

Responsive population-based cohorts as platforms for characterising pathogen- and population-level infection dynamics for epidemic prevention, preparedness and response

Ivonne Morales^{1,2}, Van K  nh Nguyen^{2,3}, Mirna Abd El Aziz¹, Ayten Sultani^{1,4}, Till B  rnighausen^{2,5,6}, Heiko Becher², Sandra Ciesek^{7,8,9}, Beate Kampmann^{10,11}, Berit Lange^{12,13}, Jan Rupp^{14,15}, Simone Scheithauer¹⁶, Helen Ward^{3,17,18,19}, Andr   Karch²⁰, Claudia M Denkinger^{1,21}

1. Department of Infectious Disease and Tropical Medicine, Center for Infectious Diseases, Heidelberg University Hospital, Heidelberg, Germany
2. Heidelberg Institute for Global Health (HIGH), Heidelberg University Hospital, Heidelberg, Germany
3. School of Public Health, Imperial College London, London, United Kingdom
4. Institute of Tropical Medicine, Travel Medicine, and Human Parasitology, University Hospital T  bingen, T  bingen, Germany
5. Department of Global Health and Population, Harvard TH Chan School of Public Health, Harvard University, Boston, United States
6. Africa Health Research Institute, KwaZulu-Natal, South Africa
7. Institute for Medical Virology, University Hospital Frankfurt, Goethe University Frankfurt, Frankfurt am Main, Frankfurt, Germany
8. German Center for Infection Research (DZIF), Frankfurt am Main, Frankfurt, Germany
9. Fraunhofer Institute for Translational Medicine and Pharmacology ITMP, Frankfurt am Main, Frankfurt, Germany
10. The Vaccine Centre, Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom
11. Medical Research Council Unit The Gambia, London School of Hygiene and Tropical Medicine, Fajara, The Gambia
12. Department of Epidemiology, Helmholtz Centre for Infection Research (HZI), Braunschweig, Germany
13. German Center for Infection Research (DZIF), Braunschweig, Germany
14. Department of Infectious Diseases and Microbiology, University Hospital Schleswig-Holstein, Campus L  beck, L  beck, Germany
15. German Center for Infection Research (DZIF), Partner Hamburg-L  beck-Borstel-Riems, Germany
16. Department of Infection Control and Infectious Diseases, University Medical Center G  ttingen (UMG), Georg-August University G  ttingen, G  ttingen, Germany
17. Imperial College Healthcare NHS Trust, London, United Kingdom
18. National Institute for Health Research Imperial Biomedical Research Centre, London, United Kingdom
19. MRC Centre for Global Infectious Disease Analysis and Abdul Latif Jameel Institute for Disease and Emergency Analytics, Imperial College London, London, United Kingdom
20. Institute of Epidemiology and Social Medicine, University of M  nster, Albert-Schweitzer-Campus 1, M  nster, Germany
21. German Center for Infection Research (DZIF), Partner Heidelberg, Heidelberg, Germany

Correspondence: Ivonne Morales (ivonne.morales@uni-heidelberg.de)

Citation style for this article:

Morales Ivonne, Nguyen Van K  nh, Abd El Aziz Mirna, Sultani Ayten, B  rnighausen Till, Becher Heiko, Ciesek Sandra, Kampmann Beate, Lange Berit, Rupp Jan, Scheithauer Simone, Ward Helen, Karch Andr  , Denkinger Claudia M. Responsive population-based cohorts as platforms for characterising pathogen- and population-level infection dynamics for epidemic prevention, preparedness and response. *Euro Surveill.* 2025;30(25):pii=2400255. <https://doi.org/10.2807/1560-7917.ES.2025.30.25.2400255>

Article received on 25 Apr 2024 / Accepted on 13 Feb 2025 / Published on 26 Jun 2025

Establishing population-based cohorts is indispensable for effective epidemic prevention, preparedness and response. Existing passive surveillance systems face limitations in their capacity to promptly provide representative data for estimating disease burden and modelling disease transmission. This perspective paper introduces a framework for establishing a dynamic and responsive nationally representative population-based cohort, with Germany as an example country. We emphasise the need for comprehensive demographic representation, innovative strategies to address participant attrition, efficient data collection and testing using digital tools, as well as novel data integration and analysis methods. Financial considerations and cost estimates for cohort establishment are discussed, highlighting potential cost savings through integration with existing research infrastructures and

digital approaches. The framework outlined for creating, operating and integrating the cohort within the broader epidemiological landscape illustrates the potential of a population-based cohort to offer timely, evidence-based insights for robust public health interventions during both epidemics and pandemics, as well as during inter-epidemic periods.

