



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Contents lists available at ScienceDirect

Personality and Individual Differences

journal homepage: www.elsevier.com/locate/paid

How did socio-demographic status and personal attributes influence compliance to COVID-19 preventive behaviours during the early outbreak in Japan? Lessons for pandemic management

Shahadat Uddin^a, Tasadduq Imam^{b,*}, Matloob Khushi^c, Arif Khan^d, Mohammad Ali Moni^e

^a School of Project Management, Faculty of Engineering, The University of Sydney, Room 5050, ABS Building, Darlington, NSW 2008, Australia

^b School of Business and Law, CQUniversity (Melbourne), VIC 3000, L4, 120 Spencer Street, Melbourne, VIC 3000, Australia

^c School of Computer Science, J12, Computer Science Building, The University of Sydney, Australia

^d Faculty of Engineering, J12, Computer Science Building, The University of Sydney, Australia

^e WHO Collaborating Centre on eHealth, School of Public Health and Community Medicine, Faculty of Medicine, The University of New South Wales, Sydney, NSW 2052, Australia

ARTICLE INFO

Keywords:

COVID-19
Social distancing
Personal protection behaviour
Social responsibility awareness
Socio-demographic status
Personal attribute
Quality of life

ABSTRACT

This study focuses on how socio-demographic status and personal attributes influence self-protective behaviours during a pandemic, with protection behaviours being assessed through three perspectives – social distancing, personal protection behaviour and social responsibility awareness. The research considers a publicly available and recently collected dataset on Japanese citizens during the COVID-19 early outbreak and utilises a data analysis framework combining Classification and Regression Tree (CART), a data mining approach, and regression analysis to gain deep insights. The analysis reveals Socio-demographic attributes – sex, marital family status and having children – as having played an influential role in Japanese citizens' abiding by the COVID-19 protection behaviours. Especially women with children are noted as more conscious than their male counterparts. Work status also appears to have some impact concerning social distancing. Trust in government also appears as a significant factor. The analysis further identifies smoking behaviour as a factor characterising subjective prevention actions with non-smokers or less-frequent smokers being more compliant to the protection behaviours. Overall, the findings imply the need of public policy campaigning to account for variations in protection behaviour due to socio-demographic and personal attributes during pandemics and national emergencies.

1. Introduction

There have been several outbreaks of coronavirus in the past, namely SARS (severe acute respiratory syndrome) and MERS (Middle East respiratory syndrome) (Liu et al., 2020). The latest outbreak known as COVID-19 is caused by a novel strain of coronavirus (SARS-CoV-2) and was first reported in Wuhan, China in late December 2019 (Burki, 2020). The outbreak was first declared as Public Health Emergency of International Concern and later as a pandemic in March 2020 by the World Health Organization (WHO) (WHO, 2020a). COVID-19 infection could

be asymptomatic or could show mild to severe pneumonia like symptoms (Kannan et al., 2020). The most common symptoms include fever, dry cough, and tiredness, while the less common symptoms include headache, chest pain, loss of speech, difficulty in breathing, sore throat, diarrhoea, conjunctivitis or other aches and pains (Mustafa & A Selim, 2020). A noticeable feature of the disease is its highly contagious nature and, while the virus outbreak is still under investigation, there are views that the spread can happen person-to-person due to close contacts, airborne particles, and contact with contaminated surfaces (Setti et al., 2020). Additionally, the high risk group of the COVID-19 is different

* Corresponding author.

E-mail addresses: shahadat.uddin@sydney.edu.au (S. Uddin), t.imam@cqu.edu.au (T. Imam), matloob.khushi@sydney.edu.au (M. Khushi), arif.khan@sydney.edu.au (A. Khan), m.moni@unsw.edu.au (M.A. Moni).

URL: <https://sydney.edu.au/engineering/about/our-people/academic-staff/shahadat-uddin.html> (S. Uddin), <https://staff-profiles.cqu.edu.au/home/view/134> (T. Imam), <https://www.sydney.edu.au/engineering/about/our-people/academic-staff/matloob-khushi.html> (M. Khushi), <https://sphcm.med.unsw.edu.au/content/who-collaborating-centre-ehealth> (M.A. Moni).

<https://doi.org/10.1016/j.paid.2021.110692>

Received 7 September 2020; Received in revised form 23 December 2020; Accepted 22 January 2021

Available online 27 January 2021

0191-8869/© 2021 Elsevier Ltd. All rights reserved.

from recent pandemics with the current strain particularly deadly for older individuals (Tang et al., 2020). Thus, while the world has confronted multiple epidemics in recent times and different countries have had several preparedness plans in place following their past experiences (Henry, 2019; Itzwerth et al., 2018; Jennings, 2009), the current pandemic has come as an unprecedented context in many countries. Indeed, even after different public health measures instituted by the governments, the virus has continued to spread across countries. As of 13th December 2020, WHO's weekly update identifies over 70 million positive cases globally and close to 1.6 million deaths from the disease (WHO, 2020b).

Despite uniqueness of the COVID-19 context, countries have adopted some common measures. These include staying at home, social distancing, contact tracing, wearing masks, washing hands, and using hand sanitiser. There have also been public and media communications to increase public awareness. However, individual behaviour during an epidemic can be influenced by various factors including trust in the Government's advice and subjective perceptions (Freimuth et al., 2014; Teasdale et al., 2012). Additionally, individual health protection behaviour during an influenza pandemic can be characterised by social contexts (Chuang et al., 2015).

Thus, with many COVID-19 cases being asymptomatic, leading to people carrying on with their daily life as normal and thereby acting as virus carriers and threat especially to elderly people, it is interesting to explore the public behaviour and their understanding of the risk associated with COVID-19 spread. Research in this context notes the influence of education level and subjective trust of the government as factors that influence the perception of COVID-19 as a conspiracy (Georgiou et al., 2020). Another work finds subjective perceptions of COVID-19 crisis rather than personality as more influential in explaining compliance to Government recommendations in Poland (Zajenkowski et al., 2020). Further, research identifies links between personality and COVID-19 behaviour (Qian & Yahara, 2020), and the positive influence of lifestyle, culture, and healthcare in flattening the COVID-19 infection curve (Tashiro & Shaw, 2020) in Japan.

Despite such explorations, it is still instructive to assess how the pandemic was perceived at subjective level in the early stages especially with the unprecedentedness of the disease. Lessons learned from such undertaking can guide policy makers to prepare for further outbreaks or a new disaster. In this study, we analyse data from Japan to understand the level of people's compliance to the government suggested preventive guidelines during the early COVID-19 outbreak and the impacts socio-demographic status and personal attributes had on their behaviours.

An outstanding feature of this work is going beyond the use of traditional statistical approaches like multiple linear regression, as often exploited in existing studies. Such approaches often assume different conditions persisting in data that may not always hold. In this research, we apply the Classification and Regression Tree (CART), a non-parametric data mining algorithm that are not restricted by such assumptions and can model complex non-linear relationship (Breiman, 2017; Yoo et al., 2018). We combine the flexibility of CART with linear regression for a robust analysis – an approach, to the best of our knowledge, is yet to be employed in assessing pandemic situations.

2. Literature review

Implementing large scale lockdown and social distancing was difficult for highly populated countries such as Indonesia. However, modelling shows early interventions have saved thousands of lives (Djalante et al., 2020). A WeChat survey of 3083 participants concludes that people living in cities have a better understanding of the disease compared to those in rural areas (Zhan et al., 2020). The study also demonstrates males (OR: 0.544, 95% CI: 0.440–0.673), younger adults (1.844, 1.466–2.320), and subjects with higher education (2.200, 1.780–2.718) exhibited better behaviour when protective measures such as washing hands, wearing masks or exercising were given (Zhan

et al., 2020). This finding indicates that health education should be strengthened among adults living in rural areas where access to medical services is limited. A survey of 976 university academics and students of the Iraqi Kurdistan region revealed that the perceived risk of contracting COVID-19 infection, serious illness and death (26.9%, 29.7%, and 41.7%, respectively) is low in the participants (Shabu et al., 2020). Similarly, a Bangladeshi cross sectional study of 320 adults identified people living in the urban area ($p < 0.01$), high education ($p < 0.01$), rich ($p < 0.01$) and joint family ($p < 0.01$) had the most contributions to good practice (Wadood et al., 2020).

