An age-specific social contact characterization and intervention patterns-how-coronavirus-spreads-n1150646 (accessed June 17, 2020).

[17] McGraw, E. What is a super spreader? An infectious disease expert explains. United States: The Conversation Media; 2010-2020. [cited 2020 January 31] Available from: https://theconversation.com/what-is-a-super-spreader-an-infectious-disease-expert-explains-130756.

[18] McGraw, E. What is a super spreader? An infectious disease expert explains. United States: The Conversation Media; 2010-2020. [cited 2020 January 31] Available from: https://theconversation.com/what-is-a-super-spreader-an-infectious-disease-expert-explains-130756.

[19] Khan A. What COVID-19 patients young and old can teach us about the coronavirus. United States: Los Angeles Times; 2020; [cited 2020 May 10] Available from https://www.latimes.com/science/story/2020-05-08/what-younger-covid-19-patients-can-teach-us-about-the-coronavirus.

[20] National Health Security Office (NHSO) Thailand. A Public Health Ministry announcement. Thailand: National Health Security Office (NHSO) Thailand; 2020. [cited 2020 Apr 30] Available from: https://www.nhso.go.th/Default.aspx.

[21] Phuket Provincial Public Health Office. Districts in Phuket and number of confirmed cases by district. Thailand: Phuket Provincial Public Health Office; 2020. [cited 2020 May 5] Available from: https://www.pktodo.moph.go.th/.

[22] The Civil Aviation Authority of Thailand (CAAT). Travel advisory to passengers planning to enter Thailand. Thailand: The Civil Aviation Authority of Thailand
[23] The World Health Organization (WHO). Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19), 2020. https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf (accessed May 1, 2020).

[24] Urton J. UW team illustrates the adverse impact of visiting ‘just one friend’ during COVID-19 lockdown. Seattle, WA: UW News, University of Washington; 2020. https://www.washington.edu/news/2020/04/13/just-one-friend-covid-19/ (accessed May 1, 2020).

[25] von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet 2007;370(9596):1453–7 PMID: 18064739.