

This is a repository copy of *Protocol for a longitudinal study examining the trajectory of COVID-19, post-COVID, multidimensional disadvantage and health-related quality of life in India:the IndiQol Project*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/219149/>

Version: Published Version

---

## Article:

Sinha, Kompal, Gutacker, Nils orcid.org/0000-0002-2833-0621, Gu, Yuanyuan et al. (3 more authors) (2024) Protocol for a longitudinal study examining the trajectory of COVID-19, post-COVID, multidimensional disadvantage and health-related quality of life in India:the IndiQol Project. *BMJ Open*. e080985. ISSN 2044-6055

<https://doi.org/10.1136/bmjopen-2023-080985>

---

## Reuse

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial (CC BY-NC) licence. This licence allows you to remix, tweak, and build upon this work non-commercially, and any new works must also acknowledge the authors and be non-commercial. You don't have to license any derivative works on the same terms. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>

## Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.

# BMJ Open Protocol for a longitudinal study examining the trajectory of COVID-19, post-COVID, multidimensional disadvantage and health-related quality of life in India: the IndiQol Project

Kompal Sinha <sup>1</sup>, Nils Gutacker <sup>2</sup>, Yuanyuan Gu <sup>3,4</sup>, Juanita Haagsma <sup>5</sup>, Kaushalendra Kumar <sup>6</sup>, Mona Aghdaee <sup>3,4</sup>

**To cite:** Sinha K, Gutacker N, Gu Y, *et al.* Protocol for a longitudinal study examining the trajectory of COVID-19, post-COVID, multidimensional disadvantage and health-related quality of life in India: the IndiQol Project. *BMJ Open* 2024;**14**:e080985. doi:10.1136/bmjopen-2023-080985

Received 19 October 2023  
Accepted 27 June 2024



© Author(s) (or their employer(s)) 2024. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

<sup>1</sup>Department of Economics, Macquarie Business School, Macquarie University, North Ryde, New South Wales, Australia

<sup>2</sup>Centre for Health Economics, University of York, York, UK

<sup>3</sup>Macquarie University Centre for the Health Economy, Macquarie Business School, Macquarie University, Macquarie Park, New South Wales, Australia

<sup>4</sup>Australian Institute of Health Innovation, Sydney, South Wales, Australia

<sup>5</sup>Department of Public Health, Erasmus MC, Rotterdam, the Netherlands

<sup>6</sup>International Institute for Population Sciences, Mumbai, Maharashtra, India

## Correspondence to

Dr Kompal Sinha;  
kompal.sinha@mq.edu.au

## ABSTRACT

**Introduction** The COVID-19 pandemic has raised concerns about the persistence of symptoms after infection, commonly referred to as ‘post-COVID’ or ‘long-COVID’. While countries in high-resource countries have highlighted the increased risk of disadvantaged communities, there is limited understanding of how COVID-19 and post-COVID conditions affect marginalised populations in low-income and middle-income countries. We study the longitudinal patterns of COVID-19, post-COVID symptoms and their impact on the health-related quality of life through the IndiQol Project.

**Methods and analysis** The IndiQol Project conducts household surveys across India to collect data on the incidence of COVID-19 and multidimensional well-being using a longitudinal design. We select a representative sample across six states surveyed over four waves. A two-stage sampling design was used to randomly select primary sampling units in rural and urban areas of each State. Using power analysis, we select an initial sample of 3000 household and survey all adult household members in each wave. The survey data will be analysed using limited dependent variable models and matching techniques to provide insights into the impact of COVID-19 pandemic and post-COVID on health and well-being of individuals in India.

**Ethics and dissemination** Ethics approval for the IndiQol Project was obtained from the Macquarie University Human Research Ethics Committee in Sydney, Australia and Institutional Review Board of Morsel in India. The project results will be published in peer-reviewed journals. Data collected from the IndiQol project will be deposited with the EuroQol group and will be available to use by eligible researchers on approval of request.

## INTRODUCTION

Socioeconomic disadvantage is associated with poor life expectancy and overall health.<sup>1</sup> The COVID-19 pandemic has disrupted lives and exacerbated health inequalities worldwide.<sup>2–3</sup> In the regions with higher socioeconomic deprivation, COVID-19-related

## STRENGTHS AND LIMITATIONS OF THIS STUDY

- ⇒ The project creates the first comprehensive longitudinal dataset on various aspects of well-being including health-related quality of life, socioeconomic conditions, COVID-19 history, post-COVID symptoms and beliefs about COVID-19 disease.
- ⇒ This population-based study surveys a representative sample of households from six states across India over four waves, thereby gathering evidence on the causal impact of COVID-19 and post-COVID on the health and well-being of individuals and their households.
- ⇒ The survey will be lengthy as it will collect household-level and individual-level self-reported data which can result in measurement error due to self-reporting bias.
- ⇒ The longitudinal nature of this survey would result in some attrition in later waves of the survey which is anticipated to affect the analysis and results of the study.

mortality rates are higher, and there are more cases reported than in less deprived regions.<sup>4–6</sup> Despite clear evidence linking deprivation and mortality, there is limited understanding of the causal impact of socioeconomic deprivation on the health-related quality of life (HRQoL) in COVID-19 survivors. This is particularly important since the long-term health consequences of COVID-19 are likely to be influenced by individuals’ social circumstances, such as their occupation or their access to healthcare services. The COVID-19 pandemic has hit low-income and middle-income countries (LMICs) hard resulting in significant economic shocks that may further exacerbate the disadvantages associated with the pandemic.<sup>7–8</sup> In India, where this study is situated, the pandemic has spread widely, with 98% of the country’s districts reported

COVID-19 patients in 2020.<sup>9</sup> India also has one of the highest rates of COVID-19 seroprevalence<sup>10</sup> and rates of COVID-19 deaths globally. Recent reports highlight the socioeconomic impact of the pandemic, revealing that less educated COVID-19 patients experience greater disease severity and mortality in India.<sup>11</sup>