The reception of public plays a key role in controlling an infectious disease such as COVID-19. The best example of this is Singapore where the disease was very well controlled compared to other countries. The key factors of this success come from the experience that the Singapore government and its public have experienced during earlier outbreaks. In 2000, 2003 and 2009 they experienced hand-foot-mouth disease, SARS and H1N1 outbreaks respectively (Lai & Tan, 2015). They learnt that response to such a pandemic has to be well-coordinated and multi-stakeholder approach including participation from public.

The issues that were caused by COVID-19 are unparalleled and unprecedented. This generation has not seen anything like it. It has been argued that the road to full recovery will be very long and multiple waves of infection are predicted (Fakhruddin et al., 2020). Hence, as the pandemic continues to progress across countries, governments face myriad of multi-dimensional challenges including surge in demand for public health system, economical and mental health issues of the public. Therefore, this study analyses the effect of personal attributes, age, social and demographics on public behaviour and their perception about the spread of the disease and means to control it.

3. Materials

3.1. Data source

This study used an open-source research dataset that was made available by Yamamoto (Yamamoto, 2020). The data was collected through an online survey in Japan in late March 2020 (Yamamoto, 2020). The provided dataset consists of the survey responses by 11,342 Japanese recruited through a cross-sectional approach from a pool of 1.2 million registered individuals. The researchers involved in data collection adopted non-probabilistic quota sampling to match the distribution of demographic status (gender, age, and work status) in the respective national statistics (Yamamoto, 2020).

3.2. Variable quantification

This study considered responses to one survey question in particular to quantify the subjective attitudes of Japanese citizens during the early COVID-19 outbreak - "Have you ever conducted anything to prevent novel coronavirus infections or outbreaks?" (Yamamoto, 2020). For the response of this question, each respondent needed to select one of the five options ('Very true', 'True', 'Neither', 'Not true' and 'Not at all') against each of 21 activities. We quantify the responses innovatively in terms of three dependent variables. Notably, a recent article on the same dataset focused on the same question to characterise the Japanese citizens' behaviour across three government recommended measures (Muto et al., 2020). We consider the same question and items are labelled exactly as in the research dataset (Isamu Yamamoto, 2020; Muto et al., 2020). However, our approach in assessment of pandemic protection behaviour is different. The published work examines the citizens' behaviour and attitudes through a common umbrella of social distancing and personal etiquettes and focuses mainly on 12 items separately within the same question (Muto et al., 2020). We, by contrast, utilise all 21 items and group these activities across three individualistic protective behaviours – 'social distancing', 'personal protection behaviour' and 'social responsibility behaviour'.

Table 1
List of activities used for different dependent variables of this study.

Social distancing	Personal protection behaviour	Social responsibility awareness
1. Avoid a poorly ventilated closed space	1. Undertake frequent handwashing	1. Stockpile surgical-style mask
2. Avoid large gatherings	2. Undertake cough etiquette	2. Stockpile food, toilet paper, tissue paper, etc.
3. Avoid conversations or shouting in proximity	3. Disinfect things around	3. Avoid contact with younger people
4. Avoid places where items 1–3 above overlap	4. Avoid going out when you have a cold	4. Avoid contact with the older people
5. Do not go to dinner with friends	5. Avoid going to clinic even when having a cold symptom	5. Get enough rest and sleep
6. Do not go to mass gatherings	6. Prepare consultation and transportation methods for when you feel ill	6. Eat a nutritious diet
7. Participate in virtual events using online tools	7. Always wear a surgical-style mask when going out	7. Do exercise such as a jogging or exercise using DVD

Table 1 details the 21 activities grouped across the three dependent variables. We interpret the meaning and impacts of the respective behaviour to form the grouping. For example, stockpiling masks, medication, food, or other essential items can be interpreted as socially irresponsible behaviour especially if the general or vulnerable populace misses out on these goods. While eating nutritious diet or getting rest and exercise may be interpreted as personal behaviour, we consider the social side of the impacts of such behaviour and the need for a healthy society – thus, these items are also grouped under social responsibility. Similarly, avoiding closed but ill-ventilated spaces, not speaking in close proximity, avoiding mass gathering, and related practices are grouped under social distancing behaviours; while taking measures to disinfect surroundings, wearing masks, avoiding travels for any reason when having cold, and related practices are deemed as personal protection behaviours.

This study considers two broad categories of independent variables interpreted from the original dataset (Yamamoto, 2020)– personal attributes and socio-demographic attributes. For the personal attribute category, four variables are used: (i) Having children younger than junior high school age; (ii) Smoking frequency; (iii) Drinking frequency; and (iv) Trust in Government policy. The first of these attributes has a

binary response (yes/no), smoking frequency has four choices ('Every day', 'Sometimes', 'Used to smoke but do not now' and 'Never smoked'); whereas, drinking frequency offers six choices ('Never drink', 'I used to drink but I quit', 'Few times per month', '1–2 times per month', '3–6 times per month' and 'Every day') (Yamamoto, 2020; Muto et al., 2020). The last one has been derived from the responses of three Government policies. They are - the Government should (a) allow mass gatherings now; (b) continue to request self-restraints of mass gatherings; and (c) limit movement in addition to mass gatherings (Muto et al., 2020). For each policy, there are five possible responses ('Agree', 'Relatively agree', 'Neither', 'Relatively disagree', 'Disagree' and 'Do not know'). An aggregate of scores for the three policies reflect values of the trust in Government policy. Notably, having children, and smoking and drinking frequency are related to quality of lifestyle an individual has chosen to lead, and can also in turn influence their compliance to any policies and rules – considering these personal attributes, hence, makes sense. Additionally, six socio-demographic variables are considered in this study. They are – (i) Age (with five choices, 20s, 30s, 40s, 50s and 60–64); (ii) Gender (female/male); (iii) Marital status (Married or Not married); (iv) Education (university/college graduate or not); (v) Work status (four choices – 'Regular employee', 'Non-regular employee', 'Self-employed or others' and 'Not working'); and (vi) Household annual income (10 ascending choices - 'less than 2,000K Japanese yen', seven choices between 2000 K and 19,999 K Japanese yen, 'More than 20,000K Japanese yen' and 'Do not know') (Yamamoto, 2020). As indicated in earlier sections, existing literature has focused on various demographic attributes like education, age, and gender in characterising public response to government actions during pandemics. Additionally, smoking has been noted as a risk factor of COVID-19 in multiple studies (Patanavanich & Glantz, 2020), while drinking and having young children (i.e., choosing to have children) can also impact the way individuals behave in different situations. Thus, our choice of personal and socio-demographic attributes variables is guided by the present literature and knowledge.

3.3. Data analysis framework

Fig. 1 shows the data analysis framework followed in this research. To convert the categorical responses into a numerical one, the RIDIT analysis has been applied to the dependent variables of this research. After this, CART, a state-of-the-art data mining approach, is applied to identify important independent variables. Finally, regression is used to model the dependent variables using the independent variables. Here is a brief about each of these three methods.

