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.

Infection, Genetics and Evolution 87 (2021) 104657

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

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

Short communication

Dynamic aspects of *ORF1ab* and *N* RNA cycle threshold values among COVID-19 patients in China

Alok Bhargava

University of Maryland School of Public Policy, College Park, MD 20742, United States of America

A B S T R A C T

The dynamics of viral loads among COVID-19 patients in Changzhou, China were evaluated using dynamic random effects models. The models were estimated by maximum likelihood methods allowing for between and within patient variations. Statistical criteria were developed for focusing on viral RNAs for clinical decision making. The empirical results showed that inflammation among patients were significant predictors of cycle threshold values for *ORF1ab* and *N* RNAs. Moreover, within subject variations were higher in Ct values of *ORF1ab* RNA indicating that assessment of *N* RNA may be adequate in resource-poor settings. The inter-relationships between *ORF1ab* and *N* RNAs were investigated and the need for developing comprehensive models for viral load dynamics is emphasized.

1. Introduction

The SARS-CoV-2 pandemic has already claimed over a million lives worldwide (Coronavirus Resource Center, 2020; Sironi et al., 2020). This virus gains entry to cells by attaching to ACE-2 receptors and its ~30 kb positive sense single stranded RNA encodes multiple proteins. ORF1a mainly encodes proteases and proteins that interfere with the innate immune response, while ORF1b mostly encodes proteins involved in RNA synthesis. Proteins such as spike, viral envelope, membrane, and Nucleocapsid are encoded by sub-genomic RNAs (Costela-Ruiz et al., 2020). Note that for simplifying the terminology, PCR products corresponding to the RNAs encoding the ORF1ab and N proteins are abbreviated as ORF1ab and N RNAs, respectively. Further, there has been interest in monitoring the viral load kinetics among Covid-19 patients especially of *ORF1ab* and *N* RNAs (He et al., 2020, Xu et al., 2020).Typically, investigators report SARS-CoV-2 RNA copies and/or cycle threshold values from RT-PCR assays, with Ct values >40 indicating undetectable viral loads.

The investigation of four sets of conceptual and methodological issues can facilitate clinical decision making. First, SARS-CoV-2 appears to induce an "imbalanced host response"(Blanco-Melo et al., 2020) in terms of production of cytokines such as IL-1, IL-2, IL-4, IL-6, IL-10, IL-17, and TNF-α, and interferon INF-I, INF-II, INF-III. The cytokines can elevate C-reactive protein levels that are indicators of inflammation and are likely to affect viral dynamics. Second, the estimation of associations between viral loads or Ct values at different time points (*i.e.*, autocorrelations) needs to account for between and within patient variations that can be large (Bhargava and Sargan, 1983; Bhargava et al., 2018). Third, it is important to investigate using statistical method the relative merits of analyzing viral loads and Ct values for *ORF1ab* and *N* RNAs, reflecting, respectively, non-structural and structural proteins, for the efficient utilization of medical resources. While the previous literature (Shi et al., 2020) has analyzed data for *ORF1ab* and *N* RNAs, the between and within-patient variations in viral loads are important for RNA selection. Fourth, the effects of explanatory variables such as patients' biomarkers for viral load kinetics need to be investigated using comprehensive empirical models. The estimation methods in this article are helpful for selection of viral RNA and for drawing robust inferences regarding the effects of biomarkers for viral load dynamics. The empirical results from Changzhou, China can facilitate the formulation of evidence-based policies.

2. Methods

Dynamic random effects models, accounting for between and within patient variations in natural logarithms of Ct values for *ORF1ab* and *N* RNAs, were estimated by maximum likelihood methods (Bhargava and Sargan, 1983) using a data set from Changzhou, China for 49 COVID-19 patients observed at 3 time points within a fortnight (Xu et al., 2020). The Specification 1 allowed Ct values to depend on patients' ages and its square to account for possible nonlinearities:

E-mail address: Bhargava@umd.edu.

https://doi.org/10.1016/j.meegid.2020.104657
Received 31 August 2020; Received in revised form 30 October 2020; Accepted 7 November 2020
Available online 1 December 2020

Table 1
Sample means and standard deviations of cycle threshold values for *ORF1ab* and *N* RNAs and for selected explanatory variables of COVID-19 patients in Changzhou, China [1].