Among those who survive the acute COVID-19 infection, some individuals experience persistent symptoms that can significantly impact their HRQoL over an extended period.<sup>12</sup> These persistent symptoms following a COVID-19 infection are termed as 'long-COVID' or 'post-COVID' ('*post-acute sequelae* of COVID-19'). The WHO has defined post-COVID as prolonged symptoms occurring approximately 3 months after the onset of COVID-19 symptoms, lasting for at least 2 months and not attributable to an alternative diagnosis.<sup>13</sup> Post-COVID symptoms can include a range of issues, including fatigue, shortness of breath and cognitive dysfunction, among others,<sup>14</sup> which may also fluctuate or relapse over time. While post-COVID is estimated to occur in 10%–20% of COVID-19 cases, it affects people of all ages, including children, with most cases occurring in patients with mild acute illness.<sup>15 16</sup> The incidence rates vary based on hospitalisation status and vaccination. For non-hospitalised cases, while the conservative estimate of post-COVID is 10%, the incidence is estimated to be 10%–30% for non-hospitalised cases; 50%–70% for hospitalised cases and 10%–12% for vaccinated cases.<sup>17</sup> In India, it is estimated that 6%–10% of COVID-19 patients go on to develop post-COVID.<sup>16</sup> Despite existing studies on post-COVID primarily focusing on patient populations in developed countries,<sup>5 18 19</sup> there has been no specific study on post-COVID in India.<sup>20</sup> India, as an LMIC country, has faced significant challenges due to the pandemic, with over 44.9 million COVID-19 patients reported as of October 2023 (according to the WHO website). Experts predict a potential intensification of COVID-19 cases in India.<sup>21 22</sup>

Socially disadvantaged communities have experienced higher COVID-19-related mortality and disease burden in less deprived areas<sup>2 3</sup> in developed (UK, USA)<sup>4</sup> and developing countries (India, China).<sup>6</sup> In the UK, the Office of National Statistics estimated 2 million people experienced long-COVID in April 2024, with prevalence being highest among people aged 35–69 years, women and people living in deprived areas.<sup>23</sup> In India, the economic shock, lockdown policies and mass migration amplified poverty and deprivation in the country leaving 121 million people out of work and 23 million people on streets. An issue of concern for public health is how the incidence and prevalence of disease symptoms among COVID-19 survivors living in deprivation. Persistent symptoms of COVID-19 among survivors can significantly diminish their quality of life for a long time. Moreover, with almost 40% of COVID-19 patients likely to experience post-COVID, the socioeconomic impact of the pandemic on multidimensional disadvantage and health inequalities will be substantial.

The condition of individuals facing hardships across multiple life domains, referred to as socioeconomic disadvantage, can influence health outcomes.<sup>24</sup> Recent literature has considered disadvantage as a multidimensional concept,<sup>24–27</sup> that is, multidimensional deprivation (MDD), based on the notion that deprived individuals should be identified using both income-based and non-income-based indicators. In India, a significant population of socioeconomically deprived people,<sup>28 29</sup> with approximately 27.5% of the population reported to be multidimensionally deprived in 2015–2016.<sup>30</sup> Specifically, 52% of households lacked proper sanitation, 12% lacked access to electricity, 10% did not have access to safe drinking water, 13% of individuals have less than 6 years of schooling and 37% had poor nutritional status during the same period.<sup>31</sup> The importance of considering the chronicity and persistence deprivation across multiple domains has been emphasised in the recent literature resulting in measures of MDD that account for the duration and persistent deprivation in multiple domains. The measure proposed by Nicholas *et al*<sup>26</sup> and Sinha *et al*<sup>24</sup> accounts for the following factors: (a) deprivation across a wide variety of dimensions/domains of life in each period; (b) deprivation for the most periods in a specific dimension and (c) chronic deprivation characterised with uninterrupted periods of deprivation. The duration and persistence of deprivation across multiple domains can lead to a health-deprivation trap.<sup>32</sup> To measure and understand the health inequalities ensuing from this disadvantage, it is important to untangle the contribution of each of the components of disadvantage, as well as their persistence and severity.<sup>26</sup> Thus, it is important to understand how severe health conditions, as experienced during COVID-19 and post-COVID, impact the pattern of MDD and how it influences the HRQoL and well-being of individuals.

### Aim of the research

The IndiQoL Project aims to examine morbidity associated with COVID-19, that is, post-COVID symptoms, multidimensional deprivation and HRQoL in India. Conducting a large-scale longitudinal household survey in India, this project will select a representative sample of the general population in India across six states, surveyed over four waves. While existing literature has analysed COVID-19 patients only, this population-based study will interview all adult household members on their COVID-19 history, socioeconomic conditions, HRQoL and behaviour. Individuals currently affected by or with a history of COVID-19 will be asked questions on their long-COVID symptoms. The first wave of the survey will include a prepandemic recall questionnaire, asking respondents about their socioeconomic circumstances and HRQoL during the annual festival period of Diwali 2019. The novelty of the current analysis lies in investigating the long-term consequences of the COVID pandemic by exploring the association between COVID-19, post-COVID conditions, multidimensional deprivation, household demographics,

socioeconomic status and HRQoL as measured by the EQ-5D-5L (using the recently developed value sets for India).<sup>33</sup>

## METHODS

### Study population

India is a large country with six broad regions—North, North-East, South, West, East and Central India. To ensure representation of the regional diversity of the country, the study population was determined by grouping the 29 Indian states into six categories based on a set of economic and health indicators. One state from each category was chosen for our sample.<sup>1</sup> Specifically, we have selected six states considering per capita net state domestic product,<sup>34</sup> child mortality (under 5 years old), life expectancy at birth, schooling and the social subclass of scheduled caste and scheduled tribe (SC/ST) population. The sampled population is drawn from six states: Uttar Pradesh, Odisha, Tamil Nadu, Haryana, Rajasthan and Maharashtra (figure 1). These six states together have around half (48%) of the total population of India.<sup>35</sup>

