ORIGINAL RESEARCH ARTICLE

QALYs for COVID-19: A Comparison of US EQ-5D-5L Value Sets

Stephen Poteet[1] · Benjamin M. Craig[2]

Abstract

Background In economic evaluations, quality-adjusted life-years (QALYs) can serve as a unit of measurement for disease burden. Obtaining QALY values for COVID-19 presents a challenge owing to the availability of two US EQ-5D-5L value sets and the potentially asymptomatic presentation of the disease. The first value set was completed allowing for the discounting of future health outcomes while the second value set is undiscounted.

Objective The objective of this study was to compare the distribution of QALY values using a national survey and the two published value sets; and to estimate the association between COVID-19 outcomes and QALY losses.

Methods Between 9 and 11 November, 2020, 1153 US adults completed the EQ-5D-5L instrument (five items and a visual analog scale) as well as self-reported their demographics, COVID-19 symptoms, and memberships to populations that are at risk of COVID-19 infection. The two US value sets were applied to the EQ-5D-5L responses to produce QALY values. We estimated the mean QALYs by visual analog scale decile and a generalized linear model of COVID-19 outcomes.

Results The discounted values are higher than the undiscounted values for each visual analog scale decile owing to methodological differences. Persons at increased risk, with a fever in the past day, and with one or more other symptoms have significantly greater QALY losses ($p < 0.01$). Overall, non-institutionalized individuals at risk of symptomatic clinical COVID-19 equal 0.68 for the 2016 value set (95% confidence interval 0.49–0.87) and 0.10 for the 2017 value set (95% confidence interval − 0.31 to 0.51) QALYs.

Conclusions Multiple studies have shown that decision makers discount future health outcomes, which increase QALY values. This study confronts the practical implications of these methodological advances for use in COVID-19 economic evaluations. Health economists will be able to use the QALY values in this study to better evaluate health interventions against COVID-19.

Key Points

Methodological differences in value sets can lead to significant disparities when calculating quality-adjusted life-years.

This study provides economists with accurate health valuations for COVID-19 that can be used to analyze different health interventions.

Future research should focus on the preferences of hospitalized patients to fully understand the burden of COVID-19.

✉ Stephen Poteet
poteet.stephen@spcollege.edu

Benjamin M. Craig
bcraig@usf.edu

1 Department of Social/Behavioral Sciences, St. Petersburg College, 2465 Drew St, ES313D, Clearwater, FL 33765, USA

2 Department of Economics, University of South Florida, Tampa, FL, USA

△ Adis

1 Introduction

Economic evaluations of medical interventions are an important tool for many healthcare systems around the world. Stakeholders allocate healthcare resources based on the information these evaluations provide. Quality-adjusted life-years (QALYs) serve as a unit of measurement for the burden of disease studies and are an integral part of the evaluative process. Quality-adjusted life-years can range from negative infinity to 1, depending on the study. Typically, 0 corresponds to "dying immediately" and 1 represents "1 year with no health problems then die." Negative values would represent health outcomes that are worse than death. Health economists continue to strive for improved accuracy, applicability, and generalizability of QALY values.

Quality-adjusted life-year values can be calculated using value sets for the EQ-5D-5L instrument [1]. The EQ-5D-5L was introduced in 2005 and is an updated version of the EQ-5D-3L instrument allowing for greater sensitivity of measurement [2]. The EQ-5D-5L contains five dimensions: mobility, self-care, usual activities, anxiety/depression, and pain/discomfort; and 5 levels: no problems (level 1), slight problems, moderate problems, severe problems, and extreme problems (level 5). Respondents would indicate for each dimension which level corresponds to their current health state. The health state "11111" indicates no health problems in any domain and "55555" indicates extreme problems in all domains. In addition to these five items, the instrument includes the EQ-VAS, a visual analog scale (VAS) that captures self-reported health status from best (100) to worst (0) imaginable health.

