Demographic and initial outbreak patterns of COVID-19 in Thailand

Pavitra Jindahra¹ · Kua Wongboonsin¹ · Patcharawalai Wongboonsin²

Accepted: 20 July 2021 / Published online: 9 November 2021
© The Author(s), under exclusive licence to Springer Nature B.V. 2021

Abstract

This study investigated the demographic heterogeneity of COVID-19 infection to reveal the role of age structure and gender on COVID-19 diffusion patterns, demonstrating that the infection is distributed unevenly across ages, genders, and outbreak times. Based on cluster analysis, we analysed the 4-month COVID-19 outbreak data (N=3017) in Thailand from January 12 to May 12, 2020, covering the early to late outbreak period of the initial wave. Results revealed that there are 7 pertinent clusters of COVID-19 outbreaks. Infection risk was classified by age, sex, and confirmed infection period. Results showed that *elderly* and *young male* clusters were at risk of becoming infected at the very beginning of the wave. *Working-age male*, *young female*, and *elderly male* clusters were key clusters controlling transmission when spreading became pervasive. Relevant clusters addressed at the end of the wave included *general public* and *younger age* clusters. Unlike other regions, the infection risk in Thailand is interestingly stronger among younger age clusters and male populations. Even though elderly individuals are at risk of becoming infected earlier than other clusters, the infection proportion was low. The findings provide new insights into the risk for COVID-19 infection.

Keywords COVID-19 · Cluster analysis · Infection pattern · Heterogeneity · Age and gender

✉ Pavitra Jindahra
pavitraj@sasin.edu

Kua Wongboonsin
kua.wongboonsin@sasin.edu

Patcharawalai Wongboonsin
patcharawalai.w@chula.ac.th

¹ Sasin School of Management, Chulalongkorn University, Sasa Patasala Building, Soi Chula 12, Phayathai Road, Pathumwan, Bangkok 10330, Thailand

² ASEAN Studies Center, Chulalongkorn University, Pathumwan, Bangkok 10330, Thailand

Introduction

The novel coronavirus (COVID-19) outbreak was declared a Public Health Emergency of International Concern by the World Health Organization (WHO) on 31 January 2020. It was declared a Global Pandemic on 11 March 2020, when there were 6463 confirmed cases with 329 deaths. By early June 2020, the number of confirmed cases had reached 116,003, with 4292 deaths across the globe. COVID-19 is the name given by the World Health Organization (WHO) on February 11, 2020 for the disease caused by a novel coronavirus called severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). Formerly referred to as '2019 novel coronavirus' or '2019-nCoV', the disease was first identified among respiratory illness cases in Wuhan City, the centre of the outbreak in Hubei Province, China, and was reported to the WHO on December 31, 2019 (World Health Organization,). As part of the worldwide pandemic, Thailand has experienced 3 waves of COVID-19 as of 11 May 2021. Figure 1 shows the daily confirmed COVID-19 cases during 12 January 2020 to 11 May 2021.

The disease reached Thailand on January 13, 2020, according to the WHO Timeline—COVID-19 (World Health Organization, 2020n). Despite Thailand being recorded as the first country outside China to report a laboratory-confirmed novel coronavirus case, it was an imported case in an individual aged 61 years

Fig. 1 Daily COVID-19 Confirmed Cases in Thailand (12 January 2020–11 May 2021)

traveling via a direct flight from Wuhan together with five family members in a tour group of 16 people (World Health Organization, 2020k). This is due to Thailand being one of the most popular holiday destinations for the Chinese, especially for the lunar New Year celebration. In addition to the reportedly more than 25,000 travellers from Wuhan and other affected Chinese cities arriving in Thailand during the period of January 3–27, there were also residents returning from China (Ratcliffe et al., 2020). The first Thai national case was confirmed on January 22, 10 days after the first case was confirmed. Spreading was noticeable when the first family cluster was identified on January 28, followed by the first local transmission in Thailand on January 31. The latter case was a 51-year-old male taxi driver who had no history of travel to China. Without travelling history to China, the taxi driver contracted the disease from his Chinese tourist passengers who exhibited symptoms of coughing. The first health provider case was reported on February 15 (National News Bureau of Thailand, 2020; World Health Organization, 2020c; Pongpirul et al., 2020).

During the initial wave, new cases were relatively low compared to other countries. They ranged from 1 to 11 cases per day and were primarily travel-related. Local transmission was limited, which was the case until March 15, when the number of new cases rose to 32, followed by a sharp increase to over one hundred cases per day, attributable to a few large transmission clusters in crowded places in Bangkok, such as entertainment venues and boxing stadiums. As of March 31, 2020, the total number of reported laboratory-confirmed cases of COVID-19 was 1652. The confirmed cases were distributed across 60 of 76 provinces, with an age range of 1 month—84 years old. Figure 2 shows the Thailand COVID-19 map per province on March 31, 2020. Approximately half of infections occurred in Bangkok. Most new cases had close contacts with previously confirmed cases, followed by travel history abroad and visiting crowded places. Among the close contact infections, major sources were from parents, followed by family members, relatives, and unspecified. Thailand saw a decrease in the number of new cases per day, dropping to less than 10 by late April (DDC, 2020a; World Health Organization, 2020e, 2020g, 2020h).

Compared to many countries, Thailand appeared to have contained the initial pandemic over a relatively short time and managed the virus well throughout most of 2020. According to the WHO Thailand situation report on 7 February 2020, Thailand had strong capacities for case detection, risk assessment, case investigation, laboratory diagnosis, clinical management, infection prevention and control, as well as risk communication. The Thai health care system was well prepared with a special advisory group consisting of former high-level health officials to support the Health Ministry in its response to the outbreak. The Thai government enacted a curfew in early March while imposing a range of measures deterring social gatherings. When prevalence started to decline in early May, certain restrictions were lifted. Given concern of a potential resurgence in the spread of the virus, vaccine security was a national priority. In early June, the National Vaccine Committee endorsed a blueprint plan for technology transfer of high-potential vaccine prototypes from overseas and for local vaccine prototype development. These strategies were in collaboration with China, the USA, and other countries, as well as development partners, such as

Fig. 2 Thailand COVID-19 Map per Province (As of March 31, 2020)

the WHO and the Coalition for Epidemic Preparedness Innovations (World Health Organization,).