Background

Timely detection of and response to emerging pathogens is central to epidemic preparedness. [1]. To support preparedness efforts, public health surveillance systems routinely collect and analyse clinical and epidemiological indicators to understand infectious disease patterns, deriving parameters like the incubation period and the basic reproductive number (R_0), used in predicting disease spread and assessing public health

countermeasures [2]. Typically, the data used for these metrics come from case-based notification systems, resulting in reporting delays and incomplete population representation [2]. In addition, digital sources like social media and alert networks can provide near real-time information and may serve as a supplementary tool for disease surveillance. However, they often lack the specificity needed for decision-making [3]. To address these gaps, prospective population-based cohorts, in which a sample of the population is selected and followed over time [1], can complement routine and digital surveillance by providing a comprehensive dataset for actively identifying symptomatic, asymptomatic and minimally symptomatic individuals and their contacts. These cohorts are invaluable for understanding the characteristics of (emerging) infections, providing more accurate estimates of disease transmission parameters, death rates and clinical epidemiologic metrics [4-6]. Population-based cohorts offer substantial added value—not only during an epidemic but also in the periods between them, when they can actively contribute to prevention. For example, cohort infrastructures can be used to monitor behaviour change and emerging risk patterns (e.g. hygiene practices or social norms related to disease transmission). This makes cohorts a strategic long-term investment—not just for preparedness, but for prevention. Beyond infectious disease surveillance, cohorts can help monitor well-being and evaluate public acceptance of health measures [7], providing early insights into their effectiveness and the need for further interventions. Indeed, the COVID-19 pandemic underscored how epidemics can disrupt people's daily lives in multiple ways, impacting their social, mental and physical well-being [8,9].

Population-based cohorts have played a crucial role in understanding the incidence, dynamics and determinants of infectious diseases, including endemic and seasonal pathogens such as influenza virus or respiratory syncytial virus (RSV), as well as emerging threats like severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The metrics we can obtain from population-based cohorts are vital for designing and monitoring targeted public health interventions for pandemic prevention and preparedness. Examples of population-based cohorts include the UK REal-time Assessment of Community Transmission (REACT) studies, the German ELISA cohort for SARS-CoV-2, and the Latin American cohorts for influenza [4-6,10]. The REACT studies addressed key knowledge gaps during the COVID-19 pandemic. Initiated early on, they delivered timely high-resolution data on infection prevalence, transmission dynamics, emerging variants, vaccine effectiveness and geographic spread, directly informing government decisions and supporting rapid public health action, e.g. the timing of lockdowns, school closure policies and social gatherings [4,10].

Creation of such a cohort, however, is time-consuming and is often complicated by multiple bottlenecks.

For instance, the 2015–16 Zika virus epidemic illustrated major challenges in funding, regulatory approvals, logistics and patient recruitment [11]. Although research consortia like ZIKAlliance (<https://zikalliance.tghn.org>) mobilised quickly, these bottlenecks delayed coordinated research efforts considerably, taking over 8 months to establish protocols. This delay meant that the peak of the epidemic had already passed in many areas before data collection could begin, severely limiting the ability to assess the full impact of Zika virus infection, particularly its association with congenital abnormalities in infants born to exposed pregnant women. An established cohort infrastructure could have enabled researchers to capture critical early-stage data, improved the precision of congenital anomaly risk estimates, and delivered timely evidence to guide public health responses and inform care for pregnant women and their families. Similarly, the development of the German National Pandemic Cohort Network (NAPKON), created to support COVID-19 research at the national level, initiated patient enrolment in the fall of 2020, once the pandemic was already well underway in the country. A key implementation challenge involved the organisation of study sites and the timely recruitment of participants, with a median time of 54 days from ethics approval to first patient enrolment. This study highlighted the critical importance of early study site preparation to minimise delays in patient recruitment [12]. These lessons underscore the need for responsive population-based cohorts as an integral part of pandemic preparedness.

In the absence of a representative cohort, a network of smaller population panels could serve as a provisional solution, which can later be ideally integrated into a robust and responsive framework. While synthesising evidence across multiple smaller population panels is possible, such efforts are often hampered by difficulties in data harmonisation, constraints on data sharing [13], and the need for methods to reconcile different study designs [13,14], limiting the power and timeliness of analyses.

In Germany, however, the IMMUNEBRIDGE project has shown that such panels can yield timely, harmonised and model-usable estimates of SARS-CoV-2 seroprevalence [15]. The effectiveness of such a strategy in future health crises will depend on the ability to harmonise broader clinical and epidemiological outcomes, and on the adaptability of existing panels to respond swiftly to emerging pathogens, while ensuring representativeness through the inclusion of dynamic and underserved populations, such as migrants and refugees.

