Quality of Life Research (2025) 34:3335–3350
https://doi.org/10.1007/s11136-025-04074-y

Design and implementation of data quality controls in the EQ-DAPHNIE study: insights from the pilot phase and 15-country analysis

Fatima Al Sayah[1] · Hilary Short[1] · Juan M. Ramos-Goñi[2] · Rosalie Viney[3] · Erica I. Lubetkin[4] · Mathieu F. Janssen[5] · Jeffrey A. Johnson[1] · the EQ-DAPHNIE Project Team

Received: 26 May 2025 / Accepted: 12 September 2025 / Published online: 17 November 2025

Abstract

Objective The EQ-DAPHNIE (EuroQol Data for Assessment of Population Health Needs and Instrument Evaluation) project is a large, multi-country survey initiative designed to generate population norms and enable comparative research using self-reported health measures. This paper describes the quality control processes and summarizes data quality metrics from the United Kingdom (UK) pilot and full implementation across 15 countries.

Methods Representative samples were recruited via Dynata, an online survey panel provider, using quota sampling by age, sex, income, community setting, and language (where applicable). The UK pilot (n = 3012) informed survey refinements ahead of full rollout (n = 68,411). Quality metrics included completion rates, bot detection, speeding, missing data, outliers, and quota achievement.

Results Across countries, response rates ranged from 80.1 to 100%, with completion rates varying widely (22.9% in Brazil to 60.8% in Japan; average 42.4%). Bot exclusions averaged 3.0%, peaking in China (11.7%). Speeding was low (0.3% average), and duplicate records were rare. Completion times ranged from 18.3 (France) to 31.4 min (New Zealand). Missing data varied substantially (0.0–48.7%), with Japan and Spain showing the least. Quota fulfillment ranged from 68.7 to 98.6%. Consistency checks showed strong agreement for repeated items—marital status (92.8–98.9%) and age (92.3–98.7%).

Conclusions The quality control measures implemented throughout the EQ-DAPHNIE project effectively addressed common issues such as bot responses, speeding, and missing data, resulting in generally high-quality and representative datasets. However, variability across countries underscores the need to account for quality indicators when using the data for norm-setting or cross-country comparisons.

Keywords Data quality · Online surveys · Panel research · Patient-reported outcome measures · Quota sampling

✉ Fatima Al Sayah
falsayah@ualberta.ca

Hilary Short
heshort@ualberta.ca

Juan M. Ramos-Goñi
jramos@mathsinhealth.com

Rosalie Viney
rosalie.viney@uts.edu.au

Erica I. Lubetkin
lubetkin@med.cuny.edu

Mathieu F. Janssen
janssen@euroqol.org

Jeffrey A. Johnson
jeff.johnson@ualberta.ca

[1] Alberta PROMs and EQ-5D Research and Support Unit (APERSU), School of Public Health, University of Alberta, Edmonton, AB, Canada

[2] Decision Analysis and Support Unit, SGH, Warsaw School of Economics, Warsaw, Poland

[3] University of Technology Sydney, Ultimo, Australia

[4] CUNY School of Medicine, New York, USA

[5] EuroQol Research Foundation, PO BOX 4443, 3006 AK Rotterdam, The Netherlands

Introduction

The EuroQol Data for Assessment of Population Health Needs and Instrument Evaluation (EQ-DAPHNIE) project is a large-scale, multi-country initiative aimed at collecting representative data on self-reported health status from diverse populations through online panel surveys [1]. The goal of this initiative is to generate high-quality data across various countries representing different global regions, using widely recognized self-report health measures alongside socio-demographic, economic, and health-related questions. The resulting data will contribute to the establishment of population norms for health status instruments, support psychometric evaluations, and enable meaningful cross-country comparisons.

Given the scale of the EQ-DAPHNIE project, recruitment in each country was conducted through online panels—a widely adopted method for large-scale survey research. Online panels consist of pre-recruited participants who regularly complete online surveys. This approach facilitates access to large, geographically dispersed populations and supports the execution of complex multi-country studies [2–5]. Despite the clear benefits of online surveys and panel-based research, several challenges persist, particularly related to sample representativeness, data quality, and participant engagement [4, 6, 7].

Despite the widespread adoption of online surveys that transformed data collection across many [8], online panels—especially non-probability (convenience) samples—are often prone to selection bias. Members of these panels tend to differ systematically from the general population [9]. A review of seven online panels found that panel members tend to have higher educational attainment and socio-economic status compared to non-panel members [10]. Additionally, convenience panels generally report lower response rates, often below 10%, compared to probability-based panels [9]. Although no universally accepted criteria exist for determining when a convenience panel is adequate for population-level inference, research suggests that bias in estimates may be influenced by the relationship between response propensity (i.e., likelihood or probability that an individual or subgroup will respond to a survey) and the variables of interest [11]. Meta-analyses have indicated that the association between response rate and bias is often weak [12]. When appropriate statistical weighting is applied to correct for coverage errors and selection bias, convenience panels can potentially serve as a basis for generating population norms [7].

Ensuring data quality in online panel research is another key concern. Factors such as non-response bias, incomplete surveys, inattentive respondents (e.g., "speeders" who rush through surveys), and fraudulent submissions (e.g., “bots” or duplicate responses) can threaten data quality [13–15]. To mitigate these risks, robust data quality control measures are necessary. These include requiring responses to key questions, monitoring dropout patterns, controlling response speed, and implementing procedures to detect bots and duplicate entries.

This paper outlines the data quality control strategies implemented in the EQ-DAPHNIE project and presents findings from the 15 countries where data collection has been completed. It also details insights from the UK pilot that informed improvements in survey design and data quality procedures across the full study. Together, these results provide a comprehensive view of the project’s efforts to ensure reliable, high-quality data in a global online panel context.

Methods

Brief overview of the EQ-DAPHNIE project

The EQ-DAPHNIE project is a multi-country initiative, aiming to gather data on a range of self-reported health measures from diverse adult populations. Detailed description of the project methods—including country selection, survey content, and the translation and local adaptation of survey questions—have been previously reported [1], and a summary of the questions and measures used across all countries is provided in Supplementary File 1. The project utilized online surveys hosted on LimeSurvey, a widely-used platform that offers advanced features such as conditional branching, randomization, and skip patterns to enhance survey design and participant experience. Additionally, a Google reCAPTCHA v3 bot detection system was implemented. This system produces a probabilistic score between 0 and 1, with higher values indicating a greater likelihood that the activity originates from a human. The platform recommends 0.5 as the default threshold, which we followed. To further minimize uncertainty, reCAPTCHA was implemented across four separate interactions, and a response was flagged as a bot if *any* of the four scores fell below 0.5.

Participants were recruited through Dynata, a leading global online panel provider, which offers several built-in quality control measures. These include incentivization through point accumulation for survey completion [3], IP address monitoring to prevent duplicate responses, and participant controls designed to improve data integrity [16, 17]. Dynata’s panel encompasses over 67 million members across 90 countries and recruits participants from more than 2000 sources worldwide, ensuring diverse representation. To improve panelist quality, Dynata employs government ID validation and a "Virtual Resume" system, which matches

profiles with real-world experience. Additionally, Dynata's AI-driven QualityScore analyzes over 175 data points to detect fraud and inattention, leading to an 85% reduction in manual data cleaning and enhancing sample accuracy.

Pilot phase design

A pilot study was conducted in the United Kingdom (UK) between April 06 and 20, 2023, targeting 3000 respondents, to evaluate survey design, data quality, and the effectiveness of specific questions and response options. The pilot study incorporated three design features:

- To manage survey length, participants were alternated between survey versions that included the EQ-HWB (short or long versions) and three EQ-5D-5L response heterogeneity vignettes with varying severity levels. Data on completion times informed the design of the full survey, aiming to keep participant burden below 20 min.
- Two versions of EQ-5D-5L response heterogeneity vignettes were tested —one including the full EQ-5D-5L descriptive system and Visual Analogue Scale (EQ VAS), and another with only EQ VAS. This allowed for the assessment of time requirements and feasibility, aiding in the selection of the most practical option for the full survey.
- The pilot included two survey versions—one with non-mandatory questions and another with mandatory questions offering a "prefer not to answer" option. This comparison aimed to identify potential dropout patterns and issues with sensitive questions.