Table 1 shows the variation of our selection parameters of human development across the selected states and India including population growth, disability-adjusted life-years (DALYs), literacy rates and proportion of SC/ST, that is, socioeconomically disadvantaged groups. The six study states are at different stages of demographic and epidemiological transition. For instance, as per the Registrar General of India,<sup>36</sup> the natural growth rate of the population varies from 18.7% (crude birth rate, CBR=25.1, crude death rate, CDR=6.1) (CBR: the number of births in a year per thousand population; CDR: the number of deaths in a year per thousand population.) in Uttar Pradesh to 7.7% (CBR=13.8, CDR=6.1) in Tamil Nadu. Additionally, the value for the ratio of DALYs due to communicable, maternal, neonatal and nutritional diseases over DALYs due to non-communicable disease and injuries varies across these states: 0.26 in Tamil Nadu, 0.33 in Maharashtra, 0.40 in Haryana, 0.66 in Rajasthan, 0.58 in Odisha and 0.68 in Uttar Pradesh.<sup>37</sup>

### Household as the sampling unit

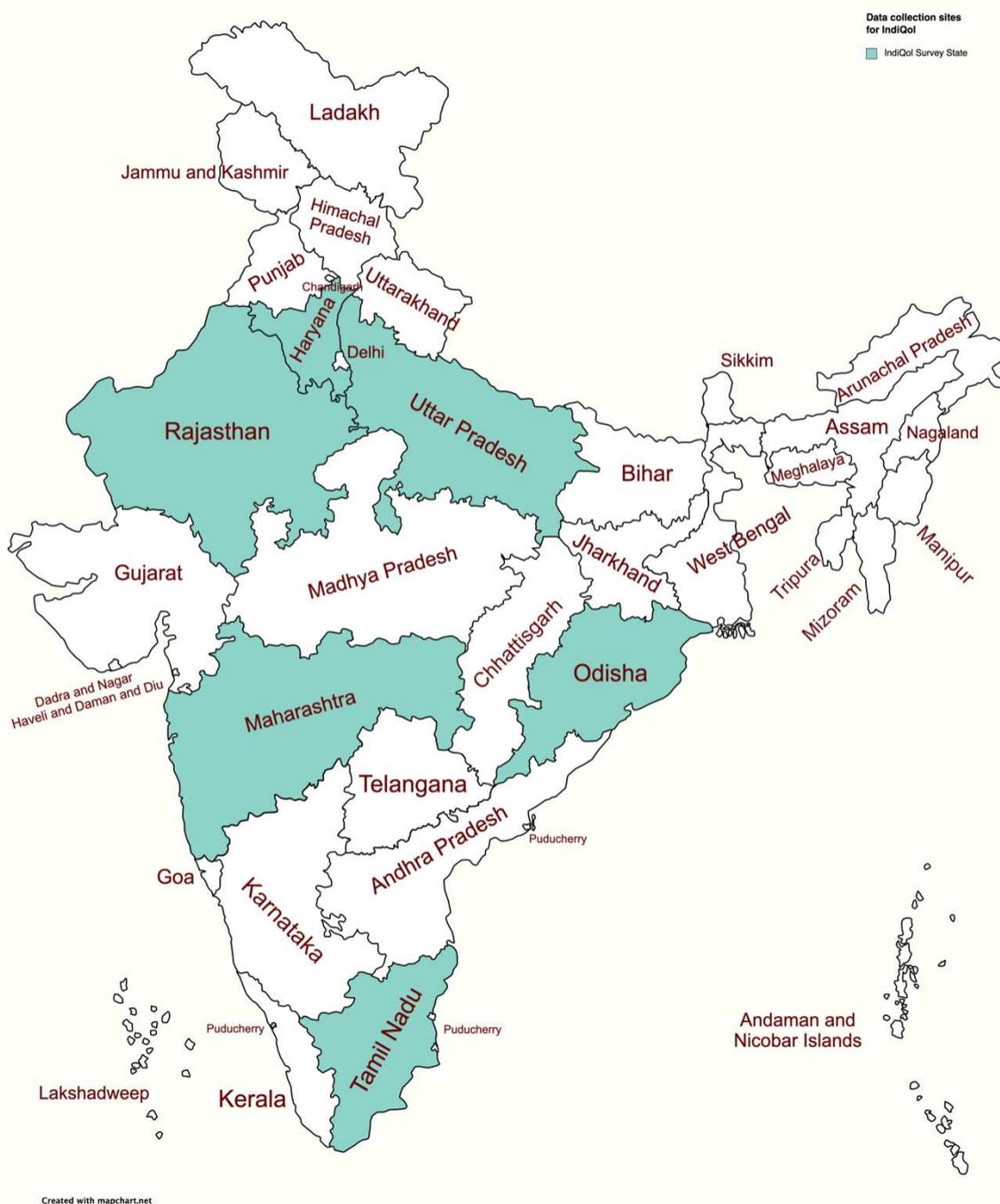
An important feature of this study is using the household as a sampling unit and interviewing all eligible (aged 18 years and older) members of the household. The motivation for considering a household rather than an eligible individual from the household lies in the contagious nature of the COVID-19 disease, post-COVID symptoms and the potential impact of changes in socioeconomic circumstances due to the pandemic. We categorise a household as a COVID-19 positive household (henceforth, COVID-positive household) if at least one member of the household reports having a history of COVID-19.

There are several advantages of the general population-based sampling approach. First, this approach allows us to assess the socioeconomic gradient in post-COVID conditions for both the overall population and the subgroup of

COVID-positive households. Second, since we analyse the association between multidimensional deprivation and experiencing COVID-19 and post-COVID, having non-COVID households in the sample allows us to benchmark our analysis against the COVID households. Third, given the highly contagious nature and rapid transmissibility of this disease, there is a high likelihood of intrahousehold transmission of COVID-19. With households losing family members, jobs and income due to COVID-19, experiencing post-COVID by an individual and/or other members of households could have vicarious effects on the health and well-being of all members of the household. It is critical to understand the variation in experiencing post-COVID symptoms and HRQoL of the members within a COVID-19 household (ie, intrahousehold HRQoL). Our approach to studying HRQoL for each member allows us to explore the vicarious mental and physical health effects of post-COVID within the household. Fourth, the impact of the pandemic on household's socioeconomic status would vary by the severity of disease incidence, that is, the number of household members experiencing post-COVID symptoms. For instance, the socioeconomic circumstances of a COVID-positive household where only one member lost their job due to COVID-19 and is now experiencing post-COVID would be different from that of a household where only one household member lost their job and multiple members of the household are experiencing post-COVID. Since the possibility of more than one family member of a COVID-positive household experiencing post-COVID is expected to be high, the impact on the HRQoL and well-being of the entire household will also be high. Finally, while most existing data on COVID-19 pandemic has collected data for COVID-19 survivors only, this population-based study will collect data on socioeconomic circumstances, COVID-19 history, post-COVID and HRQoL for the general population in India.