In December 2020 Thailand experienced the second resurgence of COVID-19 due to cross-border infection without quarantine from Malaysia and Myanmar. The key clusters were migrant worker communities in border provinces of Southwest Bangkok (DDC, 2020b). This quickly resulted in local transmission and spread to some other provinces. The second wave peaked on 26 January 2021 with 959 new cases (DDC, 2021a). This was promptly followed by the third wave, which originated in Bangkok's Thonglor nightlife venues in April 2021. Fearful of a further economic downturn, the Thai government did not implement the same degree of stringent measures with the subsequent outbreaks. It avoided lock downs while allowing inter-provincial travel during Thai New

Year holidays in April 2021 (Sivasomboon & Peck, 2021; Tan, 2021). The transmission was eminent among crowded communities and markets, passing through family members at home and colleagues at workplaces. The infection was looming with a new high daily confirmed number of cases of 2839 persons across 68 provinces on 24 April 2021. Case detection was carried out intensively while face mask policy was mandated in most provinces (DDC, 2021b, 2021c). The government aimed to achieve herd immunity against COVID-19 via vaccination, targeting 70% of the population by the end of 2021 (National News Bureau of Thailand, 2021; Tourism Authority of Thailand, 2021).

Many studies have attempted to examine the COVID-19 pandemic from various aspects. The contribution of demographics to the transmission and fatality rates of COVID-19 is a very recent development in the relevant literature. Among others, a study on the role of age structure in deaths in Italy and South Korea by Dowd et al. (2020) maintains that the interaction of demography and age-specific mortality for COVID-19 is powerful. In the UK, an investigation by Kulu and Dorey (2020) on the contribution of population age structure to mortality from COVID-19 lends support to the contribution of data disaggregation by age in infection fatality, to a high level of accuracy of a death rate projection. In a relatively large scope of geographical coverage, Dudel et al. (2020) also found that the age structure of detected infections explained over 50% of the difference in case fatality rates (CFRs) between countries with a low CFR and a high CFR, in addition to findings that the older population is at a relatively high risk. This was reflected in their study of the population-level CFR associated with COVID-19 between March 9 and April 22, 2020 in China, Germany, Italy, South Korea, Spain, the United States, and New York City. More studies are recommended on intergenerational contacts in national contexts, particularly among the squeezed 'sandwich' generation of adults who care for both the old and young. This is in addition to the need for data disaggregation by age and gender. The contribution is not limited to a real-time targeted forecast of hospitalisation numbers, but also extends to critical care needs (Dowd et al., 2020).

Despite these findings, COVID-19 transmission patterns remain unclear. Variations in COVID-19 transmission patterns were detected in different regions. To elucidate the demographic aspects of the COVID-19 outbreak in Thailand, this study investigated how demography defines infection patterns of the initial wave in terms of age and gender. This study examined distinctions among COVID-19-detected infections using a cluster analysis technique to demonstrate how the infection is unevenly distributed across age, gender, and outbreak time. The revealed outbreak patterns consequently sorted infection clusters in terms of infection risk levels. We believe that the findings of initial characteristic outbreak patterns will provide valuable insights for policy makers, the private sector, and other stakeholders to better prepare for successive waves. Better resource allocation and policy targeting should be considered that are geared towards different outbreak groups over the outbreak cycle.

Demography and transmission patterns of COVID-19

Key factors contributing to variations in the transmission pattern include regional differences in terms of population distribution across age groups, social distancing interventions, willingness of businesses and people to conform to containment measurements, and hygiene practices, especially wearing face masks. At the beginning of the pandemic, the elderly were the most vulnerable group to infection as the risk of infectivity increases with age (World Health Organization, 2020a; Dudel et al., 2020; Ferre & Rafe, 2020; Verity et al., 2020). For example, Stokes et al. (2020) reported that COVID-19 incidence in the early US pandemic was highest among older adults.

As the pandemic progresses, many studies found evidence that younger adults likely contribute to community transmission of COVID-19. A study from 6 representative cities (Wuhan, Beijing, Tianjin, Hangzhou, Suzhou, and Shenzhen) in China from December 2019–February 2020 suggested that disease trends were primarily driven by infectivity in two age groups, 23–44 and 45–64 (Liu, et al., 2020). These two age groups dominate the population size and case numbers. In particular, they are more engaged in social activities than other age groups. Similarly, Boehmer et al. (2020) found that during May–August 2020, the median age of COVID-19 cases in the US declined from 46 years in May to 37 years in July, and 38 years in August. People aged 20–29 years accounted for the largest proportion of total cases (> 20%) during June–August. A similar age shift was also reported in Europe (European Centre for Disease Prevention and Control, 2020). The median age of COVID-19 cases declined over the outbreak period from 54 years during January–May, to 39 years during June–July, as people aged 20–29 years constituted the largest proportion of cases during June–July. Younger adults are at higher risk of COVID-19 exposure due to occupational and behavioural factors. Moreover, younger adults might be less likely to follow community mitigation strategies, such as social distancing and avoiding group gatherings (Czeisler et al., 2020; Nagata, 2020).

Demographic background of Thailand

According to the Report of the Population Projections for Thailand 2010–2040 (Revision) by the Office of the National Economic and Social Development Council (NESDC Office, 2019), Thailand is estimated to have a population of 66.53 million persons in 2020. This number reflects a 0.41% annual increase from 2010 when there was a 63.79 million population. Over half (51.2%) of the 2020 population was female and 48.8% male, with a sex ratio of 95.4 males to 100 females, a slight increase from the 95.04 sex ratio in 2010 when 51.3% of the population was female and 48.7% male.

Figure 3 shows the population pyramid for Thailand in 2020, illustrating the age and gender structure of the Thai population. Thailand in 2020 is an aging

Fig. 3 Population pyramid for Thailand 2020 (%). *Sources:* Authors' calculation based on data from NESDC Office (2019); BORA (2019)

society where 18% of the total population is aged greater than 60 years old, with a much higher proportion of females (19.7%) than male elderly individuals (16.5%). The highest proportion of the population has shifted from the age group of 40–44 years old with 8.8% female and 8.6% male in 2010 to 50–54 years old in 2020 with 8.4% female and 8% male. Those in the age group of 15–19 years share the smallest proportion among working age groups in 2020, with 6.1% female and 6.7% male. This is in comparison to the 20–24 age group, who were of the smallest proportion among the working age groups in 2010 (6.7% female and 7.1% male).