Six survey versions were created and randomly assigned to participants to ensure unbiased distribution. Additionally, questions on educational level and self-reported health were duplicated at the beginning and end of the survey for response consistency checks.

The pilot also included a feedback section where respondents provided input on survey length, clarity of questions and response options, sensitivity of topics, and the ideal completion time.

Main data collection design

Following the pilot phase, the finalized survey was translated into the target languages for each of the 15 countries and adapted for local context. The target sample size for the study was set at 4500 adults per country, with the aim of achieving demographic representation of each country's general adult population. Quota sampling was employed based on age, sex, income, community setting (rural, urban, suburban), and first language in countries with more than one national language. Quotas were informed by national census data or equivalent sources to align the sample distribution with population demographics.

To ensure cultural and linguistic appropriateness, between one and three independent health measurement experts in each country—who were not involved in survey development or the broader EQ-DAPHNIE project—were invited to review the localized versions of the survey. Their feedback was used to refine translations and ensure contextual relevance.

Real-time monitoring of sample characteristics was used to track progress toward quota targets during data collection. After five weeks in the field, adjustments were made where necessary to meet demographic distributions, aiming for a minimum of 85% quota fulfillment.

A soft launch of the survey was conducted in each country with a preliminary sample of 200 respondents. These data were reviewed to identify potential issues related to survey functionality, data capture, or quality control. Upon satisfactory review, full-scale data collection commenced. Each country's data collection period was scheduled for six weeks, with a possible two-week extension if needed. After 5–6 weeks, quota thresholds were relaxed if strict adherence hindered recruitment.

Data analysis

In the pilot phase, multiple quality control checks were applied to evaluate the reliability and completeness of responses. Missing data were analyzed to determine whether missingness followed a systematic pattern (e.g., concentrated in specific question types or demographics) or occurred randomly. Dropout rates were monitored to identify questions or sections that may have contributed to participant attrition, thereby highlighting potential problems with survey length or sensitivity. Responses to open-text numerical fields such as height and weight were plotted and statistically summarized to detect extreme outliers.

To identify inattentive or automated responses, completion time was recorded for each respondent, and those who completed the survey in under five minutes were flagged as "speeders" and excluded. Duplicate entries were identified and removed from the dataset, and the percentage of responses flagged as bots was recorded. Additionally, we examined the proportion of participants clicking on the survey link, consenting, and completing the survey; average and range of completion times; and consistency in responses to duplicated questions.

During the main data collection across 15 countries, similar but expanded quality control procedures were applied to ensure consistency and cross-country comparability. Real-time monitoring dashboards tracked survey progress,

allowing the team to flag and investigate any anomalies. As in the pilot, bot detection scores from reCAPTCHAv3 were generated per respondent; any response with a bot score of 0.5 or less was automatically excluded. Speeders were identified and those below a predefined threshold of 5 min or less (established using pilot data) were flagged and automatically removed. Duplicate responses were screened using multiple identifiers, and any confirmed duplicates were also excluded after thorough review.

Survey completion times were aggregated per country, and descriptive statistics were calculated to identify deviations in participant burden. Response consistency was assessed by comparing responses to duplicated questions using percentage agreement scores. Missing data were quantified by variable and by country, and those with the highest non-response rates were flagged for further review. Outlier detection in numerical fields was also carried out by generating distribution plots and summary statistics, similar to the pilot phase. Finally, data collection duration for each country was recorded and compared to understand operational differences and inform future study planning.

Results

Pilot phase: key findings and survey design adjustments

The pilot survey was sent to 4538 individuals, of whom 96.6% clicked on the survey link, 75.1% provided consent, and 66.4% (n = 3012) completed the survey. Data from these respondents were analyzed and used to inform the final survey design.

Among the 3012 completed surveys, only 11 records (0.36%) exhibited possible bot activity. These records were flagged and excluded in the analysis. A comparison of repeated questions on education revealed a high level of consistency, with 92.7% agreement. However, there was greater variability in responses to the 1-item self-rated health, with only 78.5% consistency. To improve consistency in the final survey, repetitive questions were replaced with "date of birth" and "marital status," as these variables are less prone to framing effect.

Incorporating three response heterogeneity vignettes for the EQ-5D-5L increased survey length, and some respondents indicated that the questions felt repetitive. After consulting with an expert, it was decided to retain one vignette for a moderate health state, which included the full descriptive system and EQ VAS. Additionally, to enhance inclusivity and accommodate diverse populations in the main phase of data collection, the vignette in the final survey was revised to be gender-neutral. Streamlining the survey made it feasible to include the EQ-HWB (where translations were available). The short or long version was selected per country to balance local measures (e.g., ASCOT, WHO-5) with an acceptable survey length.

The analysis of missing data patterns revealed that, across all variables, between 0.2 and 20.2% percent of respondents selected the "prefer not to answer" option. In contrast, the average percentage of missing data across variables was 0.4%, with a range of 0.0–3.3%. This finding indicated that respondents were more likely to select the "prefer not to answer" option than to leave a non-mandatory question unanswered. As a result, all questions in the final survey were retained as non-mandatory (with no "prefer not to answer" option). The socio-demographic questions, which had the highest percentage of missing data, were originally positioned at the end of the survey. These questions were relocated to the beginning of the survey to improve response rates while retaining some at the end for consistency checking.

Several open-text fields for numerical data (e.g., height, weight, physical activity, and sedentary time) exhibited a notable number of outliers. To address this issue, the final survey substituted open-text fields with numerical response options (e.g., drop-down lists), and minimum and maximum limits were applied where appropriate to reduce outliers.

Survey respondents indicated that the survey length and instructions were generally reasonable. However, a minority expressed discomfort with specific mental health-related questions. In response to this feedback, it was decided to include only the 2-item versions of the PHQ-9 and GAD-7 in the final survey, thereby reducing the mental health-related question load. Additionally, concerns were raised regarding sensitive questions, leading to the inclusion of hover-over information icons that provided context for these questions (e.g., income, and illegal drug use), and reminded participants that they could skip them if they felt uncomfortable answering them. The question on political affiliation was also removed from the survey to respect participant sensitivity.

Summary of quality control indicators and quota achievement in the main data collection

The main data collection phase involved diverse sample sizes across the 15 countries, with panel sizes ranging from 7401 in Japan to 19,692 in Brazil, and a total of 175,392 panelists across all countries representing the sampling frame for data collection in these countries (Table 1). The percentage of respondents clicking on the survey link varied from 80.1% in China to 100% in the UK, Australia, and Japan. Most countries achieved high response rates, with more than 90% of respondents clicking the survey link.