### Sample size

The sample size is determined based on the level of confidence and statistical power analysis following Cohen's<sup>38</sup> guidelines, along with the effect sizes reported in recent literature.<sup>29 39–42</sup> Our sample size accounts for the effect size (ie, the difference between the highest and the lowest level needed to detect a meaningful difference) between the HRQoL reported by deprived and not deprived groups with COVID-19 and post-COVID. The power calculation assumed a small effect size, 95% CI and a statistical power of 90%. Based on F-test analysis of variance (fixed effect, main and interaction effects,  $\alpha$  err prob=0.05, power=0.90, effect size=0.1 (small)), a sample size of 2063 was suggested (The power analysis was carried out using the G\*Power V.3.1.9.7 software). This provides sufficient power to detect statistical effects in multiple regression models.<sup>43–45</sup> The sample size is also validated using the effect sizes reported in recent literature for Middle Eastern and North Africa (MENA) countries,<sup>39</sup> Hong Kong,<sup>40</sup> India,<sup>29</sup> Asia<sup>41</sup> and China.<sup>42</sup> The project will



**Figure 1** The six states included in the sample design.

recruit a sample of 3000 randomly selected households, representative across age, gender and education ( $\geq 12\,000$  individuals based on an average household size of 4.7 in India), through Morsel India, a field survey company in India. All members of the household ( $>18$  years old) will be interviewed. Following the norm for survey data collection in household sample surveys in India,<sup>46</sup> we aim for a 70% response rate across waves. Assuming a 30% attrition rate, which aims to collect a 70% response rate for household sample surveys in India, an initial sample of 3000

households will suffice for a final sample of at least 2063 households.

### Sampling strategy

The sample design will follow the statistical approaches recommended for designing household surveys. Henderson and Sundaresan<sup>46</sup> demonstrated that 30 clusters, that is, primary sampling units (PSUs), selected based on probability proportional to size (PPS) from strata (Henderson and Sundaresan<sup>46</sup> created 12

**Table 1** Indicators of human development (2019–2020), by state

State	Per capita net state domestic product (INR 2019–2020)	Under-5 years mortality per thousand live births (2020)	Life expectancy at birth (2018)*	Literacy rates (per cent)†	Percentage SC/ST population
Uttar Pradesh	65 704	43	66.0	73	27.3
Odisha	110 434	39	70.3	77.3	45.0
Tamil Nadu	213 396	13	73.2	82.9	29.6
Haryana	247 628	33	69.9	80.4	30.9
Rajasthan	115 356	40	69.4	69.7	35.7
Maharashtra	202 130	18	72.9	84.8	28.6
India	132 115	32	70.0	77.7	31.4

Source: Sample Registration System Statistical Report (2020) Office of the Registrar General & Census Commissioner.<sup>63</sup> Sample Registration System Based Abridged Life table 2016–2020 and 2022.<sup>64</sup> International Institute for Population Sciences (IIPS) and ICF (2021).<sup>65</sup>

\*Based on average age-specific death rates for the year 2016–2020.

†Populations age 6 and over having completed 5 or more years of schooling.

SC, scheduled caste; ST, scheduled tribe.

hypothetical stratum by considering different levels of immunisation coverage in communities (ie, PSUs) of a geographical unit (ie, state), provide an estimate of immunisation coverage with 95% confidence limits. Following Henderson and Sundaresan,<sup>46</sup> we will adopt a two-stage sampling design strategy to select PSUs in both rural and urban areas of a state. A total of 30–32 PSUs in a state will be allocated and selected from rural and urban areas according to population proportions. Before selection, all PSUs with fewer than 40 households will be linked to the nearest PSU. We will form separate strata in rural and urban areas for household selection in a state.

#### Stratification and first stage selection of the sample

The purpose of stratification is to create groups within the population that have homogeneous characteristics within a stratum and heterogeneous characteristics between strata (population subgroups). For the selection of PSUs, we based our sampling frame on the 2011 census. In each state, we will create six explicit substrata by crossing three substrata based on the number of households in a PSU, and two substrata based on percentage of PSU's population belonging to the social subgroups, that is, SC/ST. Within each state, the PSUs will be sorted by the number of households, and three explicit strata of PSUs will be formed with equal numbers of households in each. Additionally, we will sort PSUs within each substratum by the percentage of SC/ST population, resulting in two substrata in each of the three strata.

#### Implicit stratification

Within each explicit strata, the PSUs will be sorted by female literacy rate of females aged 6 years and above. The pattern of this sorting will alternate between an ascending and descending order for subsequent strata. We will then aggregate the number of households in each PSU to adopt a PPS sampling method to select the 30 PSUs in a state.

#### Second stage selection

Sample households will come from each selected PSU. Selected PSUs with more than 150 households will be divided into segments (We will follow geographical boundaries such as waterways, roads and streets) containing 60–100 households. One segment will be randomly selected for mapping and listing operation in the PSU. A complete mapping of listing of each household in the PSU will be done to select households for interview. From each of the selected PSUs or segment of a PSU, 20 households will be randomly selected with systematic sampling.