An ideal population-based cohort should be agile and adaptive, capable of monitoring various infectious diseases and selectively activating or expanding as needed by time, location and changes in population composition. Moreover, it should be representative to support robust and accurate estimation of clinical and

TABLE

Example of the estimated state-level population-representative sample size in units of households or individuals, Germany

Diagnostic test characteristics		Per German state				All 16 German states	
		Target cohort size ^a		Number of households to invite ^b		Number of households ^c	
Sensitivity	Specificity	Individuals	Households	Survey 1	Survey 1+	Target	Invite
0.7	0.7	50,802	36,577	146,310	14,631	585,239	2,340,956
0.7	0.8	25,044	18,032	72,127	7,213	288,507	1,154,028
0.7	0.9	10,113	7,281	29,125	2,913	116,502	466,007
0.7	1	545	392	1,570	157	6,278	25,114
0.8	0.7	32,573	23,453	93,810	9,381	375,241	1,500,964
0.8	0.8	17,455	12,568	50,270	5,027	201,082	804,326
0.8	0.9	7,492	5,394	21,577	2,158	86,308	345,231
0.8	1	477	343	1,374	137	5,495	21,980
0.9	0.7	22,662	16,317	65,267	6,527	261,066	1,044,265
0.9	0.8	12,870	9,266	37,066	3,707	148,262	593,050
0.9	0.9	5,783	4,164	16,655	1,666	66,620	266,481
0.9	1	424	305	1,221	122	4,884	19,538

^a The sample size calculation of the target cohort size of individuals is based on a hypothetical emerging disease with a low prevalence of 1%, assuming a 95% confidence interval with a half-width of 1%, considering a range of typical sensitivity and specificity values of diagnostic tests, recognising that lower sensitivity or specificity requires a larger sample size. The number of target households was obtained by multiplying the target sample size of individuals by the square of the design effect factor of 1.2 to reflect the reduced statistical efficiency of household-based sampling compared to simple random sampling and divided by the number of individuals per household (considering an average of 2.0 individuals per household (2021) [43]).

^b Survey 1 shows the number of households to invite considering an overall response rate of 25%, i.e. on average half of the invited household responded (50%), within each household half of the members decide to join (50%). Survey 1+ denotes the number of newly recruited participants needed per survey to maintain the cohort's representativeness due to attrition and changes in the population, considering a replacement of 10% per survey.

^c The total number of households to invite across all 16 German federal states is calculated by multiplying the number of households to invite per state (from the Survey 1 column) by 16. Similarly, the total target number of households is obtained by multiplying the per-state household target by 16.

epidemiological parameters. In this Perspective, we consider methodological and financial aspects that underpin a cohort's functionality: sample representativeness, cohort size and sampling frequency, strategies to address participant attrition, data collection and testing procedures. We present examples and discuss the challenges of establishing a long-term nationally representative cohort and possible solutions using the German experience.

Representativeness

For a cohort to be representative of a country's population, cohort enrolment needs to be based on random processes. In many contexts, single-stage random selection from the full population sampling frame is typically not feasible, even though it would statistically be the most efficient sampling process. Instead, higher levels of social organisation, typically geographically contiguous communities, are sampled in a first stage, before families or individuals are sampled within the randomly selected communities. These communities often coincide with first or second administrative unit levels in a country and should be considered for sampling, as national sampling frames are usually available for these units [4]. At the first administrative level, for example, Germany has 16 federal states, each with varying degrees of heterogeneity in population density and age distribution, as well as in governance structure, sociodemographic characteristics and level of urbanisation. Because age-specific patterns of

infection are important for disease transmission and adaptation of interventions, the cohort should include individuals of all ages, i.e. children, adolescents, adults and elderly people, through an appropriate sampling strategy. For instance, a household-based random sampling approach (probability-based), could effectively represent participants of all ages if the sampling frame can be readily defined through resident registries or census data. An example of a population cohort employing this approach is the longitudinal cohort for influenza in Peru, which included participants from randomly selected households, ranging in age from under 2 years to over 65 years [16]. Using this strategy, the cohort achieved representativeness at the level of each of four ecologically distinct regions in the country. Alternatively, a sampling approach based on population density could be employed when up-to-date census data are lacking, as was done in Mozambique [17]. In Germany and beyond, established panels that are representative of the general population include the German Social Science Infrastructure Services (GESIS) Panel [18], the German Socioeconomic Panels (SOEP) [19], and the LISS (Longitudinal Internet studies for the Social Sciences) panel in the Netherlands [20]. However, these panels are often limited to narrower age groups, geographic coverage, or originally designed for purposes other than epidemic monitoring.

Representation of minority and dynamic populations, such as refugees and immigrants, is also essential,

as these groups may be disproportionately burdened compared with the rest of the population [21]. Similar to the German Socioeconomic Panels (SOEP), targeted sampling as well as frequent refreshment sampling of minority groups, e.g. NU(M)KRAINE in Germany [22], can be conducted to ensure cohort representativeness over time [19].

Cohort size and sampling frequency

To estimate the cohort size of a population-representative cohort, the following key parameters must be considered: (i) the level of disease prevalence with its degree of uncertainty, (ii) the expected sensitivity and specificity of the diagnostic test for case detection, (iii) the response rate and (iv) the study's design effect [23]. The Table shows the estimated cohort sample size for Germany when these parameters are considered across various scenarios. For a detailed description of the calculations, see Supplementary File S1. These estimates are approximate and would require further refinement during implementation. The sample size can be adjusted for questions (disease prevalence in a more localised geographical region or among elderly people, for example) that either impact specific populations or exhibit higher prevalence rates.