Table 1 Panel size, response rates, survey completion time and quota achievement in the EQ-DAPHNIE project

Quality indicator	Pilot (UK)	UK	US	Canada	Australia	New Zealand
Panel size (N)	4538	15,587	8707	11,874	9039	7479
N (%) of link-click	4382 (96.6%)	15,587 (100.0%)	8526 (97.9%)	11,868 (99.9%)	9039 (100.0%)	7412 (99.1%)
N (%) of respondents (provided consent)	3406 (75.1%)	4667 (29.9%)	5859 (67.3%)	9405 (79.2%)	7788 (86.2%)	5856 (78.3%)
N (%) of completers	3012 (66.4%)	4505 (28.9%)	4523 (51.9%)	4707 (39.6%)	5041 (55.8%)	4514 (60.4%)
N (%) of non-completers	394 (8.7%)	10,920 (70.1%)	3704 (42.5%)	6936 (58.4%)	3661 (40.5%)	2560 (34.2%)
Survey completion time (completers), minutes*						
Outliers (5000 min or more)	0	0	0	1	5	0
Mean, SD	18.4 (10.6)	22.4 (25.5)	20.3 (12.5)	22.6 (61.1)	22.0 (15.9)	31.4 (78.4)
Median	15.9	18.7	17.1	18.3	18.4	22.2
Range	4.6–179.3	5.1–1424.9	5.02–166.7	5.03–4075.0	5.02–444.0	5.5–1634.3
% of quotas achieved (average)*	N/A	79.4%	81.6%	77.6%	78.1%	68.7%
Duration of data collection (days)	15	42	40	36	48	37

Quality indicator	France	Spain	Germany	Netherlands	China
Panel size (N)	11,189	8166	13,551	11,746	17,281
N (%) of link-click	10,883 (97.3%)	7293 (89.3%)	13,396 (98.9%)	11,383 (96.9%)	13,838 (80.1%)
N (%) of respondents (provided consent)	7688 (68.7%)	6086 (74.5%)	10,203 (75.3%)	8686 (73.9%)	9811 (56.8%)
N (%) of completers	4502 (40.2%)	4526 (55.4%)	4541 (33.5%)	4506 (38.4%)	4519 (26.2%)
N (%) of non-completers	6136 (54.8%)	2809 (34.4%)	8531 (63.0%)	6601 (56.2%)	7268 (42.1%)
Survey completion time (completers), minutes*					
Outliers (5000 min or more)	0	0	0	0	5
Mean, SD	18.3 (16.0)	21.3 (25.0)	20.0 (14.0)	18.4 (13.6)	23.8 (97.6)
Median	15.5	17.6	16.9	15.5	17.3
Range	5.0–641.6	5.1–1233.0	5.0–331.1	5.0–204.9	5.1–4973.5
% of quotas achieved (average)*	83.8%	93.4%	88.0%	81.7%	98.6%
Duration of data collection (days)	47	37	46	41	40

Quality indicator	Japan	Mexico	Chile	Argentina	Brazil
Panel size (N)	7401	12,658	9670	11,352	19,692
N (%) of link-click	6984 (94.4%)	11,907 (94.1%)	9670 (100.0%)	11,181 (98.5%)	18,281 (92.8%)
N (%) of respondents (provided consent)	5727 (77.4%)	9714 (76.7%)	8201 (84.8%)	10,189 (89.8%)	15,994 (81.2%)
N (%) of completers	4502 (60.8%)	4508 (35.6%)	4503 (46.6%)	4506 (39.7%)	4513 (22.9%)
N (%) of non-completers	2262 (30.6%)	7135 (56.4%)	4913 (50.8%)	6355 (56.0%)	13,370 (67.9%)
Survey completion time (completers), minutes*					
Outliers (5000 min or more)	0	1	1	0	0
Mean, SD	19.7 (21.4)	26.0 (21.8)	25.8 (19.2)	28.7 (58.0)	24.8 (27.9)
Median	16.6	21.4	21.6	22.5	20.6
Range	5.1–1080.4	5.0–697.0	5.1–461.8	5.0–3025.3	5.0–1292.3
% of quotas achieved (average)*	94.4%	76.9%	75.8%	70.7%	77.5%
Duration of data collection (days)	32	54	58	46	38

*In completers sample

NA not available—this data was not collected in this country

The proportion of respondents providing consent varied by country, with Argentina achieving the highest consent rate (89.8%) and the UK the lowest (29.9%). Completer rates ranged from 22.9% in Brazil to 60.8% in Japan, with an overall average completion rate of 42.4%. Conversely, non-completion rates were highest in the UK (70.1%) and lowest in New Zealand (34.2%), with an overall average of 50.5%.

Across countries, some differences were observed between survey completers and non-completers. In Australia, non-completers tended to be older, with a higher proportion of females and slightly lower self-reported health compared to completers. In New Zealand, non-completers were slightly younger, with similar sex distribution but somewhat lower self-reported health. In the Netherlands, age, sex distribution, and self-reported health were similar between completers and non-completers. In China, non-completers were somewhat older, included a slightly higher proportion of males, and reported marginally lower health.

Survey completion times varied substantially, with the mean completion time ranging from 18.3 min in France to 31.4 min in New Zealand, and an overall average of 23 min (Table 1). Quota achievement for key demographic groups varied across countries, with an average completion rate of 81.8%. The highest average proportional quota completion

rates were observed in China (98.6%), Japan (94.4%), and Spain (93.4%). The US, France, Germany, and the Netherlands met between 80 and 90% of their quotas, while other countries generally achieved between 70 and 80%. New Zealand had the lowest quota achievement at 68.7%. All countries achieved above 90% average absolute quota achievement ranging from 93.1% in Argentina and 99.6% in China. Quota achievement for each demographic variable in each country is presented in Table 4.

Data collection duration varied, ranging from 32 days in Japan to 58 days in Chile, with an overall average of 43 days. Data collection was completed within the planned 6-week period in 8 out of the 15 countries. However, in some countries, such as Mexico, Chile, and Argentina where recruiting older adults was challenging, the data collection period was extended to ensure quotas were met.

The proportion of responses flagged as bots was highest in China (11.7%) and lowest in the UK pilot (0.2%), with an average of 3.0% across all countries (Table 2). Speeding was identified in a small percentage of responses (0.3% on average), with France exhibiting the highest speeding rate (0.9%). Agreement on repeated questions was generally high. For marital status, agreement ranged from 92.8% in France to 98.9% in New Zealand, with an average of 96.3%. For age, agreement varied from 92.3% in Germany to 98.7% in the US, with an average of 96.9%. Duplicate records were uncommon across the dataset. No duplicates were identified in 10 of the 15 countries, while three countries each had one duplicate, one country had three duplicates, and another country reported 10 duplicate records.

The percentage of missing data varied by country. Japan (range across all variables 0.0–3.8%) and Spain (0.0–4.9%) exhibited the lowest levels of missing data, while higher levels were noted in Australia (0.1–48.7%), the Netherlands (0.1–13.8%), China (0.1–10.1%), and Brazil (0.1–10.6%) (Table 3). In Australia, the higher level of missing data was primarily due to missing values for height (48.7%) and weight (48.6%) variables. In the Netherlands, missing data were largely observed in educational level (13.8%) and EQ VAS (7.7%). In China, household size exhibited notable missing data (10.1%), while in Brazil, the EQ VAS had the highest missing data (10.6%). The EQ VAS was one of the three variables with the most missing data across all 15 countries, followed by the EQ VAS in the response heterogeneity vignette (in 14 out of 15 countries) and household size (in 13 out of 15 countries). Across countries, missing data were generally associated with younger age. In Australia and China, females were slightly more likely to have missing data on height/weight, household size, or EQ VAS, whereas in New Zealand fewer females had missing EQ VAS. No consistent differences by education were observed across countries. A small number of extreme outliers in height and weight were observed, with respondents reporting heights exceeding 244 cm and weights over 180 kg. The percentage of outliers ranged from 0.0 to 2.2% for height and from 0.0 to 0.8% for weight.

Discussion

This paper highlights findings on the quality of EQ-DAPHNIE data collected across 15 countries, noting both successes and challenges in the survey process. Overall, the survey performed well in terms of engagement and consent rates, though variations suggest that regional factors such as cultural attitudes and internet access may influence participation. Quality control measures—particularly bot detection and speeding checks—were effective, with low incidences of inauthentic responses, indicating that data integrity was largely maintained. While completion times varied, most fell within an acceptable range, suggesting the survey design balanced participant burden with comprehensive data collection.

Despite these successes, variations in missing data and quota achievement were observed, such as missing height and weight in Australia and household size in China. These discrepancies may reflect survey design limitations, comprehension issues, or cultural norms around sensitive information (Table 4).