In addition, the demographic structure of the Thai family can be described by urbanisation, migration, and intergenerational households. Urbanisation since the beginning of 1970s industrialisation in Thailand has driven a large number of rural Thai people, as well as international migrants from neighbouring countries, such as Cambodia, Lao PDR, and Myanmar, towards large cities in Thailand for better employment opportunities and standards of living (Wongboonsin et al., 2019).

Even though migration to urban areas increased the proportion of elderly people living alone in a household from 3.6% in 1994 to 10.4% in 2014, Keeratipongpaiboon (2012) found that more than 60% of ageing Thai people live in intergenerational household arrangements. That is, almost 30% of households house the elderly. Intergenerational households in Thailand are mostly in the form of generational coresidence (Hareven, 1996), in which elderly parents live with at least one adult child, and quasi-residence (Ofstedal et al., 1999), in which the parents' residence is adjacent to or within the same compound of their children. This is influenced by filial piety in Thai culture. The extended family living format implies that the family relationship between the old and the young is still close, allowing for filial caring.

Data

This study analysed confirmed COVID-19 case data in Thailand from January 12–May 12, 2020. The data are publicly available and updated daily (<https://data.go.th/dataset/covid-19-daily>) by the Department of Disease Control, Thailand Ministry of Public Health. Figure 4 shows daily confirmed COVID-19 cases in this 4-month period. The curve shows people infected in the initial wave over 4 months, with the pandemic rising to a peak and then falling.

The data contain daily confirmed COVID-19 cases, confirmed dates, and demographic variables of COVID-19 patients in terms of age and gender. An overview of the data summary statistics is provided in Table 1. There were 3017 confirmed COVID-19 patients with an age range of 0–97 years. The average patient from this cluster was 39 years old, and the case was confirmed on day 81 of the outbreak. Several demographic characteristics are worth noting. There were more infected males (54.29%) than females (45.71%). Male patients were also older in general. On average, male cases were confirmed earlier (day 80) than female cases (day 82). The distribution of age and confirmed days by sex are shown in Fig. 5.

The infection in Thailand was not evenly distributed over demographic variables. Interestingly, elderly individuals aged 60 years and older were infected much less often than other age groups. In particular, there were fewer elderly female cases than

Fig. 4 Daily COVID-19 confirmed cases of the initial wave in Thailand (January 12–May 12, 2020)

Table 1 Summary statistics

Variables	Overall		Male		Female	
	Mean	Std.	Mean	Std.	Mean	Std.
N	3017	1638	1379			
Size			54%	46%		
Age range	0–97	0–91	0–97			
Age	38.76	15.81	41.38	16.20	35.64	14.75
Confirmed day of the outbreak	80.73	13.14	79.72	12.85	81.94	13.40
Onset dates	Jan 12–May 12, 2020		Jan 22–May 11, 2020		Jan 12–May 12, 2020	

young female cases. For male confirmed cases, the working age group was dominant. During days 80–90, when spreading was accelerated, males became infected more than females. It is therefore conceivable that males exhibit a higher infection risk than their female counterparts, while young people present more risk than their older counterparts.

Cluster analysis

A cluster analysis was performed to characterise COVID-19 outbreak patterns in Thailand. Cluster analysis is an unsupervised machine learning method that reveals underlying patterns and properties of observations (Hastie et al., 2001). Pattern identification proceeds by grouping similar observations in the same cluster and assigning dissimilar observations to different clusters. This classification technique makes no prior assumptions about important differences within a population. Hence, cluster analysis is very suitable to organise, identify, and extract unknown pattern data, such as in the COVID-19 outbreak.

We examined distinctions among confirmed COVID-19 cases in Thailand. It is necessary to develop a representation of the Thailand COVID-19 outbreak relationship with demographic variables and outbreak time. That is, the identified outbreak occurrence patterns will share certain common characteristics of age, gender, and occurrence time. The study employs a two-step clustering approach. The details are as follows.

Two-step procedure

The analysis starts with Ward's agglomerate hierarchical clustering method (Ward, 1963) to identify the optimal number of clusters in the data. Clusters are formed based on minimising the loss of information across possible cluster combinations. Loss of information is the error sum of squares defined as the total squared deviation of every observation from the mean of the cluster. The cluster formation that minimises the error sum of squares is then chosen. The cluster formation output from

Fig. 5 Confirmed case distribution over demography

the hierarchical clustering procedure is reported by the agglomeration schedule, suggesting when to stop grouping clusters of observations. That is, the optimal number of clusters is signified when distances between clusters are reasonably large. The suggested number of clusters, together with managerial judgment, help determine an appropriate number of clusters for use in the second step.

Given the known number of clusters, K-Means clustering is then applied to assign each confirmed case observation to appropriate outbreak clusters accordingly. K-Means clustering starts with cluster centres (centroids), assigning observations closest to each cluster centre to that cluster. Reallocation is possible if it reduces the sum of internal cluster variability, which improves cluster classification. The second step is required because the K-Means clustering method is superior to the hierarchical approach in classifying observations into a cluster.

Patterns of COVID-19 Thailand outbreak

The two-step cluster analysis was performed by IBM SPSS Statistics Version 22. The cluster-based variables included age, gender, and confirmed case date. Age and confirmed date were standardised to remove unit bias. Gender is a dummy variable with a value of 1 if female and 0 otherwise. The cluster analysis suggested that there are 7 pertinent infection clusters from the initial wave in Thailand and that each cluster is variably characterised by size, age, gender, and outbreak time. The number of clusters was expounded first, followed by outbreak pattern profiling.

Number of clusters

Statistical fit, managerial relevance, and targetability are factors determining the appropriate number of clusters for the COVID-19 outbreak in Thailand. Statistically, Ward's agglomerate hierarchical clustering method was used to evaluate the optimal number of clusters. Cluster formations are shown in the agglomeration schedule. The agglomeration schedule (Table 2) reports the distance of the two clusters being merged. In particular, the recommended number of clusters corresponds to the step before a large distance increase in cluster nesting. Note that the increase in distance is quite uniform until stage 3011 when the numbers of clusters decreases from 7 to 6 with an increase of 266 units. Another jump of 621 units was detected when the number of clusters was reduced from 5 to 4. To better expound COVID-19 outbreak patterns, we opted for the seven-cluster solution because the loss of information was smaller, allowing for more data variability. The deeper characteristics of COVID-19 outbreak patterns hence entail better insights for policy makers to comprehend and design appropriate health policies.