To summarize EQ-5D-5L responses as QALY values, researchers typically use a published value set (also called tariffs). These value sets are based on health valuation studies and each study may use different elicitation

techniques such as time trade-offs (TTOs) or discrete choice experiments. Apart from the elicitation technique, there are significant differences among value sets depending on the analysis and country of origin [3–6].

Obtaining accurate QALY values for specific diseases can also be a challenge if EQ-5D-5L responses for individuals with that disease are unknown. COVID-19 outcomes present an added challenge because of their recency and heterogeneous presentation. Figure 1 is a simplistic description of two risk groups who may contract COVID-19 and how moving from one stage to another could constitute losses or gains of QALYs. This cyclical model implies knowledge of the COVID-19 burden at each node; however, QALY values for these outcomes have yet to be published to the best of our knowledge.

The goal of our study is to compare the distribution of QALY values using a national survey and the two published value sets and to estimate the association between COVID-19 outcomes and QALY losses using a generalized linear model (GLM). Based on these associations, we will calculate the QALY values for the six groups in Fig. 1 for use in future economic evaluations and burden of disease analyses.

2 Methods

This study is a secondary analysis of a protocol that received an institutional review board exemption.

2.1 US EQ-5D-5L Value Sets

The first US EQ-5D-5L valuation study was conducted in 2016 with 8222 respondents from all 50 states of USA and the District of Columbia [7]. It comprised an online survey that included a discrete choice experiment, namely a paired comparison including the five EQ-5D-5L attributes and lifespan. During the analysis, the authors allowed for

Fig. 1 COVID-19 outcomes
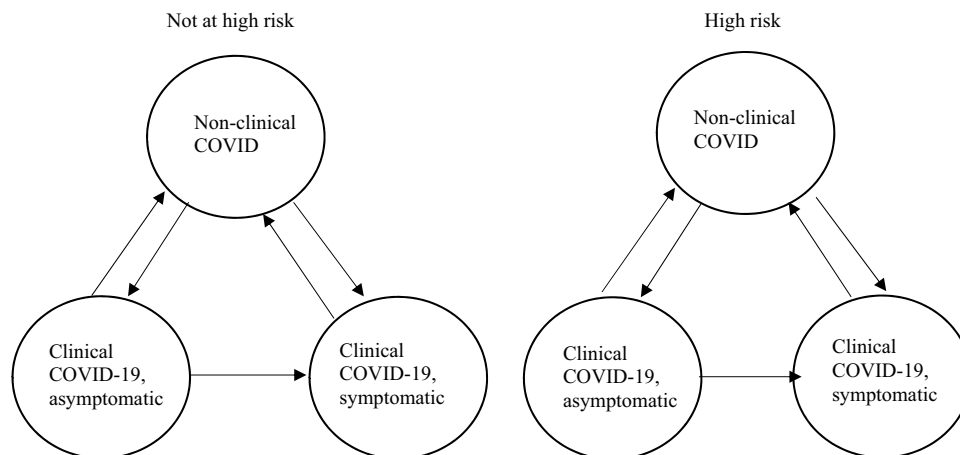
discounting of future health outcomes. The second study was conducted in 2017 with 1062 respondents from six metropolitan areas (Chicago, Philadelphia, Seattle, Birmingham, Phoenix, and Denver) [8]. It comprised an interview that included a composite TTO task, namely adaptive paired comparisons with the same six attributes trading off life-years. During the analysis, the authors did not allow for discounting. The first study also conducted a predictive model competition and differs from the standard EQ-VT protocol and the second study followed the EQ-VT protocol [7, 8]. Both studies were funded by the EuroQol Research Foundation; however, the second study also received financial support from the pharmaceutical industry, but the funding agreement ensured independence. We refer to each value set by its main methodological difference, i.e., discounted vs undiscounted.