Users of the EQ-DAPHNIE dataset should carefully consider the implications of missing data for their specific research questions. For instance, analyses may be affected differently if a variable with high levels of missingness is used as an outcome versus as a covariate in a regression model. To ensure robust analyses, researchers are encouraged to explore appropriate strategies for handling missing data, ranging from simple approaches (e.g., complete-case analysis) to advanced methods (e.g., multiple imputation). The choice of method should be guided by the research question, the role of the variable in the analysis, and the extent of missingness.

Quota achievement also varied, with New Zealand recording the lowest completion rate, underscoring challenges in respondent retention and demographic balance. It is also important to note that the initial panel sizes to which survey invitations were distributed differed substantially across countries, reflecting variation in the panel provider's capacity and reach. These differences likely influenced both the ability to meet quotas and the time required to do so. Future iterations may benefit from adaptive recruitment strategies, such as real-time quota monitoring and targeted outreach, especially where groups are harder to reach.

When comparing our data quality metrics to those reported in the literature on online panel surveys, some

Table 2 Bots, speeders, response consistency and duplicate records in the EQ-DAPHNIE project

Quality indicator	Pilot (UK)	UK	US	Canada	Australia	New Zealand
N (%) quality-control fail (bots and speeders)	15 (0.3%)	162 (1.0%)	297 (3.4%)	215 (1.8%)	337 (3.7%)	337 (4.5%)
N (%) of bots	11 (0.2%)	153 (1.0%)	282 (3.2%)	196 (1.7%)	330 (3.7%)	334 (4.5%)
N (%) of speeders	4 (0.1%)	9 (0.1%)	15 (0.2%)	19 (0.2%)	7 (0.1%)	3 (0.0%)
% agreement of repeated questions*						
Education	92.7%					
Self-reported health	78.5%					
Marital status		98.6%	97.7%	97.9%	98.1%	98.9%
Age		98.2%	98.7%	97.9%	98.2%	98.0%
Sample size of 100% agreement on duplicate Qs	2180	4279	4273	4431	4792	4230
N (%) of duplicate records	0 (0.0%)	0 (0.0%)	3 (0.0%)	10 (0.1%)	0 (0.0%)	1 (0.0%)

Quality indicator	France	Spain	Germany	Netherlands	China
N (%) quality-control fail (bots and speeders)	245 (2.2%)	232 (2.8%)	332 (2.5%)	276 (2.3%)	2051 (11.9%)
N (%) of bots	144 (1.3%)	193 (2.4%)	254 (1.9%)	142 (1.2%)	2028 (11.7%)
N (%) of speeders	101 (0.9%)	39 (0.5%)	78 (0.6%)	93 (0.8%)	23 (0.1%)
% agreement of repeated questions*					
Education					
Self-reported health					
Marital status	92.8%	96.5%	95.1%	94.7%	98.3%
Age	97.7%	96.7%	92.3%	96.7%	92.6%
Sample size of 100% agreement on duplicate Qs	4145	4213	3634	4047	4015
N (%) of duplicate records	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (0.0%)

Quality indicator	Japan	Mexico	Chile	Argentina	Brazil
N (%) quality-control fail (bots and speeders)	220 (3.0%)	264 (2.1%)	284 (2.9%)	320 (2.8%)	398 (2.0%)
N (%) of bots	200 (2.7%)	241 (1.9%)	274 (2.8%)	308 (2.7%)	336 (1.7%)
N (%) of speeders	20 (0.3%)	23 (0.2%)	10 (0.1%)	12 (0.1%)	62 (0.3%)
% agreement of repeated questions*					
Education					
Self-reported health					
Marital status	95.7%	95.8%	96.6%	93.5%	95.0%
Age	98.5%	96.4%	97.2%	97.0%	97.8%
Sample size of 100% agreement on duplicate Qs	4267	4132	4127	4047	4187
N (%) of duplicate records	0 (0.0%)	0 (0.0%)	1 (0.0%)	0 (0.0%)	0 (0.0%)

*In completers sample

interesting patterns emerge. The engagement rates in developed countries were consistent with findings from other large-scale, multinational online surveys, which tend to report higher engagement in countries with greater internet penetration [18, 19]. Conversely, regions like China and Brazil with internet penetration of 77% and 84% [20], respectively, and where cultural attitudes toward online surveys vary, showed lower engagement and completion rates. However, it is important to note that the representativeness of Dynata panels used in this study could not be directly evaluated, which is a limitation that may affect the generalizability of the findings.

The issue of bot responses and speeding has become an increasingly prominent concern in online surveys with advances in artificial intelligence (AI) technology. Our findings were consistent with previous research on the prevalence of inauthentic responses, which have become more sophisticated due to the rise of AI-driven bots. [21]. The inclusion of quality control measures such as bot detection and speed monitoring was in line with practices from other studies, and these controls proved effective in preserving data integrity [14, 22]. As AI continues to evolve, future surveys may require even more advanced bot detection systems to maintain the quality of data, particularly as bots become increasingly adept at mimicking human responses.

Dynata utilizes a nonprobability (convenience) panel which may introduce selection bias. To address this, the EQ-DAPHNIE project employed a quota sampling strategy, targeting respondents with specific demographic characteristics such as age, sex, income, and urban/rural residence. These quotas were informed by census or other national data for each country. Additionally, poststratification adjustments (weights) were applied to reduce noncoverage and

Table 3 Missing data and outliers in the EQ-DAPHNIE project

Quality indicator	Pilot (UK)	UK	US	Canada	Australia	New Zealand
% Range missing data across all variables*	0.0–3.3%	0.1–6.1%	0.1–5.9%	0.0–6.4%	0.1–48.7%	0.2–7.6%
Top 3 variables with most missing data:						
Variable 1	3.3% (Years lived in UK)	6.1% (EQ VAS)	5.9% (EQ VAS)	6.4% (EQ VAS)	48.7% (Height)	7.6% (EQ VAS)
Variable 2	2.2% (Household size)	5.2% (EQ VAS vignette)	4.2% (EQ VAS vignette)	6.1% (Weight)	48.6% (Weight)	6.2% (EQ VAS vignette)
Variable 3	2.1% (Comfort with income)	3.7% (Household size)	2.0% (Household size)	5.1% (EQ VAS vignette)	5.1% (EQ VAS)	4.2% (Years of Education)
N (%) outliers*						
Age (≥100)	0 (0.0%)	0 (0.0%)	1 (0.02%)	0 (0.0%)	1 (0.0%)	2 (0.0%)
Height (≥8ft/244 cm)	7 (0.2%)	*NA*	9 (0.2%)	7 (0.2%)	2 (0.0%)	*NA*
Weight (≥400lbs/180 kg)	23 (0.8%)	*NA*	11 (0.2%)	5 (0.1%)	3 (0.1%)	*NA*
Height (<4ft/122 cm)	344 (11.4%)	*NA*	15 (0.3%)	15 (0.3%)	89 (1.8%)	*NA*
Weight (<80lbs/36 kg)	189 (6.3%)	*NA*	167 (3.7%)	371 (7.9%)	22 (0.4%)	*NA*