Survey round:	Survey round 1		Survey round 2		Survey round 3	
Variable:	Mean	SD	Mean	SD	Mean	SD
Age, years	42.32	17.92				
Gender, 0–1	1.50	0.50				
C-reactive protein, mg/dL	3.28	1.89				
Plasma Ca, mmo/L	2.25	0.12				
Ct *ORF1ab* RNA, n	28.48	5.04	34.28	5.31	37.22	4.64
Ct *N* RNA, n	29.96	4.90	35.06	4.50	37.62	4.10

[1] The 3 survey rounds were conducted within a fortnight for the 49 patients.

$$ln(\mathrm{CtORF1ab})_{it} = a_0 + a_1 ln(\mathrm{Age})_i + a_2 [ln(\mathrm{Age})]^2_i + a_3 ln(\mathrm{CtORF1ab})_{it-1} + u_{it} \; (i = 1,\ldots,49; t = 2,3) \quad (1)$$

In eq. (1), "ln" represents natural logarithms and a_0, a_1, a_2, and a_3 are unknown coefficients, with a_3 being the coefficient of the "lagged dependent variable". The u_{it}'s can be decomposed in a random effects fashion as:

$$u_{it} = \delta_i + v_{it} \quad (2)$$

where δ_i are patient-specific random effects that were distributed with zero mean and constant variance, and v_{it} were distributed with zero mean and constant variance. Specification 2 replaced the age variable by C-reactive protein levels. A comparison of the maximized values of log-likelihood functions for Specifications 1 and 2 can afford insights into the adequacy of empirical models.

3. Results

The sample means and standard deviations of explanatory variables such as patients' age, gender, and C-reactive protein levels and for Ct values for *ORF1ab* and *N* RNAs at 3 time points are reported in Table 1. Patients' C-reactive protein levels were generally elevated at admittance, and there were rapid increases in Ct values and they approached 40 at the time of the third survey for most patients.

In Table 2, the results from Specification 1 for Ct values for *ORF1ab* and *N* RNAs showed nonlinear effects with respect to age though the coefficients were not statistically significant at the 5% level, presumably due to small number of patients. This was also true for the ratios of between to within variances that require larger sample sizes. However, coefficients of the lagged dependent variables were significant ($P < 0.05$) and the estimates 0.346 and 0.288 for *ORF1ab* and *N* RNAs, respectively, showed rapid convergence of Ct values to 40. Moreover, within patient variances for Ct values for *ORF1ab* and *N* RNAs were 0.0163 and 0.0100, respectively, thereby showing 50% higher variations for *ORF1ab* RNA.

The coefficients of CRP levels in Specification 2 were significantly associated with Ct values for *ORF1ab* and *N* RNAs, though the squared terms were not significant. Moreover, maximized values of log-likelihood functions were higher for Specifications 2 indicating the importance of inflammation rather than of age for the dynamics of viral loads. Lastly, the fit of models for Ct values for *N* RNA were better than those for *ORF1ab* RNA.

4. Discussion

The results from the empirical analyses showed the importance of modeling the dynamics of Ct values especially for *N* RNA among patients in Changzhou, China. The relatively high variations in Ct values underscore the importance of modeling viral load kinetics using data on large numbers of patients. Such analyses are essential, for example, when investigating the effects of plasma vitamin D levels (Grant et al., 2020) on viral loads. While patients' 25 (OH) D concentrations were not assessed in the study, plasma calcium levels were not significantly associated with Ct values; estimated coefficients in the models for *ORF1ab* and *N* RNAs were not statistically different from zero.

Lastly, for further comparisons of the empirical models for *ORF1ab* and *N* RNAs using statistical criteria, the bivariate correlations between Ct values of *ORF1ab* and *N* RNAs in the three survey rounds were 0.96, 0.91, and 0.92, respectively. Subsequently, enlarged models were estimated for Ct values of *ORF1ab* RNA that were explained by the Ct values for *N* RNA and *vice versa*. The estimated coefficient of Ct values for *N* RNA was 1.07 in the model for *ORF1ab* RNA and it was not statistically different from 1. In contrast, in the model for *N* RNA, coefficient of Ct values for *ORF1ab* RNA was 0.77 and was significantly less than 1. This was perhaps not surprising because the *N* protein is encoded on a sub-genomic RNA that is more abundant than the *ORF1ab* RNA. In view of the apparent lack of cross-reactivity of these RNAs with other coronaviruses (Yan et al., 2020), these results indicate that there may not be additional information in Ct values for *ORF1ab* RNA, and Ct values for *N* RNA can convey the essential information for clinical decision making and for policy formulation.

Table 2
Maximum likelihood estimates of dynamic random effects models for cycle threshold values for *ORF1ab* and *N* RNAs for COVID-19 patients in Changzhou, China[1].