Defining the sampling frequency is a challenging task that requires striking a balance between disease detection at the desired prevalence level, costs and participant burden. One possible approach is to integrate representative data from a population-representative cohort with more frequently collected surveillance data, i.e. such as the Early Warning and Response System of the Robert Koch Institute (German national public health institute), into mathematical models that allow, for example, estimating changes and trends over time [24]. In this scenario, data from a population-based cohort would provide the anchor points of the true prevalence for the model estimates, whereas routinely collected surveillance data would inform model estimates of changes and trends over time.

Response rate, participant retention, sample refreshment and rotational design

Although a response rate of 50% at either the household or individual level is used in our example in the Table, higher response rates are essential to ensure the representativeness of estimated metrics. In practice, however, lower participation rates are common. Large-scale cohort and panel studies, such as the German National Cohort (NAKO) and SOEP in Germany or the United Kingdom's REACT cross-sectional surveys, have reported varying response rates. For instance, NAKO reported response rates ranging from 9 to 32%, depending on the recruiting study centre [25]. The SOEP has reported rates between 31.5 and 80.8%, varying by population type and wave of recruitment or refreshment sample [26]. Similarly, the REACT-1 study reported response rates ranging from 11.7 to 30.5% depending on the round of data collection [4]. Therefore, sensitisation of the population to the utility of cohorts and

participation in research, as well as their engagement throughout the life of the cohort will be crucial to maximising response rates and retention. Community outreach and participant engagement through advertisement of the study, community events, healthcare providers and the media have been shown to improve participant recruitment [27].

Participant retention in a cohort is essential to maintaining the cohort's representativeness. Participant attrition is, in part, influenced by age and can be intensified by the perceived burden linked to frequent sampling, insufficient incentives for involvement, or a decline in motivation and interest. Incentives, such as providing gifts or access to study results, is a common strategy often employed in survey studies to improve participant retention [19,28]. For example, cohort studies with high retention rates have reported success by using personalised retention strategies, such as providing food, transportation support, financial compensation, and consistent communication through reminders and birthday cards. These studies also emphasise the importance of a well-trained research team that is innovative, persistent and organised [29]. Additionally, a meta-analysis evaluating retention strategies in longitudinal cohort studies found that offering multiple methods of data collection, such as a combination of face-to-face and phone-based approaches, improved participant retention by 10% [30].

Participant attrition may also lead to selection bias, especially given the migration dynamics in Europe in recent years [31]. For this reason, scheduling refreshment samples over time would be essential to maintaining the representativeness of the cohort. In line with the German SOEP's approach [19], refreshment sampling could be undertaken to increase the entire sample size or a specific subgroup of interest according to the epidemiological context. The sample can then be weighted as appropriate to ensure the representativeness of the sample. In addition, to reduce the burden on participants and increase the representativeness of the sample, a portion of the cohort could be included in the sample for a selected round of the survey and rotated ('rotational sampling') [32].

Methodology for data collection and specimen collection and testing

Digital technologies have transformed cohort study methodologies, offering modern alternatives to traditional paper-based methods for data collection, communication, and participant engagement [33]. Tools such as smartphones, web-based applications, and wearable sensors facilitate remote monitoring of physiological parameters, including heart rate, respiratory rate and body temperature, while also enabling real-time symptom reporting and improving long-term participant retention [33,34]. These platforms support direct communication between participants and research staff through SMS, instant messaging, or videoconferencing, enhancing personalised engagement

[33,34]. Additionally, digitalisation allows for in-depth, voice-recorded qualitative interviews, which can reveal themes beyond the scope of structured surveys [35]. Large language models such as GPT-4 (OpenAI) can further contribute by assisting in qualitative data analysis, translating languages for non-native speakers, and serving as virtual assistants to support research staff and participant communication [35].

A notable example from the German context is the 'Prospective Monitoring and Management App (PIA)', used in the ZIPCO study (Integrated DZIF Infection Cohort within the German National Cohort), which featured desirable features in digital cohort tools, such as adaptable questionnaires for various pathogens, integration with external data systems, user-friendly design, and good participant acceptance [33].

Adopting minimally invasive, decentralised and scalable sampling methods will be critical for acceptance and maximising cost-effectiveness. To adapt to different pathogens and required sampling methods, dedicated study sites must be available in all study regions that can respond and scale quickly. This would ensure that surveys can adapt to new testing methods whenever they become available for a new pathogen. In the later phases of epidemics, self-testing, i.e. with rapid antigen tests, approaches are often possible. However, these must first be validated and are often not yet available at the beginning of an epidemic.

A complementary test to pathogen detection itself is pathogen sequencing. During the COVID-19 pandemic, genomic surveillance was critical for identifying viral variants, surveying their global spread, and redefining our knowledge of disease transmission rates, health outcomes and vaccine efficacy, which ultimately determined the public health response to the pandemic [36]. In this regard, established sequencing infrastructures could provide their expertise and resources [37]. Furthermore, antibody surveillance is key to assessing population immunity and should go hand in hand with genomic surveillance.