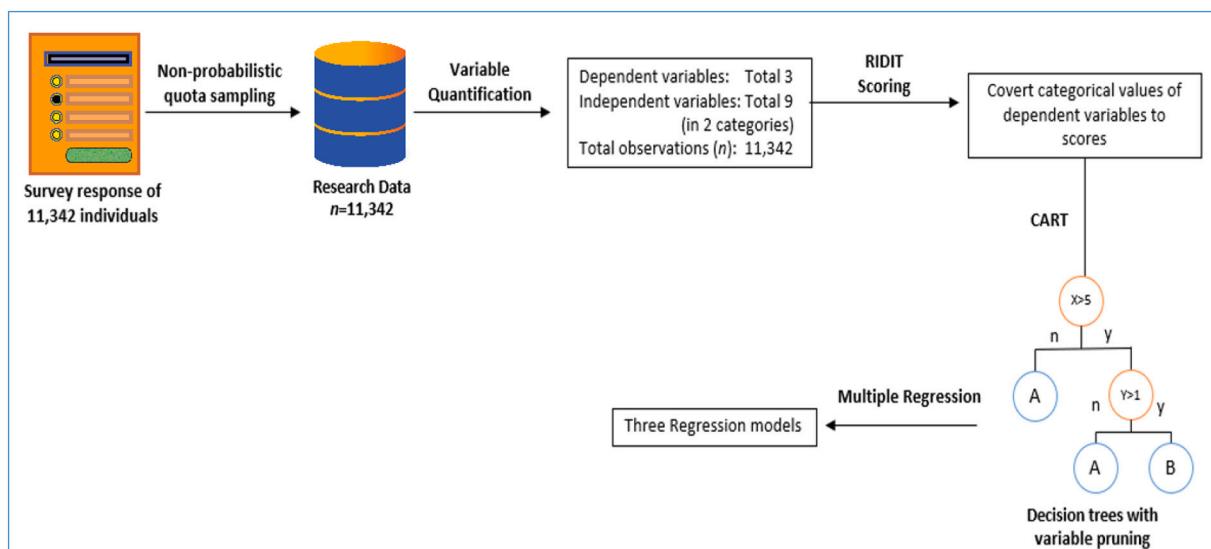


Fig. 1. Data analysis framework followed in this research.

3.3.1. RIDIT scoring

RIDIT is a statistical method for analysing subjectively categorized response variable, i.e., ordered qualitative measurements often called as ordinal data (Bross, 1958a; Flora Jr, 2014). For example, if the response variable has subjective scales with ordered categories, such as – ‘never smoked’, ‘smokes sometimes’, ‘smokes every day’ or ‘minor pain’, ‘moderate pain’, ‘severe pain’ etc., then they may not be adequately analysed by chi-square- or *t*-test-based statistical methods. RIDIT analysis is often used to transform the response variable by allocating scores relative to the identified (or reference) distribution of the data.

The term RIDIT stands for “relative to an identified distribution integral transformation” and this was termed by Bross for its analogy with other statistical transformation methods such as *probit* and *logit* (Bross, 1958b). Essentially, the process of RIDIT analysis can be summarised in the following steps:

Step 1: Select the categorical response variable whose response values are categorical, subjective and in order. For example, for the survey question – “How often do you smoke?”, the response values are – ‘never’, ‘sometimes’ and ‘daily’. Respondents giving the same response is called a group.

Step 2: Choose references or identified distribution data. Based on this set, the non-reference data will be baselined. For this reference set, each group is assigned a score or weight. Thus, each of the response values or categories is transformed into a score. Let us assume, a response variable *X* has *n* ordered responses in the reference set marked as $x_1, x_2, x_3, \dots, x_n$. n_i denotes the total number of responses for x_i ($n = n_1 + n_2 + \dots + n_i + \dots + n_n$). Then, RIDIT R_i for response x_i will be,

$$R_i = \frac{0.5n_i + \sum_{j=1}^{i-1} n_j}{n} = Prob(x_i) + \sum_{j=1}^{i-1} Prob(x_j)$$

Here $Prob(x_i)$ is the probability of choosing response x_i in the reference set.

Step 3: The score is assigned to each group in the non-reference or test set and a mean RIDIT score is calculated for each group. The score relates to the probability that a member of the group differs (i.e., have higher score) from a member of the reference population (Bross, 1958a).

For this research, we used RIDIT scores as described in step 2 above to transform the categorical responses into numerical scores. The reference set consisted of all the 11,342 respondents sampled from the whole dataset.

3.3.2. Classification and Regression Tree

CART (short for Classification and Regression Trees) is an umbrella term (Breiman et al., 1984) for referring to two types of decision trees – classification tree and regression tree. *Decision tree* or *decision tree learning* is a widely used predictive modelling approach used in machine learning, data mining and statistics. Given a dataset with multiple observation of an outcome variable with respect to one or more input (i.e., independent) variable, the decision tree essentially generates a binary tree to predict the outcome variable based on the input variable. The binary tree starts at the single root node and keeps splitting in every non-leaf node into two branches based on a logical comparison of an input variable, e.g., “if income > 800\$: Yes/No”, “if sex is male: Yes/No”, etc. The leaf nodes indicate an outcome as the prediction result of the decision tree. Thus, the outcome of a set of input variable values (of future observation) can be predicted by simply following the decision tree logic and reaching towards a particular leaf-node. Now, if the outcome variable has discrete values, i.e., discrete classes such as sex (male, female), disease outcome (benign, malignant), etc., then it is a classification tree. And, if the outcome variable is a continuous variable (price, age in years etc.) then it is a regression tree. CART is a modern term that encompasses a range of algorithms to generate both types of trees.

In our research, we generated three CART trees (regression trees as

the outcome is a continuous number) based on three outcome variables – social distance, personal protection behaviour, social responsibility awareness. In each case, the ten socio-demographic and personal attributes outlined earlier are used as independent or input variables. Applying CART can generate trees with many levels and cause overfitting to data, and pruning is applied to enhance generalisation of the outcomes (Hayes et al., 2015; Krzywinski & Altman, 2017). The pruning step, further, to improving generalisation, also identifies input variables most influential in characterising the dependent variables. The variables that are finally present in the resultant trees are used in the next step of the analysis.

For CART analyses, we used IBM SPSS Statistics (version 24) software tool. We further used the ‘Gini impurity’ measure to split off the category. This measure is widely used to split trees and computationally inexpensive to implement (Yitzhaki & Schechtman, 2012). It focuses to maximise the homogeneity of child nodes with respect to the classification performance of the underlying dependent variable (Yitzhaki & Schechtman, 2012). The variable chosen for split is known as ‘primary splitter’; and, at each split, other variables that can replace primary splitters including when values for primary splitters are missing are considered and these other predictors are known as ‘surrogate splitters’ or surrogates (Steinberg, 2009; Tran et al., 2008; Yohannes & Webb, 1999). Since we have ten independent variables, a value of nine (which is one less than the maximum number of independent variables) has been used as the maximum number of surrogates in the SPSS tree growing mechanism. This ensures consideration of most possibilities in the event of a missing value for the primary splitter. We further use the 10-fold cross-validation technique for the validation purpose.

3.3.3. Multiple regression and t-test

Multiple regression is used to explain the relationship between one continuous dependent variable with two or more independent variables. It is an extension of simple linear regression. The result of the multiple regression indicates whether the outcome or dependent variable can be predicted from the independent variables, the accuracy of the prediction, model fits and statistical significance. Application of multiple regression requires several assumptions (Laerd Statistics, 2013) including – dependent variables should be continuous and linearly dependent on independent variables, at least two or more independent variables – which can be either continuous or categorical and should not be highly correlated, independence among the observations, homoscedasticity in the data and normally-distributed residuals.

In this research, we combine the generalisability of CART and interpretability of multiple regression to gain insights – an innovative approach used in this work. More precisely, after applying CART for each of the dependent variables (social distance, personal protection behaviour, social responsibility awareness) and identifying the influential variables for the respective models in the first step, we generated three multiple regression models for each of these dependent variables with the combined set of identified independent variables as predictors. This ensures that variables that are not highly influential in characterising dependent variables are excluded, especially to reduce potential biasness in the final models. We further used one sample *t*-test (Field, 2009) to determine if there is a significant difference between two groups in following COVID-19 preventive practices.

Overall, we hypothesise that demographic attributes and individual status of respondents in Japan characterised their compliance across the three categories of pandemic protection behaviours during early stage of the pandemic. We further assume that a selected set of these variables has been highly influential particularly. Through our adopted approach, we aim at identifying these influential individuality related variables and to what extent they characterised the pandemic protection behaviour. The following section outlines the results.

Table 2
RIDIT score for the five responses against each of 21 activities of the considered survey question.

Response	Frequency	RIDIT score
Not at all	16,146	0.07
Not true	28,152	0.19
Neither	57,141	0.43
True	75,809	0.74
Very true	60,934	1.00

4. Results

Table 2 shows the corresponding RIDIT scores for the five responses against each of 21 activities. The survey respondents reported these activities in response to the considered survey question. The responses of these activities were considered to quantify the three dependent variables (*Social distance*, *Personal protection behaviour*, *Social responsibility awareness*).