#### Survey design

The survey will be conducted four times over 2023 and 2024, at equally spaced intervals, starting with winter of 2023 and ending in winter of 2024. The survey data collection team at Morsel India will conduct the survey in consultation with the project team. The survey questionnaire will be developed and conducted using computer-assisted personal interview software SurveyCTO (<https://www.surveycto.com/>). All interviewers will be trained by researchers who directly work with us. Each survey will take approximately 40–50 min using the official EQ-5D-5L (EuroQol health-related quality of life questionnaire) local language translations approved by the EQ-5D legal team. To maximise participation across all four surveys, phone numbers of respondents will be collected for the purpose of scheduling follow-up interviews. Respondents will be given a gift voucher (worth €2), at the end of follow-up surveys.<sup>47</sup> Attrition and non-response in the first follow-up will be appropriately tested following the best practices in the literature.<sup>48</sup>

The objective of collecting data over four waves (including one prepandemic recall) is twofold: (a) to monitor and evaluate the socioeconomic circumstances experienced by the population amid high rates of COVID-19 and post-COVID conditions and (b) to apply our theoretical framework for measuring

**Table 2** Core questionnaire modules, respondents and objective of each survey wave

Questionnaire* module	Respondents (aged ≥18 years)	Purpose	Topics and survey instruments
Household Questionnaire	Household head/main respondent	<ul style="list-style-type: none"> <li>▶ To record the household composition, size, location, roster (list of all members), socioeconomic characteristics, healthcare access, social capital</li> <li>▶ To compare household socioeconomic status</li> <li>▶ To identify individuals with COVID-19 history</li> </ul>	<ul style="list-style-type: none"> <li>▶ Demographic characteristics of the household members including age, gender, education level, caste/religion/social group (SC/ST/other)†</li> <li>▶ COVID-19 history</li> <li>▶ Vaccination history</li> <li>▶ Household assets</li> <li>▶ Income</li> </ul>
Individual questionnaire	All eligible‡ members of selected household	<ul style="list-style-type: none"> <li>▶ To evaluate the prevalence of COVID-19 history, COVID-19 vaccination history, post-COVID symptoms, morbidity history, health-related quality of life (EQ-5D-5L)</li> <li>▶ To recall health-related quality of life pre-COVID</li> <li>▶ To recall socioeconomic well-being pre-COVID</li> <li>▶ To identify COVID-19 stigma, post-COVID stigma</li> <li>▶ To document respondents' Disease history</li> <li>▶ To determine Individual-level socioeconomic disadvantage</li> <li>▶ To assess respondents' mental health, disorder</li> <li>▶ To measure alcohol and tobacco consumption habits</li> </ul>	<ul style="list-style-type: none"> <li>▶ EQ-5D-5L (EuroQoL health-related quality of life questionnaire).</li> <li>▶ Patient Health Questionnaire</li> <li>▶ General Anxiety Disorder</li> <li>▶ COVID-19 stigma (COVID PSS Scale)</li> <li>▶ Post-COVID, status, history and symptoms</li> <li>▶ Post-COVID stigma questionnaire</li> <li>▶ COVID-19-related symptoms</li> <li>▶ Fagerstrom test</li> <li>▶ Alcohol Use Disorder Identification Test</li> </ul>

\*Each household and eligible individuals will be surveyed four times at equally spaced intervals, over 16 months.

†Other backward class.

‡Only respondents with COVID-19 history will be eligible for questions about post-COVID symptoms.

SC, scheduled caste; ST, scheduled tribe.

of multidimensional deprivation,<sup>24</sup> which necessitates multiple waves to accurately measure the persistence and severity of multidimensional deprivation. Furthermore, since most post-COVID symptoms last at least 3 months, collecting data at equally spaced intervals will allow us to closely track health patterns and disease symptoms.

### Survey questionnaire

Each survey wave will collect data on the same set of individual and household-level socioeconomic circumstances, HRQoL, COVID-19 history and vaccination information. Detailed information on the data to be collected is presented in table 2. All members of the household (≥18 years old), including the elderly, disabled and low literacy or illiterate members, will be interviewed. The core dataset will include respondents' age, gender, educational attainment, employment status, occupation, social caste, household size, income and material well-being. Material well-being encompasses housing conditions, access to electricity, fuel, clean water, sanitation, vehicles, land and house ownership. We will also inquire about schooling, financial hardship during COVID-19 and social engagement. Respondents will be asked to report their HRQoL using the EQ-5D-5L instrument.<sup>49</sup> The resulting health profiles will be translated into index scores using the Indian EQ-5D-5L value set.<sup>50</sup> Additionally, respondents

will answer questions on chronic pre-existing health conditions, COVID history, smoking status (following the Fagerstrom test<sup>51</sup>), alcohol consumption (using the Alcohol Use Disorder Identification Test<sup>52</sup>), disability status (using the WHO disability assessment tool<sup>53</sup>) and anxiety (using the COVID-19 anxiety scale<sup>54</sup>). For respondents reporting a COVID-19 history, we will inquire about their post-COVID symptoms as established by the WHO.<sup>52</sup> Focusing on the overall population, rather than only COVID-positive households, aligns with the sampling approach used in ongoing studies in the USA and the UK; specifically, studies such as the RECOVER project (USA)<sup>55</sup> and the EuroQoL-funded POPCORN.<sup>56 57</sup> The POPCORN study investigated the direct and indirect effects of the COVID-19 pandemic on HRQoL in Greece, Italy, the Netherlands, the UK and the USA.

The first wave of the IndiQoL survey will include questions requiring respondents to recall their health and socioeconomic circumstances prior to the onset of the COVID-19 pandemic. To establish a common reference point, we will use Diwali 2019, a festival celebrated across the entire country, as the time point in the year before the COVID-19 pandemic began. To address potential bias associated with retrospective measurement, we will conduct a test–retest analysis using both retrospective and

prospective data, following recent literature. Specifically, we will use the test for concordance and Bland-Altman test.

Ensuring data quality is crucial. The survey company and our research team will independently monitor data quality after every survey wave. We will record the number of households contacted, the acceptance or rejection of participation invitations, and, where known, the reason for non-participation. We will also quantify the rate of non-response for each of the items of the questionnaire.

### Patient and public involvement

There is no patient and public involvement in the study.