Infrastructure and costs

Maintaining a large longitudinal cohort is a complex logistical undertaking. Study management, sample collection and analysis, quality assurance, sample storage and transport, and sample refreshment account for most of the costs. Based on a recent cost and cost-effectiveness analysis of active surveillance testing for SARS-CoV-2 in Germany, we estimate a cost per sample of EUR 100–150, which would be ca 5–7.6 million, 2.5–3.75 million, and EUR 1–1.5 million per survey round per German federal state, if each sample were tested under a 70/70%, 70/80% or 70/90% sensitivity/specificity scenario, respectively (Table). These estimates consider the cost per participant for study management, a hotline for study participants, development of study materials, study design, IT services and postal services for shipping RT-LAMP kits, and sample

analysis by RT-LAMP and confirmation by RT-qPCR [38]. Another reference is provided by the REACT study [4], which reported the cost of recruiting participants at GBP 26 (EUR 30.6) without including the cost of PCR and genomic testing (personal communication, Aidan Irwin-Singer, 18 Oct 2022). Integrating existing genomic testing platforms into the cohort [37] and methods to reduce the number of tests through pooling could help to reduce costs [39].

Building a population-representative cohort and link to existing research infrastructures

Building a new long-term cohort is costly and requires a strategic approach to appropriately manage the financial and organisational costs of establishing and maintaining it. An approach that could be deemed practical for mitigating the financial and organisational burdens associated with initiating and sustaining a novel long-term cohort involves leveraging existing cohorts and extending their reach (both in geographical coverage and participant age). This adaptation would involve tailoring the study protocol and incorporating innovative sampling techniques, such as self-testing. In Germany, the SOEP [19] and the NAKO Health Study [25] are two prominent large-scale longitudinal studies with robust research frameworks. These cohorts, however, were not designed for epidemic monitoring purposes, nor were they meant to achieve national representativeness; as such, they have limited age and geographic coverage (NAKO) or lack sufficient statistical power to detect diseases at a low prevalence (SOEP). Nevertheless, the SOEP has served as a major platform for supplementary surveys amid the COVID-19 pandemic. The NAKO cohort has also been used for SARS-CoV-2 surveys and contributed samples and data to other studies [15,28,40]. Although these studies held narrower scopes, they underscored the potential of leveraging large population panels, albeit requiring intricate methodological adjustments to the sampling scheme, sample size and target population. Consideration can also be given to using the research infrastructure of these cohorts for a novel adaptive population-representative cohort, which could offer cost savings while simultaneously facilitating the creation of a more suitable and population-representative cohort.

Furthermore, while such an approach may be resource-intensive, its multi-purpose functionality and adaptability aim to maximise utility and cost-effectiveness by using only a subset of the total cohort and part of the network for repeated surveys outside of epi- or pandemic situations, such as during and after the respiratory season.

To strengthen responses to future health threats, multi-country coordination of preparedness and response capacities under a unified framework will be vital across Europe and beyond [41]. Similarly, initiatives like the EU-funded ORCHESTRA (Connecting European Cohorts to Increase Common and Effective Response To SARS-CoV-2 Pandemic) project, which connected

large-scale population-based cohorts across Europe and internationally, can foster greater collaboration, facilitate data sharing, advance knowledge transfer, and guide public health decision-making [42].

Conclusions

Adopting a population-based longitudinal strategy for active infectious disease research, supported by novel and established methods and infrastructures for data integration, decentralised testing, and communication, holds the potential to substantially transform our capacity for swift, evidence-driven and efficient responses to epidemics, pandemics and interepidemic threats.

Data availability

Not applicable.

Authors' contributions

IM, VKN, MA, AS collected information and literature, drafted the initial manuscript, and reviewed drafts of the manuscript. VKN prepared the table for the manuscript. IM led the study and prepared and reviewed the final version of the manuscript. TB, HB, SC, BK, BL, JR, SS, HW, and AK reviewed and contributed to drafts of the manuscript. CD initiated the study, contributed to drafts of the manuscript, and reviewed the final document.

Conflict of interest

None declared.

Funding statement

BL acknowledges The Helmholtz Association (SO-096HZEPIAdHoc and LOKI: KA1-Co-o8), European Union's Horizon 2020 research and innovation program (grant number 101003480; grant number: 10107382; grant number: 101095606), the Federal Ministry of Education and Research (BMBF) via the IMMUNE BRIDGE project (grant number: 01KX2121), the COVIM project (grant number: 01KX202), the Federal Ministry of Education and Research (BMBF) via the RESPINOW (grant number: MV2021-012) and OptimAgent (grant number: MV2021-014) projects. HW would like to acknowledge funding from the Imperial National Institute for Health and Care Research (NIHR) Biomedical Research Centre.

The PREparedness and PAndemic REsponse in Germany (PREPARED) project, on which this publication is based, is part of the 'Network University Medicine (NUM)' and was funded by the Federal Ministry of Education and Research (BMBF) ('NUM 2.0' FKZ: 01KX2121).

The funders had no role in study design, the decision to publish, or preparation of the manuscript.

Ethical statement

No ethical approval was obtained for this work since no patient data were collected.

Use of artificial intelligence tools

None declared.