Figs. 2-4 illustrate the CART outcomes for the three dependent variables of this study. We used SPSS (Field, 2009) with default settings for generating these three trees. Four attributes (*Age*, *Income*, *Drinking frequency*, and *Education*) were not found in any of the three CART models – an indication that these variables may have low influence on the dependent variables. For this reason, these four variables were not considered in the multi regression model considered in the next stage.

There are six attributes that were found in at least one trees. Table 3 provides a statistical summary of the five of them – sex, marital status, work status, child (i.e. having children), and smoking frequency. The sixth one (*Trust in Government*) is a derived measure, which can take values ranging from 0 to 15; where 0 indicates lowest trust about the three Government policies and 15 indicates the highest trust. As notable, the dataset contains a balance of male and female respondents.

Table 4 shows the three regression models developed for each of three dependent variables of this study. Three attributes (*Trust in Government*, *Marital status*, and *Sex*) have significant positive impact on each of the three dependent variables (*Social distance*, *Personal protection*

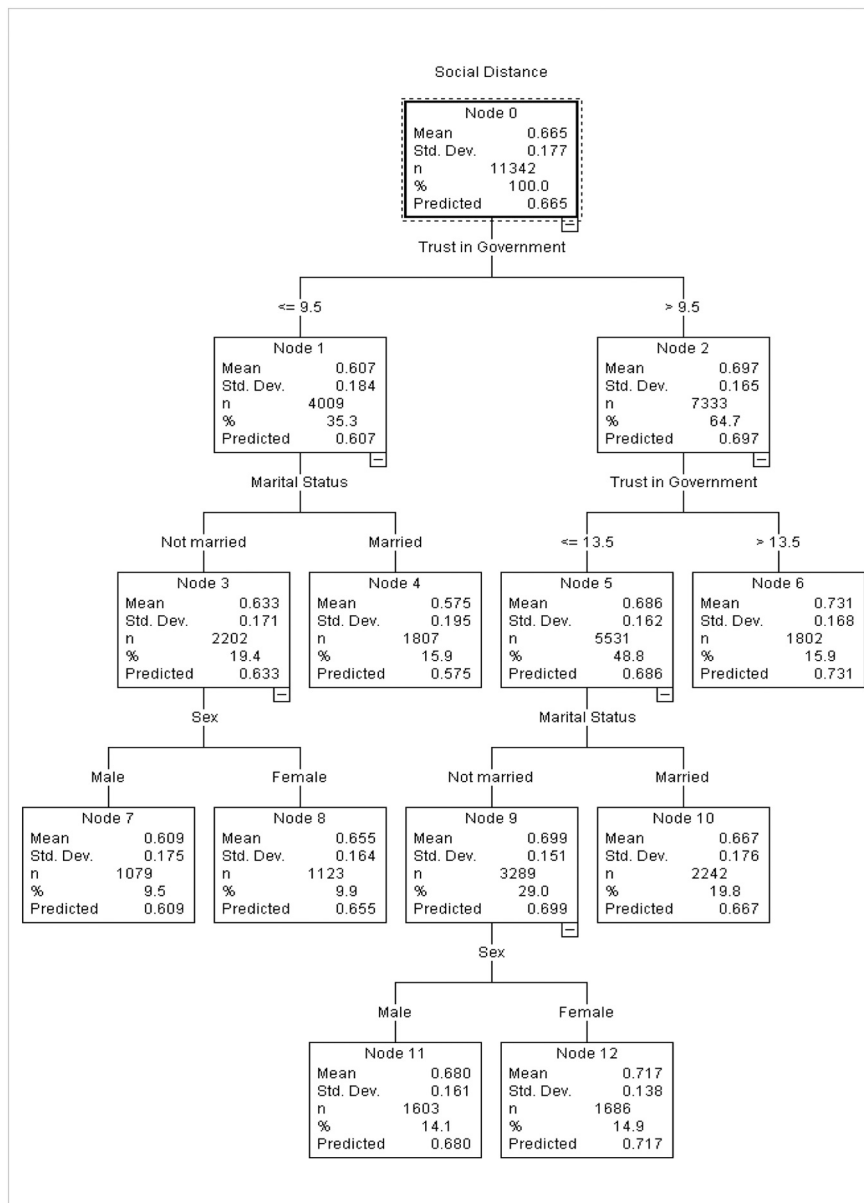


Fig. 2. CART outcome tree for the ‘*Social distance*’ dependent variable. Out of ten, three attributes (*Trust in Government*, *Marital status*, and *Sex*) were included as nodes in this tree.

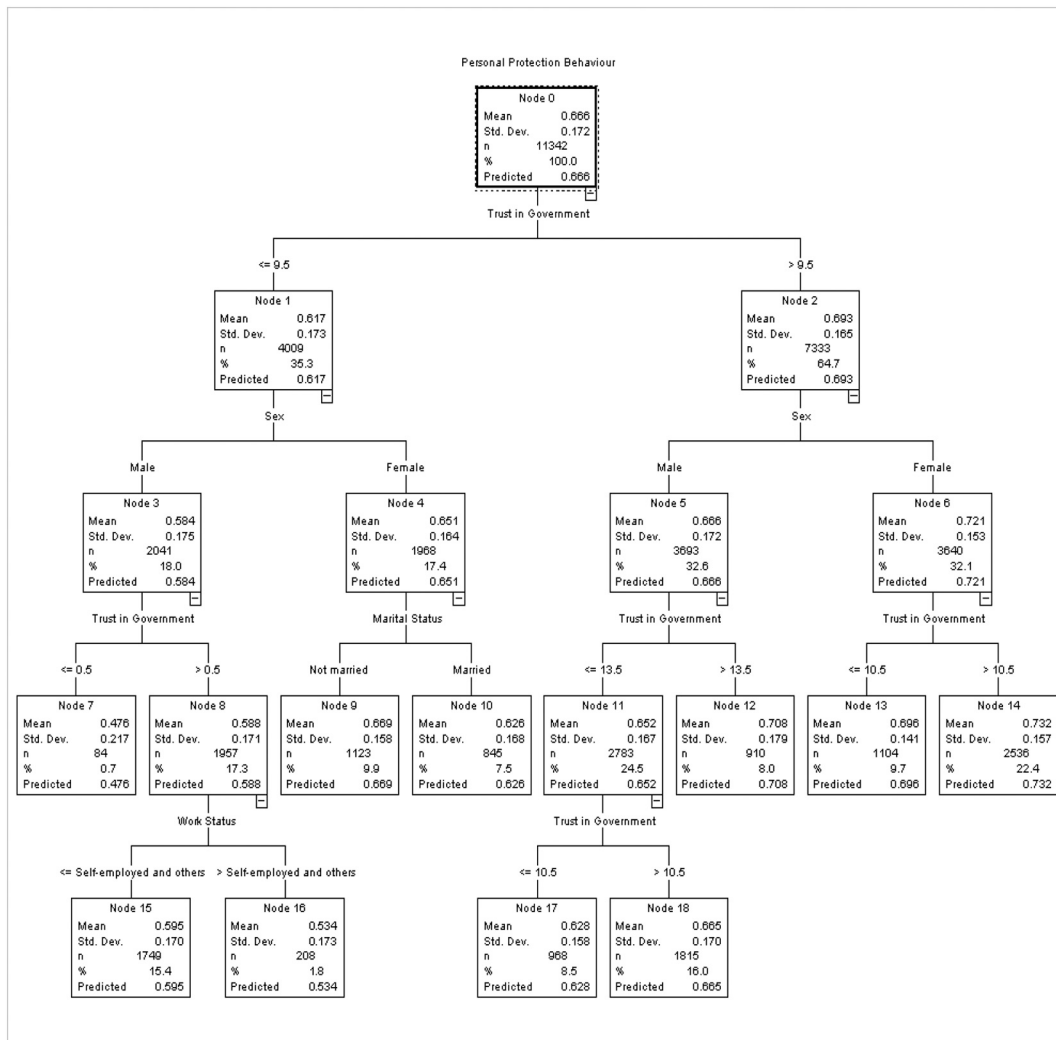


Fig. 3. CART outcome tree for the ‘Personal protection behaviour’ dependent variable. Out of ten, four attributes (Trust in Government, Sex, Marital status, and Work status) were included as nodes in this tree.

behaviour and Social responsibility awareness). Two attributes (Child and Smoking frequency) have negative impact on each of the three dependent variables. The ‘work status’ attribute has significant positive impact on only the ‘social distance’ measure. We also considered VIF (variance inflation factor), which is a statistical measure that can detect the presence of multicollinearity in a regression analysis. In regression analyses, multicollinearity is a phenomenon which is used to explore the presence of any linear relations among predictor variables (Field, 2009). In general, a VIF value of >10 indicates a high presence of multicollinearity. As revealed by the VIF (variance inflation factor) values indicated in Table 4, multicollinearity is not present in our regression analysis.