### Empirical analysis

The primary objective of estimating the average impact of COVID-19 and post-COVID conditions on HRQoL, as measured by EQ-5D-5L, will be achieved by comparing individuals with post-COVID symptoms to two reference groups: (a) those who have never had COVID-19 and (b) those who have recovered from an acute infection. Since individuals can experience multiple symptoms that may persist, fluctuate or relapse over time,<sup>52</sup> longitudinal EQ-5D-5L data will be particularly valuable for characterising and capturing the longitudinal pattern of HRQoL effects. Our framework will analyse the individual EQ-5D-5L domain responses using ordered logistic regression. To account for observed respondent characteristics (such as age and gender) and prepandemic HRQoL (recall), we will employ either a matching approach or regression adjustments. The EQ-5D-5L index scores will use the value sets recently developed for India.<sup>33</sup> The longitudinal nature of the data also allows us to control for some unobserved confounding effects through fixed effects models. To address potential missing values in multiple responses or covariates in longitudinal clustered data, we will use imputation methods.<sup>58</sup> Additionally, to capture the heterogeneity of the impact of COVID-19,<sup>59–61</sup> we will conduct stratified analysis by age, sex and across six states to provide further insights into the domain wise impact of the number of post-COVID symptoms on HRQoL.

The second objective of this study is to evaluate how disadvantage in multiple domains of well-being affects the HRQoL impact of COVID-19 and post-COVID. We consider several domains based on the global multidimensional deprivation indicators.<sup>6</sup> Specifically, the domains include income, schooling, sanitation, electricity, housing conditions, unemployment, social interaction, social capital (trust, security), access to transport, access to healthcare, domestic violence and abuse, food insecurity, social isolation, indoor air pollution and clean drinking water. This objective will use an algorithm to measure persistence augmented multidimensional deprivation developed by Nicholas *et al.*<sup>26</sup> and Sinha *et al.*<sup>24</sup> Specifically, we adopt a regression framework to estimate HRQoL values (index scores; domain responses) as a function of variables (including multidimensional deprivation,

age, gender, household size, location, health behaviours, post-COVID status). We use generalised linear models (where the dependent variable is related to covariates through a specific link function) for the EQ-5D-5L value scores and ordered logit regressions for domains. We will not only evaluate how overall disadvantage impacts the HRQoL but also explore the specific role of severity and persistence of deprivation in each domain. Using the Shapley decomposition method,<sup>62</sup> we will compute the proportional contribution of the components of multidimensional deprivation (ie, severity and persistence) on the HRQoL effect associated with post-COVID. Separately accounting for these components in our regression analyses will allow us to identify how severity and persistence in specific dimensions of deprivation have affected HRQoL for post-COVID. Further ranking of subgroups (gender, rural/urban, COVID-19 status) based on this decomposition technique will allow clear identification of COVID-sensitive dimensions of deprivation specific to each subgroup that policy could target.

### Limitations

The survey will be lengthy as it will collect household-level and individual-level self-reported data, which can result in measurement error due to self-reporting bias. The longitudinal nature of this survey would result in some attrition in later waves of the survey, which is anticipated to affect the analysis and results of the study. The first wave of survey includes a module which requires respondents to recall their pre-COVID-19 HRQoL and socioeconomic circumstances. As these events are almost 3–4 years old, there is a possibility of recall bias.

### ETHICS AND DISSEMINATION

Ethics approval for this study was given by the Macquarie University Human Research Ethics Committee in Sydney, Australia and by the Institutional Review Board of Morsel India. The results from this project will be published in peer-reviewed journals. Data collected from the IndiQoL project will be deposited with the EuroQoL group and will be available to use by eligible researchers upon approval of request.

X Kompal Sinha @kompalsinha

**Acknowledgements** We gratefully acknowledge the support of Elly Stolk, reviewers at the EuroQoL Executive Committee, blinded reviewers of our grant application, editorial assistance by Rhonda Daniels and Nyamdavaa Byambadorj. We also thank the team at Morsel India for their partnership on this project.

**Contributors** Guarantor: KS is the guarantor and takes full responsibility for the finished work and conduct of study, has access to data and controlled the decision to publish. Conception of study: KS. Study design: KS, NG, Y-YG and JH. Drafting the study: KS. Writing: NG, Y-YG and JH contributed to writing, reviewing and editing. Sample design: KS and KK. Questionnaire design and ethics approval: KS and MA. All authors agreed to be accountable for all aspects of the work. English language AI technology was used to improve the clarity of the text.

**Funding** This study is funded by the EuroQoL Group, Netherlands under the grant number 430-RA and Macquarie University, Sydney, Australia.

**Map disclaimer** The inclusion of any map (including the depiction of any boundaries therein), or of any geographic or locational reference, does not imply

the expression of any opinion whatsoever on the part of BMJ concerning the legal status of any country, territory, jurisdiction or area or of its authorities. Any such expression remains solely that of the relevant source and is not endorsed by BMJ. Maps are provided without any warranty of any kind, either express or implied.

**Competing interests** None declared.

**Patient and public involvement** Patients and/or the public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the Methods section for further details.