Quality indicator	France	Spain	Germany	Netherlands	China
% Range missing data across all variables*	0.0–6.0%	0.0–4.9%	0.1–6.0%	0.1–13.8%	0.1–10.1%
Top 3 variables with most missing data:					
Variable 1	6.0% (EQ VAS)	4.9% (EQ VAS)	6.0% (EQ VAS)	13.8% (Education)	10.1% (Household size)
Variable 2	5.5% (EQ VAS vignette)	4.1% (EQ VAS vignette)	5.9% (EQ VAS vignette)	7.7% (EQ VAS)	5.8% (EQ VAS)
Variable 3	3.9% (Household size)	3.3% (Household size)	5.0% (Household size)	6.6% (EQ VAS vignette)	5.5% (EQ VAS vignette)
N (%) outliers*					
Age (≥100)	0 (0.0%)	0 (0.0%)	2 (0.0%)	1 (0.0%)	0 (0.0%)
Height (≥8ft/244 cm)	4 (0.1%)	18 (0.4%)	5 (0.1%)	1 (0.0%)	0 (0.0%)
Weight (≥400lbs/180 kg)	2 (0.0%)	1 (0.0%)	11 (0.2%)	4 (0.1%)	3 (0.07%)
Height (<4ft/122 cm)	45 (1.0%)	328 (7.2%)	141 (3.1%)	89 (2.0%)	24 (0.5%)
Weight (<80lbs/36 kg)	27 (0.6%)	56 (1.2%)	94 (2.1%)	64 (1.4%)	17 (0.4%)

Quality indicator	Japan	Mexico	Chile	Argentina	Brazil
% Range missing data across all variables*	0.0–3.8%	0.0–8.8%	0.1–8.9%	0.2–6.7%	0.1–10.6%
Top 3 variables with most missing data:					
Variable 1	3.8% (Years of Education)	8.8% (EQ VAS)	8.9% (EQ VAS)	6.7% (Household size)	10.6% (EQ VAS)
Variable 2	3.2% (EQ VAS vignette)	7.3% (EQ VAS vignette)	7.5% (EQ VAS vignette)	5.4% (EQ VAS)	9.0% (EQ VAS vignette)
Variable 3	3.1% (EQ VAS)	5.3% (Household size)	7.2% (Household size)	4.8% (EQ VAS vignette)	8.2% (Household size)
N (%) outliers*					
Age (≥100)	0 (0.0%)	3 (0.1%)	0 (0.0%)	0 (0.0%)	2 (0.0%)
Height (≥8ft/244 cm)	1 (0.0%)	55 (1.2%)	98 (2.2%)	28 (0.6%)	2 (0.0%)
Weight (≥400lbs/180 kg)	1 (0.0%)	3 (0.1%)	3 (0.1%)	0 (0.0%)	5 (0.1%)
Height (<4ft/122 cm)	8 (0.2%)	339 (7.5%)	299 (6.6%)	285 (6.3%)	157 (3.5%)
Weight (<80lbs/36 kg)	37 (0.8%)	97 (2.2%)	53 (1.2%)	39 (0.9%)	61 (1.4%)

*In completers sample

Table 4 Quota achievement across demographic variables by country

Country	Quota variable		Overall sample n	Quota %	Quota n	Sample n	Sample %	Proportional* Quota Achievement %	Proportional Difference %	Average Proportional Quota Achievement %	Absolute†Quota Difference	Absolute Average Quota Achievement
United Kingdom	Age	18–24	4505	12.0	541	340	7.5	62.9	37.1	79.4	4.5	97.7
		25–34		17.1	770	783	17.4	101.6	1.6		0.3	
		35–44		17.9	806	842	18.7	104.4	4.4		0.8	
		45–65		32.8	1478	1529	33.9	103.5	3.5		1.1	
		65+		20.0	901	970	21.5	107.7	7.7		1.5	
	Sex	Male		49.0	2207	2097	46.5	95.0	5.0		2.5	
		Female		51.0	2298	2371	52.6	103.2	3.2		1.6	
	Income	Below £10,000		12.8	577	600	13.3	104.1	4.1		0.5	
		£10,000–19,999		38.9	1752	1270	28.2	72.5	27.5		10.7	
		£20,000–29,999		22.0	991	1057	23.5	106.6	6.6		1.5	
		£30,000–49,999		18.0	811	913	20.3	112.6	12.6		2.3	
		£50,000–69,999		4.4	198	378	8.4	190.7	90.7		4.0	
		£70,000–99,999		2.1	95	119	2.6	125.8	25.8		0.5	
		Above £100,000		1.8	81	153	3.4	188.7	88.7		1.6	
	Community Setting	Urban/suburban		79.0	3559	3449	76.6	96.9	3.1		2.4	
		Rural		21.0	946	1017	22.6	107.5	7.5		1.6	
United States	Age	18–24	4523	12.0	543	344	7.6	63.4	36.6	80.4	4.4	97.0
		25–34		17.0	769	800	17.7	104.0	4.0		0.7	
		35–44		16.0	724	772	17.1	106.7	6.7		1.1	
		45–65		33.0	1493	1512	33.4	101.3	1.3		0.4	
		65+		22.0	995	1069	23.6	107.4	7.4		1.6	
	Sex	Male		49.0	2216	2246	49.7	101.3	1.3		0.7	
		Female		51.0	2307	2225	49.2	96.5	3.5		1.8	
	Income	<$50,000 USD		37.0	1674	1704	37.7	101.8	1.8		0.7	
		$50,000-$74,999 USD		17.0	769	844	18.7	109.8	9.8		1.7	
		$75,000-$99,999 USD		12.0	543	600	13.3	110.5	10.5		1.3	
		$100,000-$149,999 USD		16.0	724	747	16.5	103.2	3.2		0.5	
		$150,000-$199,999 USD		8.0	362	552	12.2	152.6	52.6		4.2	
		>$200,000+USD		10.0	452	55	1.2	12.2	87.8		8.8	
	Community Setting	Urban/suburban		80.0	3618	3651	80.7	100.9	0.9		0.7	
		Rural		20.0	905	852	18.8	94.2	5.8		1.2	
	First Language	English		87.0	3935	4375	96.7	111.2	11.2		9.7	
		Spanish		13.0	588	67	1.5	11.4	88.6		11.5	
Canada	Age	18–24	4707	11.0	518	284	6.0	54.9	45.1	77.6	5.0	95.6
		25–34		17.0	800	674	14.3	84.2	15.8		2.7	
		35–44		17.0	800	781	16.6	97.6	2.4		0.4	
		45–65		33.0	1553	1633	34.7	105.1	5.1		1.7	
		65+		22.0	1036	1312	27.9	126.7	26.7		5.9	

Table 4 (continued)

Country	Quota variable		Overall sample n	Quota %	Quota n	Sample n	Sample %	Proportional* Quota Achievement %	Proportional Difference %	Average Proportional Quota Achievement %	Absolute†Quota Difference	Absolute Average Quota Achievement
	Sex	Male		50.0	2354	1954	41.5	83.0	17.0		8.5	
		Female		50.0	2354	2689	57.1	114.3	14.3		7.1	
	Income	<$10,000 CAD		18.0	847	324	6.9	38.2	61.8		11.1	
		$10,000-$19,999 CAD		16.0	753	531	11.3	70.5	29.5		4.7	
		$20,000-$29,999 CAD		13.0	612	641	13.6	104.8	4.8		0.6	
		$30,000-$49,999 CAD		21.0	988	1079	22.9	109.2	9.2		1.9	
		$50,000-$69,999 CAD		14.0	659	724	15.4	109.9	9.9		1.4	
		$70,000-$89,999 CAD		7.0	329	409	8.7	124.1	24.1		1.7	
		$90,000 and over CAD		11.0	518	818	17.4	158.0	58.0		6.4	
	Community Setting	Urban/suburban		82.0	3860	3864	82.1	100.1	0.1		0.1	
		Rural		18.0	847	804	17.1	94.9	5.1		0.9	
	First Language	English		80.0	3766	4006	85.1	106.4	6.4		5.1	
		French		20.0	941	304	6.5	32.3	67.7		13.5	
New Zealand	Age	18–24	4514	12.3	555	295	6.5	53.1	46.9	68.7	5.8	93.7
		25–34		18.5	835	1053	23.3	126.1	26.1		4.8	
		35–44		16.2	731	1140	25.3	155.9	55.9		9.1	
		45–65		33.0	1490	1143	25.3	76.7	23.3		7.7	
		65+		20.0	903	828	18.3	91.7	8.3		1.7	
	Sex	Male		49.0	2212	1635	36.2	73.9	26.1		12.8	
		Female		51.0	2302	2837	62.8	123.2	23.2		11.8	
	Income	$10,000 or less NZD		13.9	627	154	3.4	24.5	75.5		10.5	
		$10,001-$25,000 NZD		12.4	560	224	5.0	40.0	60.0		7.4	
		$25,001-$40,000 NZD		10.6	478	494	10.9	103.2	3.2		0.3	
		$40,001-$55,000 NZD		13.6	614	532	11.8	86.7	13.3		1.8	
		$55,001-$70,000 NZD		15.0	677	658	14.6	97.2	2.8		0.4	
		$70,001-$85,000 NZD		11.1	501	725	16.1	144.7	44.7		5.0	
		$85,001 and over NZD		23.3	1052	1664	36.9	158.2	58.2		13.6	
	Community Setting	Urban/suburban		87.0	3927	3743	82.9	95.3	4.7		4.1	
		Rural		13.0	587	759	16.8	129.3	29.3		3.8	
Australia	Age	18–24	5041	15.0	756	198	3.9	26.2	73.8	78.1	11.1	95.8
		25–34		18.0	907	944	18.7	104.0	4.0		0.7	
		35–44		17.0	857	1054	20.9	123.0	23.0		3.9	
		45–65		32.0	1613	1608	31.9	99.7	0.3		0.1	
		65+		20.0	1008	1224	24.3	121.4	21.4		4.3	
	Sex	Male		49.0	2470	2329	46.2	94.3	5.7		2.8	
		Female		51.0	2571	2650	52.6	103.1	3.1		1.6	
	Income	No income		8.8	444	78	1.5	17.6	82.4		7.3	