As revealed in Table 4, females are more likely to follow COVID-19 preventive practices than males. During the regression analyses, the numerical values of ‘1’ and ‘2’ are used for males and females, respectively. So, a positive coefficient in Table 4 for the ‘sex’ attribute indicates a higher adherence of females in following the preventive practices. As mentioned earlier, the ‘child’ attribute has also a negative impact on each of the three preventive measures of this study. We further categorise the survey respondents based on their ‘sex’ and ‘child’ information and, using *t-test*, explore the difference in abiding by the COVID-19 preventive practices among those categories. As outlined in Table 5, regardless of having a child, females tend to adhere preventive practices more than males.

The R^2 values in Table 4 range from 6.5% to 8.8%. Notably, R^2 can have different meanings and it may not necessarily preserve its common interpretation of goodness of fit when used for test data or when the training model is non-linear (Alexander et al., 2015). Research, further, reflects that high R^2 values emphasise on individual fit and can lead to faulty decision making when the focus is on model fitness at a group level (Rose & McGuire, 2019). Our use of linear model is not for predictive purpose. Rather, we use the non-linear model facility of CART to identify the important variables from the limited variables available in the source dataset and then use multiple linear regression to note any influence of independent variables on the dependent variables. Also, for the three resultant CART models, we note that the value for the ‘risk’ measure varied from 0.027 to 0.029. In tree classification, the ‘risk’ quantifies the proportion of cases incorrectly classified by the underlying tree. Thus, low values of risk reflect the robustness of the CART models, and which, adapting existing findings from the present literature (Rose & McGuire, 2019), arguably complement any low R^2 values for the regression stage.

Finally, we assess the reliability of findings. As outlined earlier, this research considered three constructs (social distancing, personal protection and social responsibility awareness). In quantifying these three constructs, we used 21 different survey items from the considered survey question and categorized these into three equal-sized groups (i.e., 7 per group) corresponding to each of the three constructs. The groupings are

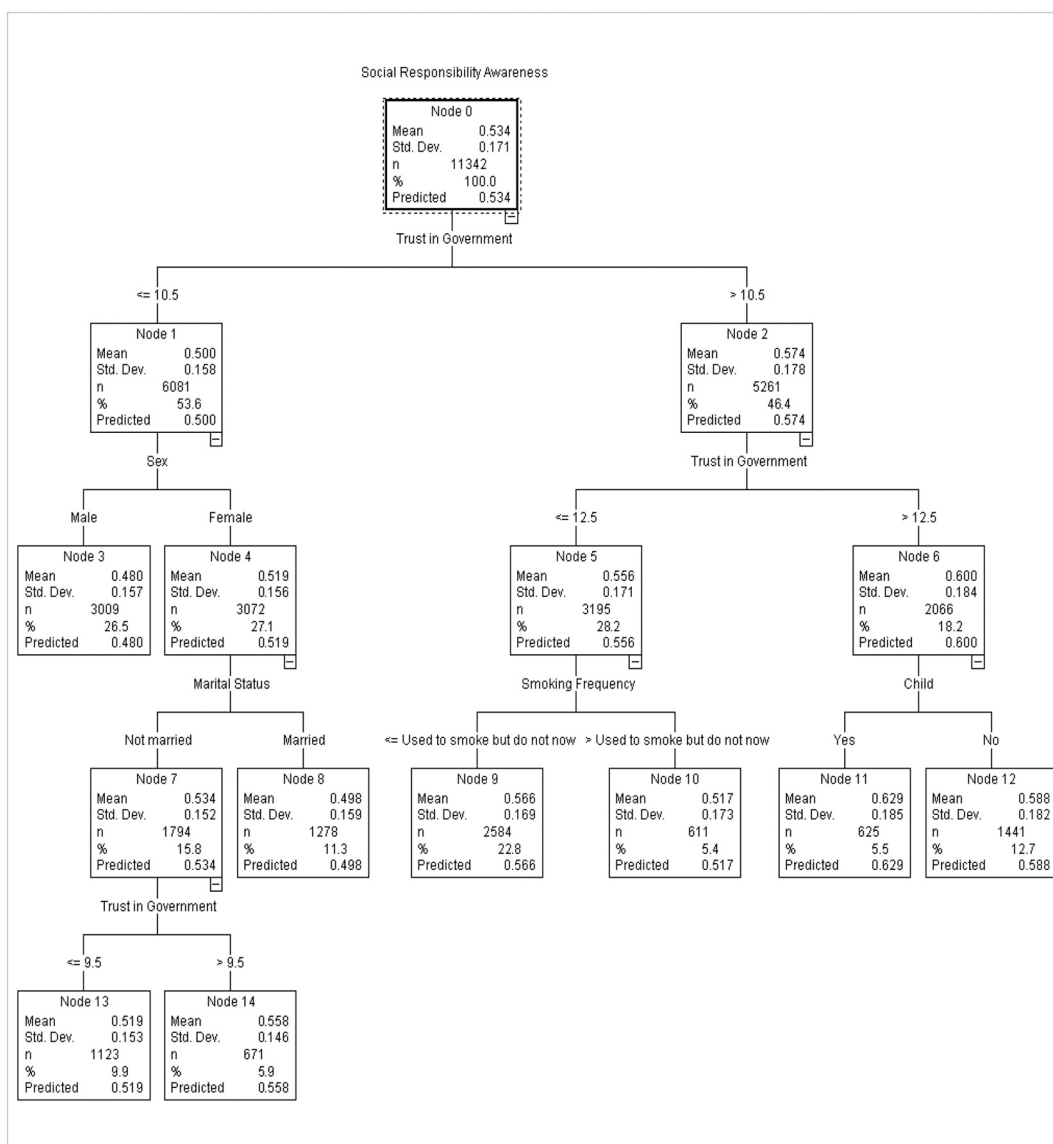


Fig. 4. CART outcome tree for the ‘Social responsibility awareness’ dependent variable. Out of ten, five attributes (Trust to Government, Sex, Marital status, Smoking frequency, and Child) were included as nodes in this tree.

statistically reliable, as revealed from the corresponding Cronbach’s alpha values outlined in Table 6. Cronbach’s alpha is a measure of internal consistency and reliability used to explore ‘how closely a set of items are related’ when they are used to quantify a construct (Cronbach, 1951). In general, a value of 0.70 or more is considered as ‘acceptable’ in most social science research situations.

5. Discussion

Overall, it appears that the demographic attributes like sex, marital and family status have played an influential role in abiding by COVID-19 protection behaviours in Japan. Research notes that, during the H1N1 pandemic, sex and marital status differently characterised intentions to get vaccinated among the Italian health workers (La Torre et al., 2009). Another research highlights that males were more adoptive of protection behaviours during the H1N1 pandemic in Saudi Arabia (Balkhy et al., 2010), while females were noted as more compliant to H1N1 prevention measures in Hong Kong (Lau et al., 2010). A research further notes that the reason for vaccination following the H1N1 pandemic varied between males and females in Japan (Iwasa & Wada, 2013). The COVID-19 outbreak, although a different disease, has some similarities to the

past pandemic and is also caused by a virus. The findings from this study therefore reaffirm the differences in pandemic protection behaviour in population due to demographic attributes like sex. Further, research highlights that women tended to show more infection protection behaviour than men across countries during the previous pandemics (Bish & Michie, 2010). Thus, this study’s finding of Japanese women being more compliant to COVID-19 protection measures matches this attitude.