The lack of discounting in health valuations causes QALY values of "poor" health to be lower (downward bias) [6]. For example, each TTO response equates 10 years with EQ-5D-5L attributes to less time in full health; therefore, "good" health trades long-term life-years and "poor" health trades near-term life-years. Under discounting, long-term life-years are typically worth less than near-term life-years. Likewise, the lack of discounting causes the unadjusted TTO analysis to produce biased estimates for poor health, known as its constant proportionality assumption [3]. These value sets can be found in Appendix 3 of the Electronic Supplementary Material.

2.1.1 Survey Instrument

The survey instrument comprised five sections: consent, screener, background, choice task, and follow-up (Screenshots) [9]. This paper is on information obtained from the screener, background, and follow-up portions of the survey, namely the EQ-5D-5L instrument. The screener included five questions that provided information on respondent demographics (Screenshots, p. 1–2). The background asked ten questions related to COVID-19 and whether the respondent was "a member of the following critical populations?" These included: person at increased risk of acquiring or transmitting COVID-19, healthcare personnel, person with limited access to routine vaccination services, person at increased risk for severe COVID-19 illness, and/or other essential worker (Screenshots, p. 3). Within the follow-up questions, the EQ-5D-5L instrument was included (five items in a fixed order followed by a VAS from best to worst imaginable health), as were other questions related to health of the respondent. Respondents were asked if they "had a fever in the past day," "have a new or worsening cough," and "any of these other symptoms" (Screenshots, p. 17–24).

On 5 November, 2020, the institutional review board at Advarra® determined that the study (Pro00047418) was exempt according to 45 CFR 46.104(d) [2]. The recruitment occurred from 9 to 11 November, 2020 by Dynata®, an online market research firm, rendering a national sample for US adults aged 18–87 years, including persons from the District of Columbia and all US states, except South Dakota. During recruitment, two exclusion criteria were applied: (1) residing outside the 50 US states or the District of Columbia and (2) being age 17 years or less (age 18 years in Nebraska or Alabama or age 20 years in Mississippi).

2.1.2 Analysis

Our analysis plan was as follows: (1) comparison of survey respondent characteristics to the 2019 American Community Survey; (2) association between VAS deciles and QALY values from the EQ-5D-5L responses and the two US value sets; (3) association between demographic variables and health/risk variables and losses in QALY values using a GLM; and (4) calculation of QALY values of COVID-19 groups adjusted for respondent demographics.

2.1.3 GLM of QALY Losses

Apart from the log link for the expected loss in QALYs (i.e., 1 minus QALY), the GLM family function was determined by completing a modified Park test [10]. This approach found that a gamma distribution was best to control for skewness and heteroskedasticity in unexplained losses [11–13].

For the base case, we selected 45 years of age; therefore, included adjusted age [i.e., (age in years − 45)/45] and its square in the GLM regression. Other demographic variables included indicators for female/other, Black or African American alone, Asian or other, and Hispanic or Latino. Thus, the constant term (0) represents a 45-year-old, non-Hispanic, white male US adult.

For the health variables, the GLM regression included indicators for whether a respondent was in a high-risk population, was considered to have had clinical COVID-19, and had a fever, cough, or at least one other symptom. A respondent was deemed high risk if they replied that they were at "increased risk for acquiring or transmitting COVID-19" or at "increased risk for severe COVID-19 illness." Respondents were deemed to have had clinical COVID-19 if they selected either "I tested positive for COVID-19" or "a doctor ordered me to quarantine for possible COVID-19." Respondents were categorized as non-clinical COVID if they did not test positive for COVID-19 or were not ordered to quarantine by their doctor because of COVID-19. Respondents had one or more symptoms if they indicated they had one of the following: congestion or runny nose, nausea or vomiting, new loss of taste or smell, headache, diarrhea, shortness of breath or difficulty breathing, muscle or body aches, fatigue, and sore throat.