Outbreak pattern profiling

Given the suggested seven clusters, K-Means clustering was applied to assign COVID-19 patients into the appropriate clusters. Generally, the timeline of the outbreak helps divide cases into 3 large clusters according to the COVID-19

Table 2 Agglomeration schedule

Stage	Agglomeration schedule							
	Number of clusters	Cluster combined		Distance	Diff. Dist	Stage cluster first appears		Next Stage
		Cluster 1	Cluster 2			Cluster 1	Cluster 2	
3007	10	1197	1451	1642.564		2999	3001	3011
3008	9	754	755	1817.238	174.674	2993	3003	3013
3009	8	1	31	2009.337	192.099	3000	3006	3012
3010	7	1822	2246	2209.705	200.368	3005	2987	3013
3011	6	1197	1548	2475.874	266.169	3007	2998	3015
3012	5	1	6	2767.719	291.844	3009	3004	3014
3013	4	754	1822	3115.877	348.158	3008	3010	3015
3014	3	1	1358	3736.587	620.710	3012	2997	3016
3015	2	754	1197	5210.136	1473.549	3013	3011	3016
3016	1	1	754	6780.685	1570.549	3014	3015	0

confirmation timeline, namely, early, mid, and late outbreak clusters. *Early outbreak clusters* are the early infected cases found in the first half of the outbreak cycle, including clusters 2 and 3. The early outbreak clusters are influential, though smallest in size (1.4%), dictating the rate of successive infection of the wave. *Mid outbreak clusters* are clusters 5–7, of which the shading of the occurrence period is highlighted during the middle of the initial wave. These clusters accounted for most cases (73.8%) found during the 4-month period. *Late outbreak clusters*, on the other hand, comprise clusters 1 and 4. Their confirmed case date was found at the end of the outbreak wave. The late outbreak clusters accounted for 24.8% of infections.

Age and gender further characterised differences among clusters. Males were dominant in clusters 3, 6, and 7, as the shading is highlighted more in males. Younger cases characterised clusters 1, 3, 5, and 7, whereas older cases were more prevalent in clusters 2, 4, and 6. Summary statistics of cluster demographic characteristics are shown in Table 3. The description of the 7 clusters is described below.

Early outbreak clusters

The early infected clusters cover the outbreak period from January 12 to March 2, 2020 and comprise Cluster 2 and Cluster 3.

Cluster 2 (elderly) Cluster 2 represents the most early infected patients whose confirmed cases occurred in the first five weeks (January 12 to February 26, 2020) of the Thailand outbreak, accounting for 0.6% (N = 19) of confirmed cases. Patients in this cluster were between 49–74 years old, with a mean age of 63 years old. Generally, they were older than an average COVID-19 patient. Males and females were found to be proportionately equal. Their cases were confirmed on average earlier than the rest, the 22nd day of the outbreak period.

Table 3 Summary statistics of cluster demographic characteristics

Variables	Early outbreak infected clusters			Mid outbreak infected clusters			Late outbreak infected clusters		
	Cluster 2	Cluster 3		Cluster 7	Cluster 5	Cluster 6	Cluster 4	Cluster 1	
N	19	24		784	905	536	455	294	
Size	0.6%	0.8%		26.0%	30.0%	17.8%	15.1%	9.7%	
Male	47%	58%		77%	30%	74%	47%	45%	
Female	53%	42%		23%	70%	26%	53%	55%	
Age	49–74	3–45		22–53	0–39	49–97	33–83	0–45	
Age mean	63	28		39	25	62	48	23	
Infected day mean	22	29		74	78	78	93	102	
Infected days	Jan 12–Feb 26, 2020	Jan 25–Mar 2, 2020		Mar 5–Apr 7, 2020	Mar 11–Apr 10, 2020	Mar 14–Apr 23, 2020	Apr 1–May 12, 2020	Apr 9–May 12, 2020	

Cluster 3 (young male) Cluster 3 is the patient cluster that became infected from January 25–March 2, 2020. Its cluster size was approximately 0.8% ($N = 24$). This cluster is male dominant (58%), with a younger age group of 3–45. The mean age of this cluster was 28 years old on average. Cases were generally confirmed on the 29th day of the outbreak period.

Mid outbreak clusters

These clusters were detected during the mid-outbreak period (March 5–April 23, 2020). According to the outbreak timeline, the mid outbreak clusters include Cluster 7, Cluster 5, and Cluster 6.

Cluster 7 (working age male) Cluster 7 is the second largest cluster, with a cluster size of 26% ($N = 784$). This cluster of COVID-19 cases was confirmed from March 5–April 7, 2020. This cluster has more males (77%) than females (23%). Interestingly, patients in this cluster were younger than average. They were 39 years old with ages ranging from 22 to 53 years old. The average confirmed date was reported to be the 74th day of the outbreak period.

Cluster 5 (young female) Cluster 5 refers to the confirmed cases during March 11–April 10, 2020, accounting for 30% of total cases. This is the largest cluster, with 905 patients. This cluster was female dominant (70%), with an average age of 25 years old. On average, patients in this group had confirmed cases on the 78th day of the outbreak period.

Cluster 6 (elderly male) Cluster 6 accounted for 17.8% ($N = 536$) of COVID-19 patients. The majority of this cluster was male (74%), with a mean age of 62 years old. They belong to the age group of 49–97 years whose cases were confirmed from March 14–April 23, 2020. Similar to those in cluster 5, patients in this group, on average, had cases confirmed on the 78th day of the outbreak period.

Late outbreak clusters

The late outbreak period refers to April 1–May 12, 2020. The late outbreak cluster comprised Cluster 4 and Cluster 1 according to the timeline.

Cluster 4 (general public) Cluster 4 is 15.1% ($N = 455$) in size. Cluster 4 had approximately equal proportions of males (47%) and females (53%) who became infected later in the outbreak (April 1–May 12, 2020). On average, they were 48 years old. The age range of this cluster was 33–83 years old. Their cases were confirmed around the 93rd day of the outbreak period.