Dependent variable:	ln (Ct *OFR1ab* RNA), n				ln (Ct *N* RNA), n			
	Specification 1[2]		Specification 2[2]		Specification 1[2]		Specification 2[2]	
Explanatory variables:	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Constant	2.394	0.494	2.314	0.430	2.564	0.462	2.551	0.335
ln (Age), years	−0.076	0.201	–		−0.025	0.207	–	
[ln (Age)]^2, n	0.020	0.030	–		0.009	0.031	–	
ln (C-reactive protein), mg/dL	–		0.035*	0.017	–		0.030*	0.016
[ln(C-reactive protein)]^2, n	–		−0.006	0.006	–		−0.007	0.005
Lagged dependent variable	0.346*	0.011	0.358*	0.124	0.288*	0.093	0.295*	0.096
Between/within variance, n	0.153	0.241	0.122	0.264	0.384	0.319	0.337	0.290
Within variance, n	0.0163		0.0166		0.0100		0.0101	
2× Maximized log-likelihood function	557.10		559.01		598.98		601.43	

[1] Slope coefficients and standard errors are reported for 49 patients in 3 survey rounds [2] Specifications 1 and 2 included, respectively, patients' ages and C-reactive protein levels and their respective squares. *$P < 0.05$.

Declaration of Competing Interest

The research in this manuscript was not supported by any funding agency and there are no conflicts of interest

Credit author statement

Alok Bhargava is responsible for all aspects of this manuscript

Acknowledgements

The author would like to thank Dr. Yuan Xue for sharing the data on patients in Changzhou, China, and Dr. Jiehong Lou for her help in creating the longitudinal database. This revision has benefited from helpful comments of the reviewers, editors, and, especially, Dr. Jonathan Dinman

References

Bhargava, A., Sargan, J.D., 1983. Estimating dynamic random effects models from panel data covering short time periods. Econometrica 51, 1635–1659.

Bhargava, A., Booysen, F., Walsh, C., 2018. Health status, food insecurity, and time allocation patterns of patients with AIDS receiving antiretroviral treatment in South Africa. AIDS Care 30, 361–368.

Blanco-Melo, D., Nilsson-Payant, B., Liu, W.-C., Ulh, S., Hoagland, D., Moller, R., Jordan, T., Oishi, K., Panis, M., Sachs, D., Wang, T., Schwartz, R., Lim, J., Albrecht, R., tenOever, B., 2020. Imbalanced host response to SARS-CoV-2 drives development of COVID-19. Cell 181, 1–10. https://doi.org/10.1016/j.cell.2020.04.026.

Coronavirus Resource Center, 2020. Johns Hopkins University. https://coronavirus.jhu.edu/.

Costela-Ruiz, V., Illescas-Montes, R., Peurta-Puerta, J.M., Ruiz, C., Melguizo-Rodriguez, L., 2020. SARS-CoV-2 infection: the role of cytokines in COVID-19 disease. Cytokine Growth Factor Rev. 54, 62–75. https://doi.org/10.1016/j.cytogfr.2020.06.001.

Grant, W., Lahore, H., McDonnell, S., Baggerly, C., French, C., Aliano, J., Bhattoa, J., 2020. Evidence that vitamin D supplementation could reduce the risk of influenza and COVID-19 infections and deaths. Nutrients 12, 988. https://doi.org/10.3390/nu12040988.

He, X., Lau, E., Wu, P., Deng, X., Wang, J., Hao, X., Lau, Y., Wong, J., Guan, Y., Tan, X., Mo, X., Chen, Y., Liao, B., Chen, W., Hu, F., Zhang, Q., Zhong, M., Wu, Y., Zhao, L., Zhang, F., Cowling, B., Li, F., Leung, G., 2020. Temporal dynamics in viral shedding and transmissibility of COVID-19. Nat. Med. 26, 672–675. https://doi.org/10.1038/s41591-020-0869-5.

Shi, J., Han, D., Zhang, R., Li, J., Zhang, R., 2020. Molecular and serological assays for SARS-CoV-2: insights from genome and clinical characteristics. Clin. Chem. 2020 (66), 1030–1046.

Sironi, M., Hasnain, S.E., Rosenthal, B., Tung, P., Luciani, F., Shaw, M.-A., Sallum, M.A., Mirhashemi, M.E., Morand, S., Gonzalez-Candelas, F., 2020. SARS-CoV-2 and COVID-19: a genetic, epidemiological, and evolutionary perspective. Infect. Genet. Evol. 84, 104384 https://doi.org/10.1016/j.meegid.2020.104384.

Xu, T., Chen, C., Zhu, Z., Cui, M., Chen, C., Dai, H., Xue, Y., 2020. Clinical features and dynamic s of viral load in imported and non-imported patients with COVID-19. Int. J. Infect. Dis. 94, 68–71, 10.101j.ijid.2020.03.022.

Yan, C., Cui, J., Huang, L., Du, B., Chen, L., Xue, G., Li, S., Zhang, W., Zhao, L., Sun, Y., Yao, H., Li, N., Zhao, H., Feng, Y., Liu, S., Zhang, Q., Liu, D., Yuan, J., 2020. Rapid and visual detection of 2019 novel coronavirus (SARS-CoV-2) by a reverse transcription loop-mediated isothermal amplification assay. Clin. Microbiol. Infect. 26, 773–779. https://doi.org/10.1016/j.cmi.2020.04.001.