3 Results

3.1 Respondent Characteristics

We first compared the respondent demographics to the 2019 American Community Survey using a chi-square test and found a few significant differences (Table 1) [14]. Respondents were more likely to be 35–54 years of age (38.68% vs 32.43%; $p < 0.01$), white (76.67% vs 73.61%; $p < 0.01$), and non-Hispanic (87.25% vs 83.60%; $p < 0.01$), and these difference may affect the interpretation of the COVID-19 QALY values.

3.2 Comparison of US EQ-5D-5L Values

In Fig. 2, the *y* axis is the mean QALY values by EQ-VAS decile of the two US value sets. Frequency of respondent per decile ranged from 8.07 to 11.10%. The discounted value set produced a mean value that ranged from 0.86 to 0.98. The mean values of the undiscounted value set ranged lower, from 0.53 to 0.94 ($p < 0.01$) (Table 2).

3.3 GLM Analysis of COVID-19 QALY Losses

The GLM analysis of COVID-19 QALY losses provides two models for each US EQ-5D-5L value set, (1) demographics only and (2) demographics with health/risk variables. The coefficients are reported as changes in ln QALY losses. For example, in the base case for model 3, a 45-year-old, non-Hispanic, white male US adult would face a ln QALY loss of -1.65. We can calculate the QALY for the base case as $1 - e^{-1.65}$, which equals a QALY of 0.81. When interpreting the coefficients, positive values would mean increases in QALY losses (i.e., lower QALYs) and negative values would mean decreases in QALY losses (i.e., higher QALYs). In terms of demographics, models 1 and 3 show significant differences in adjusted age ($p < 0.01$), but only model 3 shows slight significant differences by race ($p < 0.10$). After accounting for health/risk variables, models 2 and 4 find significant increases in QALY losses for high-risk individuals (0.55 and 0.66; $p < 0.01$), those with a fever in the past day (0.61 and 0.74; $p < 0.01$), and one or more other symptoms (0.79 and 0.84; $p < 0.01$).

3.4 QALY Values for COVID-19 Outcomes

Table 3 provides the marginal effects of different groups controlling for demographics based on the GLM analysis results. We then report the results as 1 minus the coefficient to convert into QALYs to make it easier for the reader. All QALY losses are statistically significant ($p < 0.01$). The discounted values range from 0.68 (95% confidence interval [CI] 0.49–0.87) to 0.98 (95% CI 0.97–0.98) after controlling

Table 1 Demographic characteristics

|	Completes N (%)	ACS %	p value
	100.00 (1153)		
Age, years			
18–34	29.14 (336)	29.75	< 0.01
35–54	38.68 (446)	32.43	
55 and older	32.18 (371)	37.82	
Sex			
Male	48.74 (562)	48.68	0.99
Female	51.08 (589)	51.32	
Other/prefer not to say	0.17 (2)		
Race			
White alone	76.67 (884)	73.61	< 0.01
Black or African American alone	12.06 (139)	12.45	
American Indian or Alaska Native alone	0.87 (10)	0.83	
Asian alone	6.24 (72)	5.92	
Native Hawaiian or other Pacific Islander alone	0.26 (3)	0.18	
Some other race alone	2.17 (25)	4.56	
Two or more races	1.73 (20)	2.46	
Ethnicity			
Hispanic or Latino	12.75 (147)	16.40	< 0.01
Other	87.25 (1006)	83.60	