The analysis also reveals ‘work status’ as a contributing factor concerning ‘social distancing’. In a pandemic situation, work status of individuals can impact their possibility of contacts with other individuals and their ability to absorb the economic impact of any restrictions arising from public health advice. These, in turn, shape individual perceptions of any social distancing measure and personal protection behaviour. Such has also been the case in Japan. In the considered dataset, a value of 1 is assigned to ‘regular employee’ and higher values indicate individuals who are occasionally employed or self-employed or unemployed. A positive significant coefficient for work status, hence, suggests that non-regular employees were more compliant to social distancing – potentially, which also relates to regular employees finding it difficult to maintain social distancing. Some research has noted work

Table 3
Statistics of the variables chosen following the CART training. Labels are as in original dataset (Yamamoto, 2020).

	All	Female	Male
Total respondents	11,342	5608 (49.44%)	5734 (50.56%)
Marital status			
1. Married	4722 (41.63%)	2229 (39.75%)	2493 (43.48%)
2. Not married	6620 (58.37%)	3379 (60.25%)	3241 (56.52%)
Having children younger than junior high school age			
1. No	2972 (26.20%)	1479 (26.38%)	1493 (26.04%)
2. Yes	8370 (73.80%)	4129 (73.62%)	4241 (73.96%)
Smoking frequency			
1. Never smoked	6748 (59.50%)	4120 (73.47%)	2628 (45.83%)
2. Used to smoke but not now	2188 (19.29%)	773 (13.78%)	1415 (24.68%)
3. Sometimes	270 (2.38%)	82 (1.46%)	188 (3.28%)
4. Everyday	2136 (18.83%)	633 (11.29%)	1503 (26.21%)
Work status			
1. Regular employee	5817 (51.29%)	1831 (32.65%)	3986 (69.52%)
2. Non-regular employee	2865 (25.26%)	2132 (38.02%)	733 (12.78%)
3. Self-employed and others	660 (5.82%)	238 (4.24%)	422 (7.36%)
4. Not working	2000 (17.63%)	1407 (25.09%)	593 (10.34%)

status of individuals having an impact on their compliance to pandemic protection measure (Bish & Michie, 2010). Our finding thus affirms work status as a factor to consider during pandemic management.

Also, research suggests that humans well consider sociality when facing the threat of an infectious disease (Curtis, 2014). There have further been some evidences of marital status influencing pandemic protection behaviour in Hong Kong, even though the impact of marital

Table 4
Regression outcome for the three dependent variables considered in this study.

	Model one Dependent variable: <i>Social distance</i>				Model two Dependent variable: <i>Personal protection behaviour</i>				Model three Dependent variable: <i>Social responsibility awareness</i>			
	Beta	Sig.	VIF	R ²	Beta	Sig.	VIF	R ²	Beta	Sig.	VIF	R ²
(Constant)	0.430	0.000		0.085	0.430	0.000		0.088	0.377	0.000		0.065
Child	-0.027	0.006	1.216		-0.030	0.002	1.216		-0.038	0.000	1.216	
Trust in Government	0.240	0.000	1.009		0.216	0.000	1.009		0.221	0.000	1.009	
Marital status	0.094	0.000	1.217		0.073	0.000	1.217		0.038	0.000	1.217	
Sex	0.055	0.000	1.152		0.168	0.000	1.152		0.059	0.000	1.152	
Smoking frequency	-0.040	0.000	1.076		-0.023	0.015	1.076		-0.064	0.000	1.076	
Work status	0.070	0.000	1.097		-0.014	0.129	1.097		0.018	0.055	1.097	

Table 5
t-test results to explore the differences in abiding by the COVID-19 preventive practices between different subgroups of the survey respondents.

Test no.	Independent variable	Group	N	Mean	STD	t-value	Sig.
1	Social distance	Female having child	4129	0.670	0.172	6.702	0.000
		Male having child	4241	0.644	0.187		
2	Social distance	Female not having child	1479	0.711	0.154	7.521	0.000
		Male not having child	1493	0.665	0.178		
3	Personal protection behaviour	Female having child	4129	0.690	0.162	16.645	0.000
		Male having child	4241	0.628	0.178		
4	Personal protection behaviour	Female not having child	1479	0.712	0.156	8.675	0.000
		Male not having child	1493	0.661	0.174		
5	Social responsibility awareness	Female having child	4129	0.543	0.165	8.226	0.000
		Male having child	4241	0.512	0.174		
6	Social responsibility awareness	Female not having child	1479	0.563	0.166	3.419	0.001
		Male not having child	1493	0.542	0.178		

status on infection protection behaviour in general had been inconclusive (Bish & Michie, 2010). Thus, the finding that marital status has played a role in Japanese populaces' adherence to COVID-19 protection behaviour is, hence, interesting, and potentially relates to their concern about families and the inherent sociality.

Among the personal attributes, this research particularly notes 'smoking frequency' as having a negative impact on each of the three COVID-19 preventive behaviours. As outlined in Table 2, the possible values for the 'smoking frequency' attribute have been placed in ascending order (i.e., '1' for 'never smoked', '2' for 'used to smoked but not now', '3' for 'sometimes' and '4' for 'everyday'). Thus, a negative coefficient for the 'smoking frequency' attribute indicates that individuals who are non-smoker or smoke less frequently are more likely to adhere COVID-19 preventive measures. Research notes that smoking and non-smoking behaviour can characterise the practice of seeking health support for influenza types of diseases in the USA (Biggerstaff et al., 2014). Another study finds smoking behaviour as a contributing factor towards vaccination against H1N1 among pregnant women in France (Freund et al., 2011). Thus, this study's finding of 'smoking behaviour' of Japanese citizens shaping their adherence to COVID-19 restrictions also relates to this behavioural attribute as a factor to consider when managing flu or influenza like pandemic.

Lastly, trust in government has appeared as a significant factor that positively influences public compliance across each of the protection behaviours. Indeed, various literature reflects on trust about

Table 6
Cronbach's alpha score for each of the three dependent measures considered in this study.

Dependent measure	Number of responses considered	Cronbach alpha value
Social distance	7	0.820
Personal protection behaviour	7	0.756
Social responsibility awareness	7	0.733

government within a pandemic situation. Schwartz and Yen (Schwartz & Yen, 2017), for example, call for building trust, further to collaboration across government levels and other entities, to prepare for and respond in a pandemic situation. Olganier and Mogensen (Olganier & Mogensen, 2020) suggest trust in government among other factors as having played an influential role in effectively managing the early stage of the COVID-19 in Denmark. Muto et al. (2020), in an article using the same dataset as this research, also refer to trust. Further, there have been explorations that highlighted the influence of trust in government and authorities during past pandemic situations (Chuang et al., 2015; Freimuth et al., 2014; Imai, 2020). Thus, our findings corroborate the significant influence of trust in government also in Japan during the early stage of COVID-19.

When faced with a pandemic situation, governments historically focus on adopting some public measures as has also been the case during the early outbreak of COVID-19 in Japan. However, as noted by different research covering different countries, subjective beliefs and demographic attributes can shape public adherence to these measures (Freimuth et al., 2014; Georgiou et al., 2020; Oleksy et al., 2021; Qian & Yahara, 2020). This research reaffirms the influence socio-demographic factors and personal attributes on compliance to public policies. Interestingly, individuality and subjective behaviour have not received much attention in public health policy management and institution. In a disaster or public health emergency, governments still tend to promote some standard measures and expect everyone to follow the suggested advice. Such has also been the case in Japan. Arguably, the outcomes imply the need of targeted campaigning when governments face a pandemic or a national disaster situation. Acknowledging that not everyone is likely to have the same level of trust and subjective beliefs moderated by socio-demographic and personal attributes are likely to influence individuals' compliances to any public policy, governments, rather than just instituting common public measures, may also focus on identifying demographic groups at high risk of non-compliance and adopt various social marketing strategies for effective policy outcomes. Future research can provide more insights in this regard. Overall, this article confirms a gap that has largely been overlooked in literature; and reflects the need for public health policy management and institution with a focus on individuality and personal behaviour across different cohorts in the public, further to creating an environment of trust, especially when the policies target some behavioural change to control a pandemic.