Cluster 1 (younger adult) Cluster 1 size is 9.7% (N = 294). This group was younger, with a mean age of 23 years old. There were more female (55%) than male (45%) patients in this group. On average, members of this cluster became infected on the 102nd day of the outbreak period.

Discussion

Thailand is an ageing society with 18% of the total population aged above 60 years, with a much higher proportion of female (19.7%) than male elders (16.5%). Many countries across the globe reflect older persons as the most sensitive group due to the ageing process (Ferre and Rafe, 2020). This would make one expect that the confirmed cases of COVID-19 in Thailand should be more numerous among female elderly individuals. However, the opposite was true for the COVID-19 initial wave of Thailand. The findings reveal that infections in Thailand were unevenly and heavily distributed towards younger generations, with an intensity towards males. The infection proportion of Thai elderly individuals remains low.

Early outbreak period (January 12–March 2, 2020)

The early outbreak (January 12–March 2, 2020) in Thailand was kicked off by imported cases. Despite the smallest size (1.4%), the early outbreak is the most critical period for transmission confinement, as it determines the behavior of subsequent COVID-19 spreading. Hence, early outbreak clusters are high-risk clusters that must be addressed by policy makers for swift containment measures. High-risk cluster identification will help curb the spread and even stop the infection before it becomes prevalent. This study shows that there are two clusters with a high risk of becoming infected and spreading to others. These are elderly and young male clusters (Clusters 2 and 3) with mean ages of 63 and 28 years old, respectively. Even though elderly individuals were found to be at risk of becoming infected earlier than other age groups, confirmed cases in the 60+ age group during the early outbreak period were cases imported from China. Nine of them were Thai citizens returning from China, while the rest were Chinese visitors (World Health Organization Thailand 2020d). This is a different pattern from other regions, where the infected cases in the 60+ age group were primarily indigenous.

The policy during this period emphasised screening imported cases. COVID-19 has been added to a list of dangerous infectious diseases, giving the government authority to respond quickly under the Communicable Diseases Act B. E 2558 (2015). To limit importing cases from abroad, international travel restrictions were put in place during this early outbreak period. Visitors to Thailand travelling from affected areas were screened and provided information upon arrival by the Ministry of Public Health, with additional surveillance for visitors from high-risk regions, such as China, Japan, Singapore, South Korea, and Taiwan. However, prevention

measures for local transmission were not robust. Face masks, social distancing, and hygiene practices were not widely promoted. Therefore, local transmission inevitably surfaced on January 31, 2020.

Mid outbreak period (March 5–April 23, 2020)

Mid outbreak transmission during March 5–April 23, 2020 was set off by major transmission clusters of crowded social gatherings in Bangkok, such as the entertainment venue and the boxing stadium. The accelerated pace of infection quickly became apparent. New cases rose from 1–11 cases per day to over a hundred, most of which had close contact with previously confirmed cases. Significant transmissions were recorded among family members, especially from their spouse, and those in an active lifestyle. The infectivity, as a result, soon became nationwide. Fifty percent of the confirmed cases were found in Bangkok, while the rest spread across the nation.

It took more than 70 days for mid-outbreak clusters to become infected. The risk of infectivity of mid-outbreak clusters was projected to be medium. However, the transmission risk during this period is high. The majority of confirmed cases (73.8%) were found during mid-outbreak clusters. Thus, mid-outbreak clusters are indispensable for disease control, as they affect the spreading pace, scale, and duration of wave infection. Thailand's COVID-19 situation became more severe since mid-March. The Thai government implemented many measurements to slow down the infection pace, ensuring that there are enough healthcare facilities for infected patients over time. First, on March 16, 2020, the government postponed the Thai New Year official holidays on April 13–15, 2020 to prevent social gathering and spreading of the disease. The Thai New Year Festival is like a Christmas holiday for people travelling back to their hometown to celebrate together. Second, on March 22, 2020, Thailand closed high-risk places and businesses, such as department stores, cinemas, and public service spots, including stadiums, sport facilities, playgrounds, cinemas, bars and entertainment venues. Only supermarkets, pharmacies, and take-out restaurants were allowed to continue operating. This business closure inevitably prompted younger people to return to their hometown. Third, the state of emergency was eventually enforced on March 26, 2020 for additional measures. Essentially, no gathering was allowed anywhere. Access to areas declared high risk for COVID-19 was not allowed. Vulnerable groups, such as elderly persons aged over 70 years, children under five years, and people with noncommunicable diseases, were prohibited from leaving home, except for medical treatment purposes. Social distancing and hygiene practices were enforced. Moving across provinces was impeded. On April 3, 2020, a 10 pm–4 am nationwide curfew was enforced. More importantly, COVID-19 risk communication was performed through public education, focusing on self-monitoring, practicing hygiene, and avoiding crowds.

The cluster identification performed herein suggests that working age males (Cluster 7), young females (Cluster 5), and elderly males (Cluster 6) are vital to target for transmission intervention and cycle duration truncation. The current trend of intergenerational households explains the diffusion sequence during this period.

Nearly fifty percent of households in Thailand are two-generation households, especially young immediate households (Wongboonsin et al, 2019). This supports subsequent infection of the young female cluster after the working-age male cluster. The working-age male cluster, which is the second largest cluster (26%), was infected earliest during the mid-outbreak time due to occupational and behavioural factors. Many cases are linked to entertainment venues and boxing stadiums. Young females, representing the largest cluster (30%), took on average a longer time to become infected than working-age males. Many of them contracted the virus from their spouses (Bangprapa, 2020).

The male elderly cluster was the latest cluster in the mid-outbreak period with the smallest size (17.8%). They became infected during March 4–April 23, 2020, exhibiting the lowest risk among the mid-outbreak clusters. This is the result of the return of workers from Bangkok and other urban areas to their rural hometown as businesses closed. This is attributed to the widespread transmission among this male elderly cluster, although there are less than eight percent of skip-generation households in Thailand. On the other hand, intergenerational household trends in Thailand allow for filial caring. Together with the implementation of the 70+elderly stay at home policy on March 26, 2020, this entails the delay of infection among elderly individuals and the smaller number of reported cases. Note that female elderly individuals were not representative cases during this mid-outbreak period. Their lower risk of COVID-19 exposure might be due to their less active lifestyle compared to their male counterparts.