ACS American Community Survey
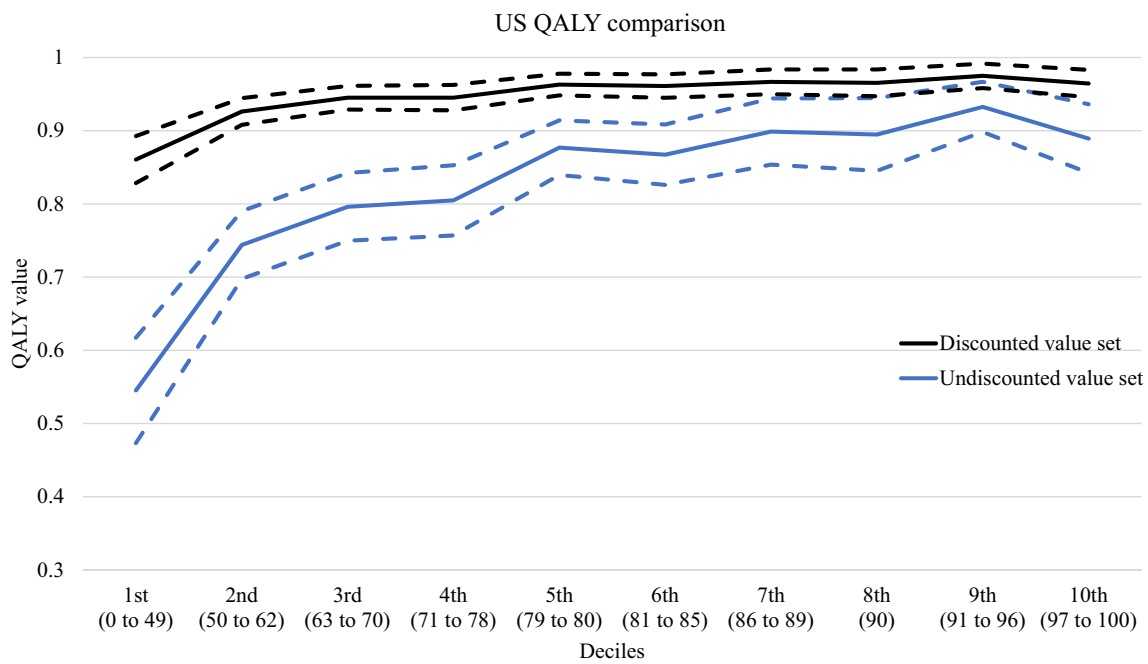
Fig. 2 Quality-adjusted life-year (QALY) values by EQ-VAS decile (mean and 95% confidence interval)

Table 2 Generalized linear model analysis of ln QALY losses

| | Discounted value set | | Undiscounted value set | |
| | Craig and Rand [7] | | Pickard et al. [8] | |
	Model 1	Model 2	Model 3	Model 4
Demographic variables				
Female/other	− 0.11	0.12	− 0.06	0.09
Age adjusted	− 0.69***	− 0.38*	− 0.59***	− 0.31*
Age adjusted squared	− 0.28	− 0.17	− 0.12	− 0.03
Black	− 0.24	− 0.17	− 0.29**	− 0.21
Asian/other	− 0.29	0.14	− 0.29*	0.06
Hispanic	− 0.11	− 0.10	− 0.01	0.03
Health/risk variables				
At high risk		0.66***		0.55***
Clinical COVID-19		0.20		0.26
Fever		0.74***		0.61***
Cough		0.18		0.13
One or more symptoms		0.84***		0.79***
Constant term	− 2.81***	− 3.77***	− 1.65***	− 2.47***

QALY quality-adjusted life-year

Coefficients are reported in changes in ln QALY losses; the negative values on the coefficients (female, Hispanic) would correspond to higher QALYs; age adjusted = (age − 45)/45

*p < 0.10; **p < 0.05; ***p <0.01

for demographic variables. The undiscounted values are lower for all groups, from 0.10 (95% CI − 0.31 to 0.51) to 0.91 (95% CI 0.90–0.92). In summary, a symptomatic person with COVID-19 is in "fair" health (0.68) based on the discounted value set and close to death (0.10) based on the undiscounted value set.