Table 4 (continued)

Country	Quota variable		Overall sample n	Quota %	Quota n	Sample n	Sample %	Proportional* Quota Achievement %	Proportional Difference %	Average Proportional Quota Achievement %	Absolute†Quota Difference	Absolute Average Quota Achievement
		Less than $16,000 AUD		1.9	96	108	2.1	112.8	12.8		0.2	
		$16,000-$25,999 AUD		7.7	388	347	6.9	89.4	10.6		0.8	
		$26,000-$41,999 AUD		9.9	499	575	11.4	115.2	15.2		1.5	
		$42,000-$64,999 AUD		12.8	645	694	13.8	107.6	7.6		1.0	
		$65,000-$89,999 AUD		12.7	640	769	15.3	120.1	20.1		2.6	
		$90,000 or more AUD		46.2	2329	2243	44.5	96.3	3.7		1.7	
	Community Setting	Urban/suburban		70.0	3529	4205	83.4	119.2	19.2		13.4	
		Rural		30.0	1512	800	15.9	52.9	47.1		14.1	
France	Age	18–24	4502	8.1	365	391	8.7	107.2	7.2	83.7	0.6	96.6
		25–34		11.5	518	717	15.9	138.5	38.5		4.4	
		35–44		12.4	558	781	17.3	139.9	39.9		4.9	
		45–65		25.7	1157	1660	36.9	143.5	43.5		11.2	
		65+		20.7	932	938	20.8	100.7	0.7		0.1	
	Sex	Male		48.3	2174	1875	41.6	86.2	13.8		6.7	
		Female		51.7	2328	2556	56.8	109.8	9.8		5.1	
	Income	<€11,999		10.0	450	521	11.6	115.7	15.7		1.6	
		€12,000–20,999		30.0	1351	1085	24.1	80.3	19.7		5.9	
		€21,000–29,999		30.0	1351	1200	26.7	88.8	11.2		3.3	
		€30,000–39,999		10.0	450	570	12.7	126.6	26.6		2.7	
		€40,000–52,999		10.0	450	487	10.8	108.2	8.2		0.8	
		€53,000+		10.0	450	463	10.3	102.8	2.8		0.3	
	Community Setting	Urban/suburban		81.5	3669	3553	78.9	96.8	3.2		2.6	
		Rural		18.5	833	859	19.1	103.1	3.1		0.6	
Brazil	Age	18–24	4513	20.1	907	920	20.4	101.4	1.4	77.5	0.3	96.0
		25–54		52.8	2383	2847	63.1	119.5	19.5		10.3	
		55–64		13.8	623	544	12.1	87.3	12.7		1.7	
		65+		13.2	596	183	4.1	30.7	69.3		9.1	
	Sex	Male		49.1	2216	2153	47.7	97.2	2.8		1.4	
		Female		50.9	2297	2304	51.1	100.3	0.3		0.2	
	Income	<R$ 900		10.0	451	341	7.6	75.6	24.4		2.4	
		R$ 900–R$ 1499		30.0	1354	757	16.8	55.9	44.1		13.2	
		R$ 1500–R$ 2199		20.0	903	971	21.5	107.6	7.6		1.5	
		R$ 2200–R$ 2799		10.0	451	521	11.5	115.4	15.4		1.5	
		R$ 2800–R$ 3999		10.0	451	515	11.4	114.1	14.1		1.4	
		R$ 4000–R$ 6199		10.0	451	579	12.8	128.3	28.3		2.8	
		R$ 6200+		10.0	451	629	13.9	139.4	39.4		3.9	
	Community Setting	Urban/suburban		87.6	3953	4083	90.5	103.3	3.3		2.9	

Table 4 (continued)

Country	Quota variable		Overall sample n	Quota %	Quota n	Sample n	Sample %	Proportional* Quota Achievement %	Proportional Difference %	Average Proportional Quota Achievement %	Absolute†Quota Difference	Absolute Average Quota Achievement
		Rural		12.4	560	253	5.6	45.2	54.8		6.8	
Japan	Age	18–64	4502	58.6	2638	3028	67.3	114.8	14.8	94.4	8.7	98.1
		65+		28.8	1297	1455	32.3	112.2	12.2		3.5	
	Sex	Male		48.6	2188	2198	48.8	100.5	0.5		0.2	
		Female		51.4	2314	2258	50.2	97.6	2.4		1.2	
	Income	<¥3 million JPY		34.3	1544	1490	33.1	96.5	3.5		1.2	
		¥3-5million JPY		23.0	1035	1085	24.1	104.8	4.8		1.1	
		¥5–8 million JPY		21.9	986	1016	22.6	103.0	3.0		0.7	
		>¥8 million JPY		21.1	950	873	19.4	91.9	8.1		1.7	
	Community Setting	Urban/suburban		91.9	4137	4139	91.9	100.0	0.0		0.0	
		Rural		8.1	365	339	7.5	93.0	7.0		0.6	
China	Age	18–34	4519	24.7	1116	1113	24.6	99.7	0.3	98.6	0.1	99.6
		35–49		27.2	1229	1225	27.1	99.7	0.3		0.1	
		50–64		27.6	1247	1281	28.3	102.7	2.7		0.7	
		65+		20.5	926	900	19.9	97.2	2.8		0.6	
	Sex	Male		51.1	2309	2346	51.9	101.6	1.6		0.8	
		Female		48.9	2210	2142	47.4	96.9	3.1		1.5	
	Income	<¥60,000 CNY		20.0	904	909	20.1	100.6	0.6		0.1	
		¥60,000–95,900 CNY		20.0	904	867	19.2	95.9	4.1		0.8	
		¥96,000–143,900 CNY		20.0	904	913	20.2	101.0	1.0		0.2	
		¥144,000–239,000 CNY		20.0	904	902	20.0	99.8	0.2		0.0	
		>¥240,000 CNY		20.0	904	902	20.0	99.8	0.2		0.0	
	Community Setting	Urban/suburban		66.2	2992	2983	66.0	99.7	0.3		0.2	
		Rural		33.8	1527	1513	33.5	99.1	0.9		0.3	
Netherlands	Age	18–39	4506	33.0	1487	1571	34.9	105.7	5.7	81.7	1.9	96.4
		40–65		40.0	1802	1749	38.8	97.0	3.0		1.2	
		65+		27.0	1217	1143	25.4	93.9	6.1		1.6	
	Sex	Male		49.7	2239	1977	43.9	88.3	11.7		5.8	
		Female		50.3	2267	2487	55.2	109.7	9.7		4.9	
	Income	<€17,999		11.5	518	480	10.7	92.6	7.4		0.8	
		€18,000–€29,999		34.6	1559	914	20.3	58.6	41.4		14.3	
		€30,000–€49,999		40.4	1820	1704	37.8	93.6	6.4		2.6	
		€50,000–€69,999		10.6	478	557	12.4	116.6	16.6		1.8	
		>€70,000+		2.9	131	220	4.9	168.4	68.4		2.0	
	Community Setting	Urban/suburban		93.0	4191	4008	88.9	95.6	4.4		4.1	
		Rural		7.0	315	438	9.7	138.9	38.9		2.7	
Spain	Age	18–24	4526	13.7	620	624	13.8	100.6	0.6	93.4	0.1	98.0