Finally, it is worth noting the technical contribution made in this research. As outlined earlier, while a recently published article on the same dataset also explores Japanese populaces' response to pandemic protection measures and advice (Muto et al., 2020), we zoom into the pandemic protection measures from three behavioural perspectives and considered more items of the respective survey question. In addition, rather than opting only for descriptive statistics and multiple linear regression, we use the CART analysis to extract features that are influential in characterising the dependent variables.

Notably, while the CART algorithm incepted in 1980s (Breiman et al., 1984), its application in social science domain has been recent and, to the best of our knowledge, very limited in subjective behaviour assessment. There have also been other algorithms which followed CART including Random Forest (RF) (Breiman, 2001) – another well-known tree-based machine learning approach. Fernández-Delgado et al. (Fernández-Delgado et al., 2014) claimed RF as the best classifier among the many machine learning approaches. A later work, however, has refuted this claim by identifying weaknesses in the previous study (Wainberg et al., 2016). Another recent research compared CART and RF and noted CART as a better performing approach for data with different conditions including when there are missing information (Hayes et al., 2015). Considering these, we focus on CART in our research. The choice is also due to the easy interpretability of CART outcomes. In this context, it may be argued that structural equation modelling (SEM) can also lead to similar hierarchical explanation of

relationships in data. However, as noted by Medina-Borja and Pasupathy (Medina-Borja & Pasupathy, 2007), CART is applicable when the aim is to extract knowledge from data as compared to SEM which can support or contradict any predetermined linear interrelationship between variables. With our focus on identifying influential variables, the use of CART is hence justified. Further, as also outlined earlier, CART can conceptualise non-parametric and nonlinear relationship among data (van Engelsdorp et al., 2010), and can therefore avoid limitations in a linear regression. Overall, our results complement the existing finding (Muto et al., 2020), and our approach may also motivate similar interpretative categorisation of pandemic protection behaviour in future research.

6. Conclusion

This study assesses the behavioural perspectives of population differentiated by socio-demographic and personal attributes when confronting a pandemic situation. Using a publicly available and recently collected data on Japanese citizens during the COVID-19 early outbreak and exploiting both CART and regression analysis, the study notes that socio-demographic and personal attributes of individuals indeed shape the subjective prevention actions and thereby the control of spread of a pandemic. Socio-demographic attributes – sex, marital family status, and having children – appear to have played an influential role in Japanese citizens' abiding by the COVID-19 protection behaviours, especially with women having children being noted more conscious than the male counterparts. Among the personality attributes, smoking behaviour appeared as a contributing factor with non-smokers or less-frequent smokers more compliant to the protection behaviours. Work status also appears to have some impact especially concerning social distancing. Further, trust in government appears as a significant factor.

There are some limitations of the research. The conclusions drawn are dependent on information recorded in the publicly available dataset. Thus, the study may have sustained the same limitations as the data collector including limitations due to self-reporting by participants, not covering all age groups, random sampling, and timing of data collection (Muto et al., 2020). Further, the way the items for the considered questionnaire have been interpreted and grouped across the three behaviours can be subjective, as is not uncommon in qualitative coding.

Despite the limitations, the findings provide useful insights especially concerning the need for public policy campaigning to account for variations in responses and protection behaviour due to socio-demographic and personal attributes during a pandemic. Such consideration, rather than a general campaign, may lead to more effective pandemic management – an issue to be explored in future research.

CRedit authorship contribution statement

Shahadat Uddin: Conceptualization, Investigation, Methodology, Software, Data curation, Formal analysis, Project administration, Writing – original draft, Writing – review & editing, Visualization.
Tasadduq Imam: Conceptualization, Investigation, Methodology, Project administration, Writing – original draft, Writing – review & editing.
Matloob Khushi: Investigation, Writing – original draft, Writing – review & editing.
Arif Khan: Investigation, Writing – original draft, Writing – review & editing.
Mohammad Ali: Investigation, Writing – original draft, Writing – review & editing.

References

- Alexander, D. L. J., Tropsha, A., & Winkler, D. A. (2015). Beware of R2: Simple, unambiguous assessment of the prediction accuracy of QSAR and QSPR models. *Journal of chemical information and modeling*, 55, 1316–1322.
- Balkhy, H. H., El-Saed, A., & Sallah, M. (2010). Epidemiology of H1N1 (2009) influenza among healthcare workers in a tertiary care center in Saudi Arabia: a 6-month surveillance study. *Infection control and hospital epidemiology*, 31, 1004–1010.