4 Discussion

The differences in QALY values between the two US EQ-5D-5L value sets are apparent based on the VAS deciles or COVID-19 outcomes. When comparing mean QALY values in Fig. 2, the undiscounted values have several areas where they are non-monotonic, and the 95% CI is much wider than the discounted values. This is to be expected owing to the differences in discounting, representation, and sample size. There continues to be methodological advancement as economists strive for greater accuracy in health valuation. One area in particular is the assumption of linear time preferences and how this may violate basic human behavior. Recent studies have shown that allowing for nonlinear time preferences, through the use of discounting functions, provides a better statistical fit [6, 15]. These differences could lead to misallocation of resources and incorrect implementation of interventions against COVID-19 depending on which value set is chosen. Health economists will have to decide between

Table 3 QALY values for COVID-19 outcomes

	Discounted value set			Undiscounted value set		
	Craig and Rand [7]			Pickard et al. [8]		
	QALY	Std. error	95% CI		QALY	Std. error	95% CI	
Not at high risk								
Non-clinical COVID-19	0.98	< 0.00	0.97	0.98	0.91	0.01	0.90	0.92
Clinical, asymptomatic	0.97	0.01	0.96	0.99	0.89	0.02	0.84	0.93
Clinical, symptomatic	0.83	0.05	0.73	0.93	0.48	0.12	0.24	0.72
At high risk								
Non-clinical COVID-19	0.96	0.01	0.94	0.97	0.85	0.01	0.82	0.88
Clinical, asymptomatic	0.95	0.02	0.91	0.98	0.81	0.02	0.72	0.89
Clinical, symptomatic	0.68	0.10	0.49	0.87	0.10	0.10	− 0.31	0.51

CI confidence interval, *QALY* quality-adjusted life-year, *Std.* standard

Clinical COVID-19 if individual selected either: "I tested positive for COVID-19" or "a doctor ordered me to quarantine for possible COVID-19;" otherwise they were labeled non-clinical COVID-19

the conventional approach of using undiscounted values, which would be biased, or choosing discounted values that are more accurate.

There were also dramatic losses of QALYs for those non-institutionalized individuals in the at-risk group that developed COVID-19 symptoms. When using the undiscounted QALY values, we see that these individuals are at a state close to zero (dead). COVID-19 is a unique disease, there are differences not only in infection rate and severity when compared with influenza, but some individuals experience long-term side effects [16, 17]. We can clearly see the enormous potential gains in QALYs by preventing symptomatic COVID-19 in individuals at high risk and those not at high risk.

4.1 Limitations

Our study is limited in the generalizability to the entire US population as there are significant differences between the survey respondents and the American Community Survey. This is because there are individuals that are more likely to take online surveys. We thought about using sampling weights but determining the proper sampling weights can be difficult and we questioned if this would actually make the data more representative of the population. We also do not have parameters for hospitalized patients with COVID-19 within this study, thus we are unable to predict the full burden of the disease. This study also only includes six different subgroups. This could be an issue if these do not align with future modeling studies. We also must be cognizant of when this survey was taken by respondents. This occurred right after the 2020 US election; therefore, the context of the survey is important when understanding a disease that has been highly politicized.

5 Conclusions

This study is one of the first to determine the QALY values for COVID-19 outcomes in the USA or elsewhere. Health economists will be able to use this information to better model the merits of health interventions against COVID-19. There are already studies taking place that use QALYs to determine the cost effectiveness of different COVID-19 suppression strategies. These measures may not be accurate as they are not using QALYs adjusted for COVID-19, instead using valuations adjusted for pneumonia, which is only one of the potential complications of this disease [18]. Future research should explore the preferences of hospitalized patients and caregivers to fully understand the burden of COVID-19.

Declarations

Consent to participate Under a double opt-in process, respondents consented when they first joined the panel and on the first page of the survey instrument. Participants could leave the survey at any time by closing their internet browser or by leaving a comment.

Availability of data and material The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request.

Code availability The code generated in the study datasets is available from the corresponding author on reasonable request.

Consent for publication Not applicable.

Authors' contributions SP and BMC developed the research question, methodology, and design of the study as well as conducted the data analysis, interpreted the results, and wrote the final manuscript.