Acknowledgements

We thank Aidan Irwin-Singer, Steven Riley, and Francoise Balfe of the UK Department of Health for providing the cost estimate for the REACT study.

References

1. Szklo M. Population-based cohort studies. *Epidemiol Rev.* 1998;20(1):81-90. <https://doi.org/10.1093/oxfordjournals.epirev.a017974> PMID: 9762511
2. Beesley LJ, Osthus D, Del Valle SY. Addressing delayed case reporting in infectious disease forecast modeling. *PLoS Comput Biol.* 2022;18(6):e1010115. <https://doi.org/10.1371/journal.pcbi.1010115> PMID: 35658007
3. Abat C, Chaudet H, Rolain J-M, Colson P, Raoult D. Traditional and syndromic surveillance of infectious diseases and pathogens. *Int J Infect Dis.* 2016;48:22-8. <https://doi.org/10.1016/j.ijid.2016.04.021> PMID: 27143522
4. Elliott P, Whitaker M, Tang D, Eales O, Steyn N, Bodinier B, et al. Design and implementation of a National SARS-CoV-2 Monitoring Program in England: REACT-1 Study. *Am J Public Health.* 2023;113(5):545-54. <https://doi.org/10.2105/AJPH.2023.307230> PMID: 36893367
5. Razuri H, Romero C, Tinoco Y, Guezala MC, Ortiz E, Silva M, et al. Population-based active surveillance cohort studies for influenza: lessons from Peru. *Bull World Health Organ.* 2012;90(4):318-20. <https://doi.org/10.2471/BLT.11.097808> PMID: 22511830
6. Klein C, Borsche M, Balck A, Föh B, Rahmüller J, Peters E, et al. One-year surveillance of SARS-CoV-2 transmission of the ELISA cohort: A model for population-based monitoring of infection risk. *Sci Adv.* 2022;8(15):eabm5016. <https://doi.org/10.1126/sciadv.abm5016> PMID: 35427158
7. Moser A, Carlander M, Wieser S, Hämmig O, Puhan MA, Höglinger M. The COVID-19 Social Monitor longitudinal online panel: Real-time monitoring of social and public health consequences of the COVID-19 emergency in Switzerland. *PLoS One.* 2020;15(11):e0242129. <https://doi.org/10.1371/journal.pone.0242129> PMID: 33175906
8. Penninx BWJH, Benros ME, Klein RS, Vinkers CH. How COVID-19 shaped mental health: from infection to pandemic effects. *Nat Med.* 2022;28(10):2027-37. <https://doi.org/10.1038/s41591-022-02028-2> PMID: 36192553
9. Kuper H, Lopes Moreira ME, Barreto de Araújo TV, Valongueiro S, Fernandes S, Pinto M, et al. The association of depression, anxiety, and stress with caring for a child with Congenital Zika Syndrome in Brazil: Results of a cross-sectional study. *PLoS Negl Trop Dis.* 2019;13(9):e0007768. <https://doi.org/10.1371/journal.pntd.0007768> PMID: 31568478
10. Ward H, Atchison C, Whitaker M, Davies B, Ashby D, Darzi A, et al. Design and implementation of a national program to monitor the prevalence of SARS-CoV-2 IgG antibodies in England using self-testing: The REACT-2 Study. *Am J Public Health.* 2023;113(11):1201-9. <https://doi.org/10.2105/AJPH.2023.307381> PMID: 37733993
11. Koopmans M, de Lamballerie X, Jaenisch T, Rosenberger KD, Morales I, Marques ETA, et al. Familiar barriers still unresolved—a perspective on the Zika virus outbreak research response. *Lancet Infect Dis.* 2019;19(2):e59-62. [https://doi.org/10.1016/S1473-3099\(18\)30497-3](https://doi.org/10.1016/S1473-3099(18)30497-3) PMID: 30420230
12. Tilch K, Hopff SM, Appel K, Kraus M, Lorenz-Depiereux B, Pilgram L, et al. Ethical and coordinative challenges in setting up a national cohort study during the COVID-19 pandemic in Germany. *BMC Med Ethics.* 2023;24(1):84. <https://doi.org/10.1186/s12910-023-00959-0> PMID: 37848886
13. Tacconelli E, Gorska A, Carrara E, Davis RJ, Bonten M, Friedrich AW, et al. Challenges of data sharing in European Covid-19 projects: A learning opportunity for advancing pandemic preparedness and response. *Lancet Reg Health Eur.* 2022;21:100467. <https://doi.org/10.1016/j.lanepe.2022.100467> PMID: 35942201
14. Rico-Urbe LA, Morillo-Cuadrado D, Rodríguez-Laso Á, Vorstenbosch E, Weser AJ, Fincias L, et al. Worldwide mapping of initiatives that integrate population cohorts. *Front Public Health.* 2022;10:964086. <https://doi.org/10.3389/fpubh.2022.964086> PMID: 36262229