Table 4 (continued)

Country	Quota variable		Overall sample n	Quota %	Quota n	Sample n	Sample %	Proportional* Quota Achievement %	Proportional Difference %	Average Proportional Quota Achievement %	Absolute†Quota Difference	Absolute Average Quota Achievement
		25–54		47.4	2145	2273	50.2	106.0	6.0		2.8	
		55–64		16.8	760	796	17.6	104.7	4.7		0.8	
		65+		22.3	1009	833	18.4	82.5	17.5		3.9	
	Sex	Male		51.7	2340	2274	50.2	97.2	2.8		1.5	
		Female		48.3	2186	2206	48.7	100.9	0.9		0.4	
	Income	<€1400		30.0	1358	1218	26.9	89.7	10.3		3.1	
		€1400–3000		50.0	2263	2351	51.9	103.9	3.9		1.9	
		>€3000		20.0	905	920	20.3	101.6	1.6		0.3	
	Community Setting	Urban/suburban		81.1	3671	3811	84.2	103.8	3.8		3.1	
		Rural		18.9	855	677	15.0	79.1	20.9		3.9	
Mexico	Age	18–24	4508	23.5	1059	1170	26.0	110.4	10.4	76.9	2.5	93.9
		25–54		47.6	2146	2483	55.1	115.7	15.7		7.5	
		55–64		14.8	667	657	14.6	98.5	1.5		0.2	
		65+		14.2	640	197	4.4	30.8	69.2		9.8	
	Sex	Male		48.5	2186	2008	44.5	91.8	8.2		4.0	
		Female		51.5	2322	2485	55.1	107.0	7.0		3.6	
	Income	<$10,000 MXN		30.0	1352	1473	32.7	108.9	8.9		2.7	
		$10,000–19,999 MXN		60.0	2705	2097	46.5	77.5	22.5		13.5	
		$20,000+MXN		10.0	451	647	14.4	143.5	43.5		4.4	
	Community Setting	Urban/suburban		82.0	3697	4080	90.5	110.4	10.4		8.5	
		Rural		18.0	811	352	7.8	43.4	56.6		10.2	
Chile	Age	18–24	4503	18.8	847	1019	22.6	120.4	20.4	75.8	3.8	94.3
		25–54		47.5	2139	2983	66.2	139.5	39.5		18.7	
		55–64		16.9	761	432	9.6	56.8	43.2		7.3	
		65+		16.8	757	69	1.5	9.1	90.9		15.3	
	Sex	Male		49.2	2215	2197	48.8	99.2	0.8		0.4	
		Female		50.8	2288	2287	50.8	100.0	0.0		0.0	
	Income	<$400,000 CLP		30.0	1351	793	17.6	58.7	41.3		12.4	
		$400,000–599,999 CLP		20.0	901	900	20.0	99.9	0.1		0.0	
		$600,000–999,999 CLP		30.0	1351	1370	30.4	101.4	1.4		0.4	
		>$1,000,000 CLP		20.0	901	1293	28.7	143.6	43.6		8.7	
	Community Setting	Urban/suburban		85.0	3828	3828	85.0	100.0	0.0		0.0	
		Rural		15.0	675	616	13.7	91.2	8.8		1.3	
Argentina	Age	18–24	4506	21.2	955	1110	24.6	116.2	16.2	70.7	3.4	93.1
		25–54		45.6	2055	2702	60.0	131.5	31.5		14.4	
		55–64		15.1	680	530	11.8	77.9	22.1		3.3	
		65+		18.1	816	164	3.6	20.1	79.9		14.5	

Table 4 (continued)

Country	Quota variable		Overall sample n	Quota %	Quota n	Sample n	Sample %	Proportional* Quota Achievement %	Proportional Difference %	Average Proportional Quota Achievement %	Absolute†Quota Difference	Absolute Average Quota Achievement
	Sex	Male		49.6	2235	2222	49.3	99.4	0.6		0.3	
		Female		50.4	2271	2259	50.1	99.5	0.5		0.3	
	Income	<$100,000 ARS		30.0	1352	829	18.4	61.3	38.7		11.6	
		$100,000–$199,999 ARS		30.0	1352	830	18.4	61.4	38.6		11.6	
		$200,000–$299,999 ARS		20.0	901	1078	23.9	119	19.6		3.9	
		>$300,000 ARS		20.0	901	1714	38.0	190.2	90.2		18.0	
	Community Setting	Urban/suburban		95.0	4281	4260	94.5	99.5	0.5		0.5	
		Rural		5.0	225	196	4.3	87.0	13.0		0.7	
Germany	Age	18–24	4541	11.4	518	456	10.0	88.1	11.9	88.0	1.4	96.5
		25–39		23.4	1063	1208	26.6	113.7	13.7		3.2	
		40–64		38.7	1757	1915	42.2	109.0	9.0		3.5	
		65+		26.5	1203	958	21.1	79.6	20.4		5.4	
	Sex	Male		49.2	2234	2403	52.9	107.6	7.6		3.7	
		Female		50.8	2307	2082	45.8	90.3	9.7		5.0	
	Income	<€59,999		40.0	1816	1996	44.0	109.9	9.9		4.0	
		€60,000–€99,999		40.0	1816	1883	41.5	103.7	3.7		1.5	
		€100,000+		20.0	908	592	13.0	65.2	34.8		7.0	
	Community Setting	Urban/suburban		78.0	3542	3427	75.5	96.8	3.2		2.5	
		Rural		22.0	999	1078	23.7	107.9	7.9		1.7	

*Proportional quota achievement: quota n/sample n

†Absolute quota difference: quota %—sample %

nonresponse, aligning the panel respondents' demographic distribution with the target population.

While the use of a probability sample is generally considered more representative, studies have shown that, once weighted appropriately, demographic estimates from convenience panels are often similar to those from probability samples [23]. Similarly, a comparison of responses to the Patient-Reported Outcomes Measurement Information System (PROMIS) global health items across four surveys found comparable estimates of physical and mental health outcomes, despite differences in sampling methods (probability vs. nonprobability) [24].

Nevertheless, it is important to recognize that weighting cannot fully eliminate all sources of bias. Residual biases may persist due to unobservable characteristics that differ between online panel participants and the broader population, such as health literacy, digital access, socioeconomic vulnerability, or underlying propensity to participate in surveys. These limitations underscore the need for caution when using the EQ-DAPHNIE data to make population inferences. In particular, researchers seeking to generate population norms should carefully assess the appropriateness of the sample for their specific questions. Potential strategies to mitigate these limitations include triangulating results with other population-based datasets, conducting sensitivity analyses to test the robustness of findings under different weighting schemes, and, where feasible, supplementing panel data with probability-based samples in key subpopulations.