**Patient consent for publication** Not applicable.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

#### ORCID iDs

Kompal Sinha <http://orcid.org/0000-0003-4318-6100>

Nils Gutacker <http://orcid.org/0000-0002-2833-0621>

Yuanyuan Gu <http://orcid.org/0000-0002-3816-9106>

Juanita Haagsma <http://orcid.org/0000-0002-2055-548X>

Kaushalendra Kumar <http://orcid.org/0000-0001-5913-0297>

Mona Aghdaee <http://orcid.org/0000-0002-2570-1685>

#### REFERENCES

- Schneider P, Love-Koh J, McNamara S, et al. Decomposing the socioeconomic gradient of health-related quality of life in England: regression analysis of health survey data. *J Epidemiol Community Health* [Preprint] 2021.
- Cash R, Patel V. Has COVID-19 subverted global health? *Lancet* 2020;395:1687–8.
- Marmot M. Health equity in England: the Marmot review 10 years on. *BMJ* 2020;368:m693.
- Iacobucci G. COVID-19: deprived areas have the highest death rates in England and Wales. *BMJ* 2020;m1810.
- Kim D. Who gets long COVID and suffers its mental health and socioeconomic consequences in the United States? Preliminary findings from a large nationwide study. *medRxiv* [Preprint] 2023.
- Upshaw TL, Brown C, Smith R, et al. Social determinants of COVID-19 incidence and outcomes: a rapid review. *PLoS ONE* 2021;16:e0248336.
- Budhiraja S, Aggarwal M, Wig R, et al. Long term health consequences of COVID-19 in hospitalized patients from North India: a follow up study of Upto 12 months. *medRxiv* [Preprint] 2021.
- Mukhra R, Krishan K, Kanchan T. COVID-19 sets off mass migration in India. *Arch Med Res* 2020;51:736–8.
- Acharya R, Porwal A. A vulnerability index for the management of and response to the COVID-19 epidemic in India: an ecological study. *Lancet Glob Health* 2020;8:e1142–51.
- Murhekar MV, Bhatnagar T, Selvaraju S, et al. SARS-CoV-2 antibody seroprevalence in India, August–September, 2020: findings from the second nationwide household serosurvey. *Lancet Glob Health* 2021;9:e257–66.
- Sharma AK, Gupta R, Baig VN, et al. Educational status and COVID-19 related outcomes in India: hospital-based cross-sectional study. *BMJ Open* 2022;12:e055403.
- Malik P, Patel K, Pinto C, et al. Post-acute COVID-19 syndrome (PCS) and health-related quality of life (HRQoL)—a systematic review and meta-analysis. *J Med Virol* 2022;94:253–62.
- World Health Organization. A clinical case definition of post COVID-19 condition by a Delphi consensus. 2021. Available: [https://www.who.int/publications/i/item/WHO-2019-nCoV-Post\\_COVID-19\\_condition-Clinical\\_case\\_definition-2021.1](https://www.who.int/publications/i/item/WHO-2019-nCoV-Post_COVID-19_condition-Clinical_case_definition-2021.1) [Accessed 28 Feb 2023].
- Raveendran AV, Jayadevan R, Sashidharan S. Long COVID: an overview. *Diabetes Metab Syndr* 2021;15:869–75.
- Ledford H. Long COVID is a double curse in low-income nations—here's why. *Nature New Biol* 2024;625:20–2.
- Arjun MC, Singh AK, Roy P, et al. Long COVID following Omicron wave in Eastern India—a retrospective cohort study. *J Med Virol* 2023;95:e28214.
- Davis HE, McCorkell L, Vogel JM, et al. Long COVID: major findings, mechanisms and recommendations. *Nat Rev Microbiol* 2023;21:133–46.
- Bryson WJ. Long-term health-related quality of life concerns related to the COVID-19 pandemic: a call to action. *Qual Life Res* 2021;30:643–5.
- Tsuzuki S, Miyazato Y, Terada M, et al. Impact of long-COVID on health-related quality of life in Japanese COVID-19 patients. *medRxiv* [Preprint] 2021.
- Hindustan Times. Long COVID: range of symptoms. Hindustan Times. 2022. Available: <https://www.hindustantimes.com/india-news/long-covid-range-of-symptoms-medical-burden-101639894995286.html>
- Aljazeera: India's Omicron wave may intensify in coming weeks. Available: <https://www.aljazeera.com/news/2022/1/24/india-omicron-wave-may-intensify-coming-weeks-covid-experts> [Accessed 24 Jan 2022].
- Johns Hopkins University. COVID-19 map. Johns Hopkins Coronavirus Resource Center; 2021. Available: <https://coronavirus.jhu.edu/map.html> [Accessed Feb 2022].
- Ayuobkhani D. Prevalence of ongoing symptoms following coronavirus (COVID-19) infection in the UK. Office of National Statistics; 2021. Available: <https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/prevalenceofongoingsymptomsfollowingcoronaviruscovid19infectionintheuk/1april2021>
- Sinha K, Davillas A, Jones AM, et al. Do socioeconomic health gradients persist over time and beyond income? A distributional analysis using UK biomarker data. *Econ Hum Biol* 2021;43:101036.
- Alkire S, Kanagaratnam U, Suppa N. The global multidimensional poverty index (MPI) 2019, OPHI MPI methodological note 47. Oxford Poverty and Human Development Initiative, University of Oxford; 2019. Available: [https://www.ophi.org.uk/wp-content/uploads/OPHI\\_MPI\\_MN\\_47\\_2019\\_vs2.pdf](https://www.ophi.org.uk/wp-content/uploads/OPHI_MPI_MN_47_2019_vs2.pdf)
- Nicholas A, Ray R, Sinha K. Differentiating between dimensionality and duration in multidimensional measures of poverty: methodology with an application to China. *Rev Income Wealth* 2019;65:48–74.
- Ray R, Sinha K. Multidimensional deprivation in China, India and Vietnam: a comparative study on micro data. *J Hum Dev Capabil* 2015;16:69–93.
- Alkire S, Oldiges C, Kanagaratnam U. Examining multidimensional poverty reduction in India 2005/6–2015/16: insights and oversights of the headcount ratio. *World Dev* 2021;142:105454.
- Hegde S, Sreeram S, Bhat KR, et al. Evaluation of post-COVID health status using the EuroQol-5D-5L scale. *Pathog Glob Health* 2022;116:498–508.
- Pathak D, Vasishtha G, Mohanty SK. Association of multidimensional poverty and tuberculosis in India. *BMC Public Health* 2021;21:2065.
- Das P, Ghosh S, Paria B. Multidimensional poverty in India: a study on regional disparities. *GeoJournal* 2022;87:3987–4006.
- Love-Koh J, Schneider P, McNamara S, et al. Decomposition of quality-adjusted life expectancy inequalities by mortality and health-related quality of life dimensions. *Pharmacoeconomics* 2023;41:831–41.
- Jyani G, Yang Z, Sharma A, et al. Evaluation of EuroQol valuation technology (EQ-VT) designs to generate national value SETS: learnings from the development of an EQ-5D value set for india using an extended design (DEVINE) study. *Med Decis Making* 2023;43:692–703.
- RBI. *Handbook of statistics on Indian States*. Reserve Bank of India, 2022. Available: <https://m.rbi.org.in/scripts/AnnualPublications.aspx?head=Handbook+of+Statistics+on+Indian+State>
- Ministry of Health and Family Welfare. Population projections for India and states 2011–2036, report of the technical group on population projections, national commission on population. Ministry of Health and Family Welfare; 2020.
- RGI. SRS bulletin, sample registration system, office of the registrar general of India, New Delhi. 2022.
- Dandona L, Dandona R, Kumar GA, et al. Nations within a nation: variations in epidemiological transition across the States of India, 1990–2016 in the global burden of disease study. *Lancet* 2017;390:2437–60.
- Cohen J. *Statistical power analysis for the behavioural sciences*. 2nd edn. Hillsdale, 1988.
- Al Dhaheri AS, Bataineh MF, Mohamad MN, et al. Impact of COVID-19 on mental health and quality of life: is there any effect? A cross-sectional study of the MENA region. *PLoS One* 2021;16:e0249107.
- Choi EPH, Hui BPH, Wan EYF, et al. Covid-19 and health-related quality of life: a community-based online survey in Hong Kong. *Int J Environ Res Public Health* 2021;18:3228.
- Lim SL, Woo KL, Lim E, et al. Impact of COVID-19 on health-related quality of life in patients with cardiovascular disease: a multi-ethnic Asian study. *Health Qual Life Outcomes* 2020;18:387.