15. Lange B, Jaeger VK, Harries M, Rücker V, Streeck H, Blaschke S, et al. Estimates of protection levels against SARS-CoV-2 infection and severe COVID-19 in Germany before the 2022/2023 winter season: the IMMUNE BRIDGE project. *Infection*. 2024;52(1):139-53. <https://doi.org/10.1007/s15010-023-02071-2> PMID: 37530919
16. Tinoco YO, Azziz-Baumgartner E, Uyeki TM, Rázuri HR, Kasper MR, Romero C, et al. Burden of influenza in 4 ecologically distinct regions of Peru: household active surveillance of a community cohort, 2009-2015. *Clin Infect Dis*. 2017;65(9):1532-41. <https://doi.org/10.1093/cid/cix565> PMID: 29020267
17. Wagenaar BH, Augusto O, Ásbjörnsdóttir K, Akullian A, Manaca N, Chale F, et al. Developing a representative community health survey sampling frame using open-source remote satellite imagery in Mozambique. *Int J Health Geogr*. 2018;17(1):37. <https://doi.org/10.1186/s12942-018-0158-4> PMID: 30373621
18. Bosnjak M, Dannwolf T, Enderle T, Schaurer I, Struminskaya B, Tanner A, et al. Establishing an Open Probability-Based Mixed-Mode Panel of the General Population in Germany: The GESIS Panel. *Soc Sci Comput Rev*. 2018;36(1):103-15. <https://doi.org/10.1177/0894439317697949>
19. Goebel J, Grabka MM, Liebig S, Kroh M, Richter D, Schröder C, et al. The German Socio-Economic Panel (SOEP). *Jahrb Natl Okon Stat*. 2019;239(2):345-60. <https://doi.org/10.1515/jbnst-2018-0022>
20. Centerdata. LISS-Longitudinal Internet studies for the Social Sciences panel n.d. Tilburg: Centerdata. [Accessed: 8 May 2025]. Available from: <https://www.lissdata.nl>
21. European Centre for Disease Prevention and Control (ECDC). Assessing the burden of key infectious diseases affecting migrant populations in the EU/EEA. Stockholm: ECDC; 2014. Available from: <https://www.ecdc.europa.eu/en/publications-data/assessing-burden-key-infectious-diseases-affecting-migrant-populations-eueea>
22. Universitätsmedizin Greifswald. NU(M)KRAINE: Infektionsmedizinisches Screeningprogramm des Netzwerks Universitätsmedizin für Flüchtlinge der Ukraine. [NU(M)KRAINE - Infectious Disease Screening Program of the University Medicine Network for Refugees from Ukraine]. Greifswald: Universitätsmedizin Greifswald. [Accessed: 8 May 2025]. German. Available from: <https://www.mezizin.uni-greifswald.de/num/forschungsprojekte/netzwerk-universitaetsmedizin/numkraine>
23. Rogan WJ, Gladen B. Estimating prevalence from the results of a screening test. *Am J Epidemiol*. 1978;107(1):71-6. <https://doi.org/10.1093/oxfordjournals.aje.a112510> PMID: 623091
24. Eaton JW, Bao L. Accounting for nonsampling error in estimates of HIV epidemic trends from antenatal clinic sentinel surveillance. *AIDS*. 2017;31(Suppl 1) Suppl 1:S61-8. <https://doi.org/10.1097/QAD.0000000000001419> PMID: 28296801
25. Peters A, Peters A, Greiser KH, Göttlicher S, Ahrens W, Albrecht M, et al. Framework and baseline examination of the German National Cohort (NAKO). *Eur J Epidemiol*. 2022;37(10):1107-24. <https://doi.org/10.1007/s10654-022-00890-5> PMID: 36260190
26. Selin K, Zimmermann S. SOEP Group. SOEPcompanion (v38) 2023. Berlin: DIW Berlin/SOEP. [Accessed: 6 May 2025]. Available from: <https://www.econstor.eu/bitstream/10419/273538/1/185378723X.pdf>
27. Tigges BB, Kaar JL, Erbstein N, Silberman P, Winseck K, Lopez-Class M, et al. Effectiveness of community outreach and engagement in recruitment success for a prebirth cohort. *J Clin Transl Sci*. 2017;1(3):184-91. <https://doi.org/10.1017/cts.2017.7> PMID: 29082032
28. Bartig S, Brücker H, Butschalowsky H, Danne C, Gößwald A, Goßner L, et al. Corona Monitoring Nationwide (RKI-SOEP-2): seroepidemiological study on the spread of SARS-CoV-2 across Germany. *Jahrb Natl Okon Stat*. 2022;243(3-4):431-49. <https://doi.org/10.1515/jbnst-2022-0047>
29. Abshire M, Dinglas VD, Cajita MIA, Eakin MN, Needham DM, Himmelfarb CD. Participant retention practices in longitudinal clinical research studies with high retention rates. *BMC Med Res Methodol*. 2017;17(1):30. <https://doi.org/10.1186/s12874-017-0310-z> PMID: 28219336
30. Teague S, Youssef GJ, Macdonald JA, Sciberras E, Shatte A, Fuller-Tyszkiewicz M, et al. Retention strategies in longitudinal cohort studies: a systematic review and meta-analysis. *BMC Med Res Methodol*. 2018;18(1):151. <https://doi.org/10.1186/s12874-018-0586-7> PMID: 30477443