Late outbreak period (April 1–May 12, 2020)

It was not until April that the COVID-19 situation in Thailand was under control. Implementation of a curfew from April 3, 2020 was immediately followed by an order of suspension of all commercial international flights and lockdown measures to varying degrees. The transmission rate started to slow down in April. The number of new cases was reduced from over 100 to 7 on April 30, 2020. The first phase of easing the lockdown was then implemented on May 1, 2020.

The arrival of the late outbreak period was due to many factors. In addition to the enforcement of confinement measures, mechanisms against the widespread nature of the virus in the general public included the active role of social media, public education focusing on self-monitoring for at-risk groups, and that of the 1,040,000 Village Health Volunteers (VHVs) nationwide and 15,000 Public Health Volunteers (PHVs) in Bangkok. The VHVs and PHVs conduct home visits, provide health education, deliver medicines, and make reports to public health authorities. They were managed by the Department of Health Service Support, Ministry of Public Health (World Health Organization Thailand 12 April 2020). More importantly, cooperation from the public is the key to success. Taweetilp Visanuyothin, spokesman for the Centre for COVID-19 Situation Administration, said that controlling an epidemic required 90% cooperation from the population—and the Thai people had given it (Kendall, 2020). Hygiene habits become viral in Thai society. A survey of 12,999 ASEAN adults conducted by YouGov/Imperial College London during 30

March–27 April 2020 reports that 95% Thai citizens wore face masks and 89% were frequently washing their hands (YouGov, 2020).

Confirmed cases during the late outbreak period accounted for 24.8%. Late infected clusters included the general public (Cluster 4) and younger aged (Cluster 1) groups. These clusters exhibited a low risk of infectivity among confirmed cases, as they developed infections the latest. Note that when easing of lockdown measures was implemented, the younger age group reappeared as an important cluster of transmission from April 9–May 12, 2020. As a result, the easing of lockdown likely contributed to the transmission characterised by this younger group (<45), whose lifestyle in general is more social.

Conclusions

This study decomposes the outbreak patterns of COVID-19 in terms of age, gender, and day of infection. The confirmed case composition was portrayed by seven outbreak pattern profiles. The clusters exhibited uneven COVID-19 infection risk. Distinctly, the study shows that the infection risk in Thailand is stronger among younger age groups than in older counterparts and more robust among males, than in females. In particular, the working-age male likely drove the dynamics of transmission during the first wave of COVID-19 in Thailand. Young females are the most vulnerable to infectivity. The younger aged groups (<45), on the other hand, were found to characterise transmission when containment measures were weak or relaxed, as evidenced during the early and late outbreak periods.

Even though Thailand is an ageing society, elderly individuals in Thailand are less vulnerable than one would expect. They accounted for only 10% of the total confirmed cases. This is different from other regions in which infection risk was looming among elderly individuals, especially during the early phase of the pandemic. Moreover, females in Thailand (46%) are less susceptible to infectivity than their male counterparts (54%). This was true for both people of working age and retirees. This pattern is in contrast with that of European countries where people of working age diagnosed with COVID-19 substantially outnumbered infected men (Sobotka, et al, 2020).

COVID-19 transmission occurs person to person rather than from a common source. Successive waves are expected until the pool of susceptible people is exhausted, or control measures are implemented. Findings of the initial wave characteristics would therefore be helpful in preparing for successive waves. That is, when local cases are not detected for quite some time and control measures are relaxed, it is likely that the subsequent wave will be triggered by imported cases. COVID-19 screening in imported cases should be enforced. Intervention policies should be geared toward the younger age groups, as they are more susceptible to infectivity and should receive attention. Social gatherings with the presence of infected cases are expected to lead to widespread transmission if people do not continue hygiene practices (e.g., wearing face masks and hand sanitization). When spread is unavoidable and widespread, intervention policies should be targeted, particularly among working-age male and younger female clusters.

Our study contributes new insights to unknown COVID-19 infection patterns in Thailand. The infection pattern identification method in this study can be applied to other countries, recognising relevant clusters likely contributing to transmission dynamics. Infection pattern identification helps detect who is at risk of becoming infected before others and who is likely to trigger transmission. This implies cluster priorities for prevention and intervention implications, allowing for policy targeting. Resource allocation can be better prioritised according to the cluster risk and outbreak cycle. It must be noted that the identified transmission patterns are influenced by the transmission environment structure, including the population structure, the intervention measures enforced, and the willingness of citizens and businesses to accept restrictions. Therefore, the detected patterns in different regions could vary.

Funding This research was funded by the Sasin School of Management.

Availability of data and material <https://data.go.th/dataset/covid-19-daily>.

Declarations

Conflict of interest The authors declare that they have no conflicts of interest.

References

- Bangprapa, M. (2020). *Govt warns of Covid-19 threat from family members at home*. Bangkok Post. Retrieved May 10, 2020, from <https://www.bangkokpost.com/thailand/general/1897860/govt-warns-of-covid-19-threat-from-family-members-at-home>.
- Boehmer, T. K., DeVies, J., Caruso, E., van Santen, K. L., Tang, S., Black, C. L., et al. (2020). Changing Age Distribution of the COVID-19 Pandemic — United States, May–August 2020. *Morbidity and Mortality Weekly Report*, 69(39), 1404–1409. <https://doi.org/10.15585/mmwr.mm6939e1>.
- BORA. (2019). *Official statistics registration system: Number of population by age as of December 2019*. Bureau of Registration and Administration, Department of Provincial Administration, Ministry of Interior, December.
- Czeisler, M. É., Tynan, M. A., Howard, M. E., Honeycutt, S., Fulmer, E. B., Kidder, D. P., Robbins, R., Barger, L. K., Facer-Childs, E. R., Baldwin, G., Rajaratnam, S. M. W., & Czeisler, C. A. (2020). Public attitudes, behaviors, and beliefs related to COVID-19, stay-at-home orders, nonessential business closures, and public health guidance—United States, New York City, and Los Angeles, May 5–12, 2020. *Morbidity and Mortality Weekly Report*, 69(24), 751–758. <https://doi.org/10.15585/mmwr.mm6924e1>.
- DDC. (2020a). *The coronavirus disease 2019 situation report no. 155*. Department of Disease Control, Ministry of Public Health. Retrieved June 10, 2020, from <https://ddc.moph.go.th/viralpneumonia/eng/file/situation/situation-no155-060663.pdf>.
- DDC. (2020b). *The coronavirus disease 2019 situation report no. 349*. Department of Disease Control, Ministry of Public Health. Retrieved May 14, 2021, from <https://ddc.moph.go.th/viralpneumonia/eng/file/situation/situation-no349-191263.pdf>.
- DDC. (2021a). *The coronavirus disease 2019 situation report no. 383*. Department of Disease Control, Ministry of Public Health. Retrieved May 14, 2021, from <https://ddc.moph.go.th/viralpneumonia/eng/file/situation/situation-no383-260164.pdf>.
- DDC. (2021b). *The coronavirus disease 2019 situation report no. 471*. Department of Disease Control, Ministry of Public Health. Retrieved May 14, 2021, from <https://ddc.moph.go.th/viralpneumonia/eng/file/situation/situation-no471-240464.pdf>.