- 42 Qu G, Zhen Q, Wang W, *et al.* Health-related quality of life of COVID-19 patients after discharge: a multicenter follow-up study. *J Clin Nurs* 2021;30:1742–50.
- 43 Faul F, Erdfelder E, Lang A-G, *et al.* G\*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods* 2007;39:175–91.
- 44 Green SB. How many subjects does it take to do a regression analysis. *Multivariate Behav Res* 1991;26:499–510.
- 45 Wilson Van Voorhis CR, Morgan BL. Understanding power and rules of thumb for determining sample sizes. *TQMP* 2007;3:43–50.
- 46 Henderson RH, Sundaresan T. Cluster sampling to assess immunization coverage: a review of experience with a simplified sampling method. *Bull World Health Organ* 1982;60:253–60.
- 47 McGovern ME, Canning D, Bärnighausen T. Accounting for non-response bias using participation incentives and survey design: an application using gift vouchers. *Econ Lett* 2018;171:239–44.
- 48 Ghanem D, Hirshleifer S, Ortiz-Becerra K. Testing attrition bias in field experiments. CEGA WOR king paper series no.WPS-113. Center for Effective Global Action. University of California, Berkeley; 2021. Available: <https://doi.org/10.26085/C38C76>
- 49 Herdman M, Gudex C, Lloyd A, *et al.* Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res* 2011;20:1727–36.
- 50 Jyani G, Sharma A, Prinja S, *et al.* Development of an EQ-5D value set for India using an extended design (DEVINE) study: the Indian 5-level version EQ-5D value set. *Value Health* 2022;25:1218–26.
- 51 Heatherton TF, Kozlowski LT, Frecker RC, *et al.* The Fagerstrom test for nicotine dependence: a revision of the Fagerstrom tolerance questionnaire. *Br J Addict* 1991;86:1119–27.
- 52 World Health Organization. *AUDIT: the alcohol use disorders identification test: guidelines for use in primary health care*. World Health Organization, 2001.
- 53 Ustün TB, Chatterji S, Kostanjsek N, *et al.* Developing the world health organization disability assessment schedule 2.0. *Bull World Health Organ* 2010;88:815–23.
- 54 Lim SL, Woo KL, Lim E, *et al.* Impact of COVID-19 on health-related quality of life in patients with cardiovascular disease: a multi-ethnic asian study. *Health Qual Life Outcomes* 2020;18:387:387:.
- 55 Perlis RH, Santillana M, Ognyanova K, *et al.* Prevalence and correlates of long COVID symptoms among US adults. *JAMA Netw Open* 2022;5:e2238804.
- 56 Long D, Haagsma JA, Janssen MF, *et al.* Health-related quality of life and mental well-being of healthy and diseased persons in 8 countries: does stringency of government response against early COVID-19 matter? *SSM Popul Health* 2021;15:100913.
- 57 Long D, Bonsel GJ, Lubetkin EI, *et al.* Health-related quality of life and mental well-being during the COVID-19 pandemic in five countries: a one-year longitudinal study. *J Clin Med* 2022;11:6467.
- 58 Schafer JL, Yucel RM. Computational strategies for multivariate linear mixed-effects models with missing values. *J Comput Graph Stat* 2002;11:437–57.
- 59 Jaffrelot C, Thakker H. COVID-19, amplifying the return of mass poverty in India. 2020. Available: <https://www.institutmontaigne.org/en/expressions/covid-19-amplifying-return-mass-poverty-india>
- 60 Rao GV, Gella V, Radhakrishna M, *et al.* Post-COVID-19 symptoms are not uncommon among recovered patients—a cross-sectional online survey among the Indian population. *medRxiv* [Preprint] 2021.
- 61 Costa-Font J, Hernández-Quevedo C. Measuring inequalities in health: what do we know? What do we need to know *Health Policy* 2012;106:195–206.
- 62 Shorrocks AF. Decomposition procedures for distributional analysis: a unified framework based on the shapley value. *J Econ Inequal* 2013;11:99–126.
- 63 Sample registration system statistical report. New Delhi Office of the Registrar General & Census Commissioner, Govt. of India; 2020.
- 64 Sample registration system based abridged life tables 2016-20. New Delhi Office of the Registrar General & Census Commissioner, Govt. of India; 2022.
- 65 International Institute for Population Sciences (IIPS) and ICF. National family health survey (NFHS-5), 2019-21. India, Mumbai IIPS; 2021.