31. European Centre for Disease Prevention and Control (ECDC). Assessing the burden of key infectious diseases affecting migrant populations in the EU/EEA. Stockholm: ECDC; 2014. Available from: <https://www.ecdc.europa.eu/en/publications-data/assessing-burden-key-infectious-diseases-affecting-migrant-populations-eueea>
32. Eurostat. Methodological guidelines and description of EU-SILC target variables. Brussels: European Commission; 2022. [Accessed: 26 May 2025]. Available from: <https://circabc.europa.eu/d/a/workspace/SpacesStore/94141a49-a4a7-48bc-89f7-df858c27d016/Methodological%20guidelines%202022%20operation%20v4.pdf>
33. Ortmann J, Heise J-K, Janzen I, Jenniches F, Kemmling Y, Frömke C, et al. Suitability and user acceptance of the eResearch system "Prospective Monitoring and Management App (PIA)"-The example of an epidemiological study on infectious diseases. *PLoS One*. 2023;18(1):e0279969. <https://doi.org/10.1371/journal.pone.0279969> PMID: 36595548
34. Mirjalali S, Peng S, Fang Z, Wang CH, Wu S. Wearable Sensors for Remote Health Monitoring: Potential Applications for Early Diagnosis of Covid-19. *Adv Mater Technol*. 2022;7(1):2100545. <https://doi.org/10.1002/admt.202100545> PMID: 34901382
35. Mathis WS, Zhao S, Pratt N, Weleff J, De Paoli S. Inductive thematic analysis of healthcare qualitative interviews using open-source large language models: How does it compare to traditional methods? *Comput Methods Programs Biomed*. 2024;255:108356. <https://doi.org/10.1016/j.cmpb.2024.108356> PMID: 39067136
36. Tosta S, Moreno K, Schuab G, Fonseca V, Segovia FMC, Kashima S, et al. Global SARS-CoV-2 genomic surveillance: What we have learned (so far). *Infect Genet Evol*. 2023;108:105405. <https://doi.org/10.1016/j.meegid.2023.105405> PMID: 36681102
37. Scheithauer S, Dilthey A, Bludau A, Ciesek S, Corman V, Donker T, et al. Etablierung der Genomischen Erreger-Surveillance zur Stärkung des Pandemie- und Infektionsschutzes in Deutschland. [Establishment of genomic pathogen surveillance to strengthen pandemic preparedness and infection prevention in Germany]. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz*. 2023;66(4):443-9. German. <https://doi.org/10.1007/s00103-023-03680-w> PMID: 36811648
38. Nguyen HT, Denking CM, Brenner S, Koeppl L, Brugnara L, Burk R, et al. Cost and cost-effectiveness of four different SARS-CoV-2 active surveillance strategies: evidence from a randomised control trial in Germany. *Eur J Health Econ*. 2023;24(9):1545-59. <https://doi.org/10.1007/s10198-022-01561-8> PMID: 36656403
39. Cherif A, Grobe N, Wang X, Kotanko P. Simulation of Pool Testing to Identify Patients With Coronavirus Disease 2019 Under Conditions of Limited Test Availability. *JAMA Netw Open*. 2020;3(6):e2013075. <https://doi.org/10.1001/jamanetworkopen.2020.13075> PMID: 32573706
40. Mikolajczyk R, Diexer S, Fricke J, Ahnert P, Pischon T, Karch A. Description of the COVID 2.0 survey in the NAKO and first results. *Eur J Public Health*. 2023;33(Supplement_2):ckad160.502. <https://doi.org/10.1093/eurpub/ckad160.502>
41. Jit M, Ananthakrishnan A, McKee M, Wouters OJ, Beutels P, Teerawattananon Y. Multi-country collaboration in responding to global infectious disease threats: lessons for Europe from the COVID-19 pandemic. *Lancet Reg Health Eur*. 2021;9:100221. <https://doi.org/10.1016/j.lanpe.2021.100221> PMID: 34642675
42. CORDIS - EU research results. Connecting European cohorts to increase common and effective response to SARS-CoV-2 pandemic: ORCHESTRA. Brussels: European Commission; 2020. Available from: <https://cordis.europa.eu/project/id/101016167>
43. Bundesinstitut für Bevölkerungsforschung. Zahl der Privathaushalte und durchschnittliche Haushaltsgröße in Deutschland (1991-2040). [Number of private households and average household size in Germany (1991-2040)]. Wiesbaden: BiB; 2023. German. Available from: <https://www.bib.bund.de/DE/Fakten/Fakt/L50-Privathaushalte-Haushaltsgroesse-1991-Vorausberechnung.html>

License, supplementary material and copyright

This is an open-access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0) Licence. You may share and adapt the material, but must give appropriate credit to the source, provide a link to the licence and indicate if changes were made.

Any supplementary material referenced in the article can be found in the online version.

This article is copyright of the authors or their affiliated institutions, 2025.