References

1. Devlin NJ, Brooks R. EQ-5D and the EuroQol group: past, present and future. Appl Health Econ Health Policy. 2017;15(2):127–37.
2. Gerlinger C, Bamber L, Leverkus F, Schwenke C, Haberland C, Schmidt G, et al. Comparing the EQ-5D-5L utility index based on value sets of different countries: impact on the interpretation of clinical study results. BMC Res Notes. 2019;12(1):18.
3. Augustovski F, Belizán M, Gibbons L, Reyes N, Stolk E, Craig BM, et al. Peruvian valuation of the EQ-5D-5L: a direct comparison of time trade-off and discrete choice experiments. Value Health. 2020;23(7):880–8.
4. Kiadaliri AA, Eliasson B, Gerdtham U-G. Does the choice of EQ-5D tariff matter? A comparison of the Swedish EQ-5D-3L index score with UK, US, Germany and Denmark among type 2 diabetes patients. Health Qual Life Outcomes. 2015;13:145.
5. Zhao Y, Li S-P, Liu L, Zhang J-L, Chen G. Does the choice of tariff matter? A comparison of EQ-5D-5L utility scores using Chinese, UK, and Japanese tariffs on patients with psoriasis vulgaris in Central South China. Medicine. 2017;96(34):e7840.
6. Craig BM, Rand K, Bailey H, Stalmeier PFM. Quality-adjusted life-years without constant proportionality. Value Health. 2018;21(9):1124–31.
7. Craig BM, Rand K. Choice defines QALYs: a US valuation of the EQ-5D-5L. Med Care. 2018;56(6):529–36.
8. Pickard AS, Law EH, Jiang R, Pullenayegum E, Shaw JW, Xie F, et al. United States valuation of EQ-5D-5L health states using an international protocol. Value Health. 2019;22(8):931–41.
9. Schmitz LPTC. LimeSurvey: an open source survey tool. Hamburg: LimeSurvey Project; 2012. http://www.limesurvey.org. Accessed 8 Mar 2021.
10. Jones A. Models For Health care, health, econometrics and data group (HEDG) Working Papers, HEDG, c/o Department of Economics, University of York; 2010. https://EconPapers.repec.org/RePEc:yor:hectdg:10/01
11. Manning WG, Mullahy J. Estimating log models: to transform or not to transform? J Health Econ. 2001;20(4):461–94.
12. McCaffrey N, Kaambwa B, Currow DC, Ratcliffe J. Health-related quality of life measured using the EQ-5D–5L: South Australian population norms. Health Qual Life Outcomes. 2016;14(1):133.
13. Willan AR, Briggs AH. Statistical analysis of cost-effectiveness data. New York: Wiley; 2006.
14. Bureau USC. American Community Survey 1-year public use microdata samples. 2019. https://www.census.gov/programs-surveys/acs/microdata.html. Accessed 1 Nov 2020
15. Jonker MF, Donkers B, de Bekker-Grob EW, Stolk EA. Advocating a paradigm shift in health-state valuations: the estimation of time-preference corrected QALY tariffs. Value Health. 2018;21(8):993–1001.
16. Pormohammad A, Ghorbani S, Khatami A, Razizadeh MH, Alborzi E, Zarei M, et al. Comparison of influenza type A and B with COVID-19: a global systematic review and meta-analysis on clinical, laboratory and radiographic findings. Rev Med Virol. 2020;e2179.
17. Yelin D, Wirtheim E, Vetter P, Kalil AC, Bruchfeld J, Runold M, et al. Long-term consequences of COVID-19: research needs. Lancet Infect Dis. 2020;20(10):1115–7.
18. Zala D, Mosweu I, Critchlow S, Romeo R, McCrone P. Costing the COVID-19 pandemic: an exploratory economic evaluation of hypothetical suppression policy in the United Kingdom. Value Health. 2020;23(11):1432–7.