- Biggerstaff, M., Jhung, M. A., Reed, C., Garg, S., Balluz, L., Fry, A. M., et al. (2014). Impact of medical and behavioural factors on influenza-like illness, healthcare-seeking, and antiviral treatment during the 2009 H1N1 pandemic: USA, 2009–2010. *Epidemiology & Infection*, 142, 114–125.
- Bish, A., & Michie, S. (2010). Demographic and attitudinal determinants of protective behaviours during a pandemic: A review. *British Journal of Health Psychology*, 15, 797–824.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5–32.
- Breiman, L. (2017). *Classification and regression trees*. NW: CRC Press.
- Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). *Classification and regression trees*: CRC press.
- Bross, I. D. (1958a). How to use ridit analysis. *Biometrics*, 18–38.
- Bross, I. D. (1958b). How to use ridit analysis. *Biometrics*, 14, 18–38.
- Burki, T. (2020). China's successful control of COVID-19. *The Lancet Infectious Diseases*, 20, 1240–1241.
- Chuang, Y.-C., Huang, Y.-L., Tseng, K.-C., Yen, C.-H., & Yang, L.-h. (2015). Social capital and health-protective behavior intentions in an influenza pandemic. *PLoS One*, 10.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *psychometrika*, 16, 297–334.
- Curtis, V. A. (2014). Infection-avoidance behaviour in humans and other animals. *Trends in Immunology*, 35, 457–464.
- Djalante, R., Lassa, J., Setiamarga, D., Sudjatma, A., Indrawan, M., Haryanto, B., et al. (2020). Review and analysis of current responses to COVID-19 in Indonesia: Period of January to March 2020. *Progress in Disaster Science*, 6, 100091.
- Fakhrudin, B., Blanchard, K., & Ragupathy, D. (2020). Are we there yet? The transition from response to recovery for the COVID-19 pandemic. *Progress in Disaster Science*, 7, 100102.
- Fernández-Delgado, M., Cernadas, E., Barro, S., & Amorim, D. (2014). Do we need hundreds of classifiers to solve real world classification problems? *The journal of machine learning research*, 15, 3133–3181.
- Field, A. (2009). *Discovering statistics using SPSS*: Sage publications ltd.
- Flora, D., Jr. (2014). *J. Ridit analysis*. Wiley StatsRef: Statistics Reference Online.
- Freimuth, V. S., Musa, D., Hilyard, K., Quinn, S. C., & Kim, K. (2014). Trust during the early stages of the 2009 H1N1 pandemic. *Journal of Health Communication*, 19, 321–339.
- Freund, R., Ray, C. L., Charlier, C., Avenell, C., Truster, V., Tréluyer, J.-M., et al. (2011). Determinants of non-vaccination against pandemic 2009 H1N1 influenza in pregnant women: A prospective cohort study. *PLoS One*, 6, Article e20900.
- Georgiou, N., Delfabbro, P., & Balzan, R. (2020). COVID-19-related conspiracy beliefs and their relationship with perceived stress and pre-existing conspiracy beliefs. *Personality and Individual Differences*, 166, 110201.
- Hayes, T., Usami, S., Jacobucci, R., & McArdle, J. J. (2015). Using classification and regression trees (CART) and random forests to analyze attrition: Results from two simulations. *Psychology and aging*, 30, 911–929.
- Henry, B. (2019). Canadian pandemic influenza preparedness: Public health measures strategy. *Canada Communicable Disease Report*, 45(6), 159–163.
- Imai, H. (2020). Trust is a key factor in the willingness of health professionals to work during the COVID-19 outbreak: Experience from the H1N1 pandemic in Japan 2009. *Psychiatry and Clinical Neurosciences*, 74(5), 329–330.
- Isamu Yamamoto, K. U. (2020). *Japanese citizen's behavioral changes*.
- Itzwerth, R., Moa, A., & MacIntyre, C. R. (2018). Australia's influenza pandemic preparedness plans: An analysis. *Journal of Public Health Policy*, 39(1), 111–124.
- Iwasa, T., & Wada, K. (2013). Reasons for and against receiving influenza vaccination in a working age population in Japan: A national cross-sectional study. *BMC Public Health*, 13, 647.
- Jennings, LC (2009). Influenza A(H1N1)09: New Zealand's response to the pandemic threat. *Journal of the New Zealand Medical Association*, 122(1298), 11–16.
- Kannan, S., Ali, P. S. S., Sheeza, A., & Hemalatha, K. (2020). COVID-19 (novel coronavirus 2019)-recent trends. *European Review for Medical and Pharmacological Sciences*, 24, 2006–2011.
- Krzywinski, M., & Altman, N. (2017). Classification and regression trees. *Nature Methods*, 14, 757–758.
- La Torre, G., Di Thiene, D., Cadeddu, C., Ricciardi, W., & Boccia, A. (2009). Behaviours regarding preventive measures against pandemic H1N1 influenza among Italian healthcare workers, October 2009. *Eurosurveillance*, 14, 19432.
- Laerd Statistics. (2013). Multiple regression analysis using SPSS statistics [cited 2020 21-06-2020]; Available from: <https://statistics.laerd.com/spss-tutorials/multiple-regression-using-spss-statistics.php>.
- Lai, A. Y.-H., & Tan, S. L. (2015). Impact of disasters and disaster risk Management in Singapore: A case study of Singapore's experience in fighting the SARS epidemic. In D. P. Aldrich, S. Oum, & Y. Sawada (Eds.), *Resilience and recovery in Asian disasters: Community ties, market mechanisms, and governance* (pp. 309–336). Tokyo: Springer Japan.
- Lau, J. T., Griffiths, S., Choi, K. C., & Tsui, H. Y. (2010). Avoidance behaviors and negative psychological responses in the general population in the initial stage of the H1N1 pandemic in Hong Kong. *BMC Infectious Diseases*, 10, 139.
- Liu, Y., Gayle, A. A., Wilder-Smith, A., & Rocklöv, J. (2020). The reproductive number of COVID-19 is higher compared to SARS coronavirus. *Journal of Travel Medicine*, 27.
- Medina-Borja, A., & Pasupathy, K. S. (2007). *Uncovering complex relationships in system dynamics modeling: Exploring the use of CART, CHAID and SEM2007*.
- Mustafa, N. M., & A Selim, L. (2020). Characterisation of COVID-19 pandemic in paediatric age group: A systematic review and meta-analysis. *J Clin Virol*, 128, 104395.
- Muto, K., Yamamoto, I., Nagasu, M., Tanaka, M., & Wada, K. (2020). Japanese citizens' behavioral changes and preparedness against COVID-19: An online survey during the early phase of the pandemic. *PLoS One*, 15, Article e0234292.
- Olagnier, D., & Mogensen, T. H. (2020). The Covid-19 pandemic in Denmark: Big lessons from a small country. *Cytokine & Growth Factor Reviews*, 53, 10–12.
- Oleksy, T., Wnuik, A., Maison, D., & Łys, A. (2021). Content matters. Different predictors and social consequences of general and government-related conspiracy theories on COVID-19. *Personality and Individual Differences*, 168, 110289.
- Patanavanich, R., & Glantz, S. A. (2020). Smoking is associated with COVID-19 progression: A meta-analysis. *Nicotine & Tobacco Research*, 22(9), 1653–1656.
- Qian, K., & Yahara, T. (2020). Mentality and behavior in COVID-19 emergency status in Japan: Influence of personality, morality and ideology. *PLoS One*, 15, Article e0235883.
- Rose, S., & McGuire, T. G. (2019). Limitations of P-values and R-squared for stepwise regression building: A fairness demonstration in health policy risk adjustment. *Amer. Statist.*, 73, 152–156.
- Schwartz, J., & Yen, M.-Y. (2017). Toward a collaborative model of pandemic preparedness and response: Taiwan's changing approach to pandemics. *Journal of Microbiology, Immunology and Infection*, 50, 125–132.
- Setti, L., Passarini, F., De Gennaro, G., Barbieri, P., Perrone, M. G., Borelli, M., et al. (2020). Airborne transmission route of COVID-19: why 2 meters/6 feet of interpersonal distance could not be enough. 17, 2932.
- Shabu, S., Amen, K., Mahmood, K., & Shabila, N. (2020). *Risk perception and behavioral response to COVID-19 in Iraqi Kurdistan region*.
- Steinberg, D. (2009). CART: Classification and regression trees. In X. Wu, & V. Kumar (Eds.), *The top ten algorithms in data mining* (pp. 179–201). New York: Chapman & Hall/CRC.
- Tang, X., Du, R., Wang, R., Cao, T., Guan, L., Yang, C., et al. (2020). Comparison of hospitalized patients with ARDS caused by COVID-19 and H1N1. *Chest*, 158(1), 195–205.
- Tashiro, A., & Shaw, R. (2020). COVID-19 pandemic response in Japan: What is behind the initial flattening of the curve? *Sustainability*, 12, 5250.
- Teasdale, E., Yardley, L., Schlotz, W., & Michie, S. (2012). The importance of coping appraisal in behavioural responses to pandemic flu. *British Journal of Health Psychology*, 17, 44–59.
- Tran, C., Abraham, A., & Jain, L. (2008). Soft computing paradigms and regression trees in decision support systems. In V. Sugumaran (Ed.), *Intelligent information technologies: Concepts, methodologies, tools, and applications: Concepts, methodologies, tools, and applications* (pp. 1011–1035). Hershey: IGI Global.
- van Engelsdorp, D., Speybroeck, N., Evans, J. D., Kim Nguyen, B., Mullin, C., Frazier, M., et al. (2010). Weighing risk factors associated with bee colony collapse disorder by classification and regression tree analysis. *Journal of Economic Entomology*, 103, 1517–1523.
- Wadood, M. A., Lee, L. L., Huq, M., Mamun, A., Mohd, S., & Hossain, M. (2020). *Practice and perception of Bangladeshi adults toward COVID-19: A cross-sectional study*.
- Wainberg, M., Alipanahi, B., & Frey, B. J. (2016). Are random forests truly the best classifiers? *The Journal of Machine Learning Research*, 17, 3837–3841.
- WHO. (2020a). *Timeline: WHO's COVID-19 response* [updated 2020].
- WHO. (2020b). *Weekly epidemiological update - 15 December 2020* [updated 2020].
- Yamamoto, I. (2020). *Japanese citizens' behavioral changes*. Ann Arbor, MI: Inter-university Consortium for Political and Social Research (distributor).
- Yitzhaki, S., & Schechtman, E. (2012). *The Gini methodology: A primer on a statistical methodology*: Springer Science & Business Media.
- Yohannes, Y., & Webb, P. (1999). *Classification and regression trees, CART: A user manual for identifying indicators of vulnerability to famine and chronic food insecurity: Intl food policy res Inst.*
- Yoo, K., Yoo, H., Lee, J. M., Shukla, S. K., & Park, J. (2018). Classification and regression tree approach for prediction of potential hazards of urban airborne Bacteria during Asian dust events. *Scientific Reports*, 8, 11823.
- Zajenkowski, M., Jonason, P. K., Leniarska, M., & Kozakiewicz, Z. (2020). Who complies with the restrictions to reduce the spread of COVID-19? Personality and perceptions of the COVID-19 situation. *Personality and Individual Differences*, 166, 110199.
- Zhan, S., Yang, Y. Y., & Fu, C. (2020). Public's early response to the novel coronavirus-infected pneumonia. *Emerg Microbes Infect.*, 9, 534.