- DDC. (2021c). *The coronavirus disease 2019 situation report no. 488*. Department of Disease Control, Ministry of Public Health. Retrieved May 14, 2021, from <https://ddc.moph.go.th/viralpneumonia/eng/file/situation/situation-no488-110564.pdf>.
- Dowd, J. B., Adriano, L., Brazel, D. M., Rotondi, V., Block, P., Ding, X., et al. (2020). Demographic science aids in understanding the spread and fatality rates of COVID-19. *Proceedings of the National Academy of Sciences of the United State of America*, 117(18), 9696–9698. <https://doi.org/10.1073/pnas.2004911117>
- Dudel, C., Riffe, T., Acosta, E., Raalte, A., Strozza, C., & Myrskylä, M. (2020). Monitoring trends and differences in COVID-19 case fatality rates using decomposition methods: Contributions of age structure and age-specific fatality. *PLoS ONE*, 15(9), e0238904. <https://doi.org/10.1371/journal.pone.0238904>
- European Centre for Disease Prevention and Control. (2020). *Rapid risk assessment: coronavirus disease 2019 (COVID-19) in the EU/EEA and the UK—eleventh update: resurgence of cases*. Retrieved October 21, 2020, from <https://www.ecdc.europa.eu/en/publications-data/rapid-risk-assessment-coronavirus-disease-2019-covid-19-eueea-and-uk-eleventh>.
- Ferre, J., & Rafe, A. A. (2020). *COVID-19 and older persons: A defining moment for an informed, inclusive and targeted response*. UN DESA Policy Brief #68. United Nations, Department of Economic and Social Affairs. Retrieved June 15, 2020, from https://www.un.org/development/desa/ageing/wp-content/uploads/sites/24/2020/05/PB_68.pdf.
- Hareven, T. K. (1996). Introduction: Aging and generational relations over the life course. In T. K. Hareven (Ed.), *Aging and generational relations over the life course: A historical and cross cultural perspective* (pp. 1–12). Walter de Gruyter & Co.
- Hastie, T., Tibshirani, R., & Friedman, J. (2001). *The elements of statistical learning: Data mining, inference, and prediction*. Springer-Verlag.
- Keeratipongpaiboon, T. (2012). *Population ageing: Changes in household composition and economic behaviour in Thailand*. Ph.D. Thesis, SOAS University of London.
- Kendall, D. (2020). *How has Thailand kept the lid on Covid?* Bangkok Post. Retrieved November 17, 2020, from <https://www.bangkokpost.com/thailand/special-reports/1933736/how-has-thailand-kept-the-lid-on-covid>.
- Kulu, H., & Dorey, P. (2020). The Contribution of age structure to the number of deaths from COVID-19 in the UK by geographical units. *medRxiv* preprint, April 22. doi: <https://doi.org/10.1101/2020.04.16.20067991>.
- Liu, Y., Gu, Z., Xia, S., Shi, B., Zhou, X., Shi, Y., & Liu, J. (2020). What are the underlying transmission patterns of COVID-19 Outbreak? An age-specific social contact characterization. *Eclinical Medicine*, 22(2020), 1–10. <https://doi.org/10.1016/j.eclinm.2020.100354>
- Nagata, J. M. (2020). Supporting young adults to rise to the challenge of COVID-19. *Journal of Adolescent Health*, 67(2), 297–298. <https://doi.org/10.1016/j.jadohealth.2020.04.020>
- National News Bureau of Thailand. (2020). *Confirmed Coronavirus cases in Thailand remain at 32*. Retrieved May 5, 2020, from <https://thainews.prd.go.th/en/news/detail/TCATG200210190841054>.
- National News Bureau of Thailand. (2021). *1 Million Doses of Sinovac Vaccine Arrive in Thailand*. Retrieved May 11, 2021, from <https://thainews.prd.go.th/en/news/detail/TCATG210410190955642>.
- NESDC Office. (2019). *Report of the population projections for Thailand 2010–2040 (revision)*. Office of the National Economic and Social Development Council August.
- Ofstedal, M., Knodel, J., & Chayovan, N. (1999). Intergenerational support and gender: A comparison of four Asian countries. *Asian Journal of Social Science*, 27, 21–41. <https://doi.org/10.1163/030382499X00039>
- Pongpirul, W., Pongpirul, K., Ratnarathon, A., & Prasithsirikul, W. (2020). Journey of a Thai taxi driver and novel coronavirus. *The New England Journal of Medicine*, 382(11), 1067–1068. <https://doi.org/10.1056/NEJMc2001621>
- Ratcliffe, R., McCurry, J., & Devlin, H. (2020). *Coronavirus: first human transmission in Thailand as death toll hits 258*. The Guardian. Retrieved May 29, 2020, from <https://www.theguardian.com/world/2020/jan/31/coronavirus-thailand-confirms-first-human-to-human-transmission>.
- Sivasomboon, B., & Peck, G. (2021). *Bangkok nightlife clusters expose Thailand's virus stumbles*. AP News. Retrieved May 12, 2021, from <https://apnews.com/article/world-news-coronavirus-pandemic-prayuth-chaan-ocha-bangkok-thailand-fdd0bc6333d378172aa36a4593ff0c97>.
- Sobotka, T., Brzozowska, Z., Muttaraka, R., Zeman, K., & di Lego V. (2020). Age, gender and COVID-19 infections. *medRxiv*. <https://doi.org/10.1101/2020.05.24.20111765>.
- Stokes E. K., Zambrano, L. D., Anderson, K. N., Marder, E. P., Raz, K. M., Felix, S. E. B., Tie, Y., Fullerton, K. E. (2020). *Coronavirus disease 2019 case surveillance—United States, January 22–May 30, 2020*.