The findings of this study have important implications for the future use of the EQ-DAPHNIE data. The rigorous data quality controls—such as bot detection, speed monitoring, and duplicate response checks—have helped ensure the accuracy and reliability of the dataset for subsequent analyses. However, the observed variations in missing data and quota achievement suggest that further refinements in survey design and recruitment strategies may be necessary. For instance, countries with higher missing data for specific questions may benefit from revised question formats or additional prompts to increase participant engagement and completion rates. Moreover, more flexible recruitment methods, such as using alternative panels or offline data collection, may be necessary to achieve representative samples in certain regions, particularly those with lower internet penetration or literacy levels [25].

Conclusion

The EQ-DAPHNIE project provides critical insights into the challenges and opportunities associated with conducting large-scale international surveys using online panels. The quality control measures and design adjustments implemented throughout the study effectively addressed many common issues such as bot responses, speeding, and missing data, resulting in a generally high-quality dataset. Lessons learned offer important considerations for future multi-country health surveys, and underscore the need for continued innovation in survey design and quality control measures to ensure reliability and effectiveness of online surveys for large-scale population health research.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11136-025-04074-y.

Acknowledgements We extend our sincere thanks to Professor Paula Lorgelly (University of Auckland, New Zealand) for her valuable guidance and contributions to the design of the response heterogeneity vignettes used in the EQ-DAPHNIE project. We also gratefully acknowledge the efforts of all the researchers across the 15 participating countries who reviewed the survey instruments and ensured their accurate and culturally appropriate local adaptation. We would also like to acknowledge contributions of the EQ-DAPHNIE Project Team for the methodological design and planning of data collection. The EQ-DAPHNIE Project Team: Principal Investigators: Drs. Jeffrey Johnson (University of Alberta) and M.F. Bas Janssen (EuroQol Research Office); Co-Investigators: Al Sayah (University of Alberta); Bailey (The University of the West Indies), Ghandi (Duke-NUS Medical School), Golicki (Medical University of Warsaw), Gutacker (University of York), Lubetkin (CUNY School of Medicine), Mulhern (University of Technology Sydney), Purba (Universitas Padjadjaran), Scott (University of Cape Town), Sullivan (University of Otago), Viney (University of Technology Sydney), Yang (Guizhou Medical University), Zarate (Merk & Co).

Author contributions All listed authors collaboratively contributed to the design of the data quality control measures implemented in this study. FAS and HS conducted the data analysis. FAS led the drafting of the manuscript, with all co-authors contributing to its content. All authors have reviewed and approved the final version of the manuscript.

Funding This project was supported by operating grants from the EuroQol Research Foundation (reference # 367-RA). Views expressed by the authors in the publication do not necessarily reflect the views of the EuroQol Foundation.

Data availability The data that supports the findings of this study are available from the EuroQol Group, but restrictions apply to the availability of these data. Data are currently only available to members of the EuroQol Group, however, data are available for use in collaboration with EuroQol member(s). The authors can advise upon reasonable request. The complete core survey content is available upon request.

Declarations

Ethics approval and consent to participate This study was conducted in line with the principles of the Declaration of Helsiniki. Approval was granted by the University of Alberta (Health Research Ethics Board Pro00123401), University of Otago (Human Ethics Research Committee H23/130), University of Technology Sydney (Human Research Ethics Committee EC00146), and Advarra (Pro00077236).

Consent to participate Informed consent was obtained from all participants included in the study.

Competing interests All authors except HS are members of the EuroQol Group.

References

1. Johnson, J. A., et al. (2025). *EuroQol data for assessment of population health needs and instrument evaluation (EQ-DAPHNIE): A study for enhancing population health assessment*. Quality of Life Research.
2. Blom, A. G., Gathmann, C., & Krieger, U. (2015). Setting up an online panel representative of the general population: The German Internet Panel. *Field Methods, 27*(4), 391–408.
3. Stanley, M., et al. (2020). The effectiveness of incentives on completion rates, data quality, and nonresponse bias in a probability-based internet panel survey. *Field Methods, 32*(2), 159–179.
4. Svensson, J. (2013). Web panel surveys: Can they be designed and used in a scientifically sound way?. In *59th ISI World Statistical Congress*. Hong Kong, China.
5. Svensson, J. (2014). Web panel surveys: A challenge for official statistics. In *Statistics Canada Symposium*. Sweden.
6. Eysenbach, G., & Wyatt, J. (2002). Using the internet for surveys and health research. *Journal of Medical Internet Research*. https://doi.org/10.2196/jmir.4.2.e13
7. Hays, R. D., Liu, H., & Kapteyn, A. (2015). Use of internet panels to conduct surveys. *Behavior Research Methods, 47*(3), 685–690.
8. Lane, T. S., Armin, J., & Gordon, J. S. (2015). Online recruitment methods for web-based and mobile health studies: A review of the literature. *Journal of Medical Internet Research*. https://doi.org/10.2196/jmir.4359
9. Baker, R., et al. (2013). Summary report of the AAPOR task force on non-probability sampling. *Journal of Survey Statistics and Methodology, 1*(2), 90–143.
10. Craig, B. M., et al. (2013). Comparison of US panel vendors for online surveys. *Journal of Medical Internet Research*. https://doi.org/10.2196/jmir.2903
11. Bethlehem, J., et al. (2002). Weighting nonresponse adjustments based on auxiliary information. In R. Groves (Ed.), *Survey nonresponse* (pp. 275–288). Wiley.
12. Groves, R. M., & Peytcheva, M. (2008). The impact of nonresponse rates on nonresponse bias. *Public Opinion Quarterly, 72*, 167–189.
13. Arevalo, M., et al. (2022). Strategies and lessons learned during cleaning of data from research panel participants: Cross-sectional web-based health behavior survey study. *JMIR Formative Research*. https://doi.org/10.2196/35797
14. Douglas, B. D., Ewell, P. J., & Brauer, M. (2023). Data quality in online human-subjects research: Comparisons between MTurk, Prolific, CloudResearch, Qualtrics, and SONA. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0279720
15. Peer, E., et al. (2022). Data quality of platforms and panels for online behavioral research. *Behavior Research Methods, 54*(4), 1643–1662.
16. Dynata. (2024). Dynata's approach to data quality, Dynata, Editor.
17. Fawson, B., & Lorch, J. (2024). The new dynamics of online sample quality, Dynata, Editor. Dynata.
18. Wu, M. J., Zhao, K., & Fils-Aime, F. (2022). Response rates of online surveys in published research: A meta-analysis. *Computers in Human Behavior Reports*. https://doi.org/10.1016/j.chbr.2022.100206
19. Daikeler, J., Silber, H., & Bošnjak, M. (2022). A meta-analysis of how country-level factors affect web survey response rates. *International Journal of Market Research, 64*(3), 306–333.
20. Mei, B., & Brown, G. T. L. (2018). Conducting online surveys in China. *Social Science Computer Review, 36*(6), 721–734.
21. Pozzar, R., et al. (2020). Threats of bots and other bad actors to data quality following research participant recruitment through social media: Cross-sectional questionnaire. *Journal of Medical Internet Research*. https://doi.org/10.2196/23021
22. Griffin, M., et al. (2022). Ensuring survey research data integrity in the era of internet bots. *Quality and Quantity, 56*(4), 2841–2852.
23. Chang, L., & Krosnick, J. A. (2009). National surveys via RDD telephone interviewing versus the internet: Comparing sample representativeness and response quality. *Public Opinion Quarterly, 73*(4), 641–678.
24. Riley, W., et al. (2014). Sources of comparability between probability sample estimates and nonprobability web samples estimates. In *Federal Committee on Statistical Methodology (FCSM) Research Conference*
25. Sammut, R., Griscti, O., & Norman, I. J. (2021). Strategies to improve response rates to web surveys: A literature review. *International journal of nursing studies., 123*, 104058.