- Morbidity and Mortality Weekly Report*, 69(24), 759–765. <https://www.cdc.gov/mmwr/volumes/69/wr/mm6924e2.htm>.
- Tan, H. Y. (2021). Thai government's reluctance to impose Covid-19 lockdown stirs unease. *Straits Times*. Retrieved May 12, 2021, from <https://www.straitstimes.com/asia/se-asia/softly-does-it-bangkoks-covid-19-containment-strategy-stirs-unease>.
- Tourism Authority of Thailand. (2021). *Thailand reopening roadmap goes ahead amid COVID-19 vaccine rollout*. TAT Press Release. Retrieved May 11, 2021, from <https://www.tatnews.org/2021/05/thailand-reopening-roadmap-goes-ahead-amid-covid-19-vaccine-rollout/>.
- Verity, R., Okell, L. C., Dorigatti, I., Winskill, P., Whittaker, C., Imai, N., et al. (2020). Estimates of the severity of coronavirus disease 2019: A model-based analysis. *The Lancet Infectious Diseases*, 20(6), 669–677. [https://doi.org/10.1016/S1473-3099\(20\)30243-7](https://doi.org/10.1016/S1473-3099(20)30243-7)
- Ward, J. (1963). Hierarchical grouping to optimize and objective function. *Journal of the American Statistical Association*, 58, 236–244.
- World Health Organization. (2020a). *Coronavirus disease 2019 (COVID-19): Situation report, 51*. Retrieved May 2, 2020, from <https://apps.who.int/iris/handle/10665/331475>.
- World Health Organization. (2020b). *Novel Coronavirus (2019-nCoV) WHO Thailand Situation Report*. Retrieved May 2, 2020, from https://www.who.int/docs/default-source/searo/thailand/20200207-tha-sitrep-02-ncov-final.pdf?sfvrsn=b94cadff_0.
- World Health Organization. (2020c). *Coronavirus disease 2019 (COVID-19): WHO Thailand Situation Report*. Retrieved May 2, 2020, from https://www.who.int/docs/default-source/searo/thailand/20200215-tha-sitrep-05-ncov-final.pdf?sfvrsn=b4e36c2e_0.
- World Health Organization. (2020d). *Coronavirus disease 2019 (COVID-19) WHO Thailand Situation Report*. Retrieved May 2, 2020, from https://www.who.int/docs/default-source/searo/thailand/20200225-tha-sitrep-08-covid-19.pdf?sfvrsn=ff115175_0.
- World Health Organization. (2020e). *Coronavirus disease 2019 (COVID-19) WHO Thailand Situation Report*. Retrieved May 2, 2020, from https://www.who.int/docs/default-source/searo/thailand/2020-03-31-tha-sitrep-38-covid19-final5dcb2e0b5a5149c793287f0a7dbec0e6.pdf?sfvrsn=789a672a_0.
- World Health Organization (2020f). *Coronavirus disease 2019 (COVID-19) WHO Thailand Situation Report*. Retrieved May 2, 2020, from https://www.who.int/docs/default-source/searo/thailand/2020-04-12-tha-sitrep-50-covid19-finala0161210534642308b23a2965a57a44c.pdf?sfvrsn=5150506c_0.
- World Health Organization. (2020g). *Coronavirus disease 2019 (COVID-19): WHO Thailand Situation Report*. Retrieved May 2, 2020, from https://www.who.int/docs/default-source/searo/thailand/2020-04-29-tha-sitrep-67-covid19-en-final.pdf?sfvrsn=9c4bd08_0.
- World Health Organization. (2020h). *Coronavirus disease 2019 (COVID-19): WHO Thailand Situation Report*. Retrieved May 2, 2020, from https://www.who.int/docs/default-source/searo/thailand/2020-03-30-tha-sitrep-37-covid19-final-with-revision.pdf?sfvrsn=94dc7aba_0.
- World Health Organization. (2020i). *Coronavirus disease 2019 (COVID-19): WHO Thailand Situation Report*. Retrieved May 2, 2020, from https://www.who.int/docs/default-source/searo/thailand/2020-06-01-tha-sitrep-88-covid19.pdf?sfvrsn=1bab05b9_2.
- World Health Organization. (2020j). *Coronavirus disease 2019 (COVID-19): WHO Thailand Situation Report*. Retrieved May 2, 2020, from https://www.who.int/docs/default-source/searo/thailand/2020-06-03-tha-sitrep-89-covid19-finalcd290d4ebce14a8f9d14f01711ee05e4.pdf?sfvrsn=81c84725_0.
- World Health Organization. (2020k). *Novel Coronavirus—Thailand (ex-China)*. Retrieved May 2, 2020, from <https://www.who.int/csr/don/14-january-2020-novel-coronavirus-thailand/en/>.
- World Health Organization. (2020l). *WHO Coronavirus Disease (COVID-19) Dashboard*. Retrieved June 4, 2020, from <https://covid19.who.int/>.
- World Health Organization. (2020m). *WHO Director-General's remarks at the media briefing on 2019-nCoV on 11 February 2020*. Retrieved May 2, 2020, from <https://www.who.int/dg/speeches/detail/who-director-general-s-remarks-at-the-media-briefing-on-2019-ncov-on-11-february-2020>.
- World Health Organization. (2020n). *WHO Timeline - COVID-19*. Retrieved May 2, 2020, from <https://www.who.int/news-room/detail/27-04-2020-who-timeline---covid-19>.
- Wongboonsin, P., Keeratipongpaiboon, T., & Wongboonsin, K. (2019). Changes in family composition and care relations in the Kingdom of Thailand. In P. Wongboonsin & J. Tan (Eds.). *Care relations in South-east Asia: The family and beyond* (Chapter 3, pp. 99–154). Brill.
- YouGov (2020). *Thais most likely to wear facemasks in ASEAN*. Retrieved November 17, 2020, from <https://th.yougov.com/en-th/news/2020/05/19/thais-most-likely-wear-facemasks-asean/>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.