# Experiences of NHS healthcare services in England QMI

Quality and Methodology Information (QMI) for the Health Insight Survey (HIS), including the strengths and limitations of the data, methods, and data uses.

Contact:
Health Insight Survey team
Health.Studies@ons.gov.uk
+44 8081 961270

Release date:
26 February 2026

Next release:
To be announced

## Table of contents

1. Output information

2. About this QMI report

3. Important points

4. Quality summary

5. Quality characteristics of the Health Insight Survey

6. Methods used to produce the data

7. Thematic analysis method and quality summary

8. Related links

9. Cite this methodology

# 1 . Output information

- Statistical designation: official statistics in development.
- Survey name: Health Insight Survey (HIS).
- Data collection: the HIS is an online survey of around 100,000 people aged 16 years and over in England.
- Time period: the HIS started in July 2024.
- Frequency: published every four weeks.
- How compiled: we derive estimates from questionnaire responses of HIS participants who were initially enrolled on the Office for National Statistics (ONS) Coronavirus (COVID-19) Infection Survey (CIS).
- Geographic coverage: England, with selected analysis broken down to the NHS regions of England, and integrated care boards.

# 2 . About this QMI report

This quality and methodology information report contains information on the quality characteristics of the data (including the European Statistical System's five dimensions of quality), as well as the methods used to create it.

The information in this report will help you to:

- understand the strengths and limitations of Health Insight Survey (HIS) statistics
- reduce the risk of misusing data
- help you to decide suitable uses for the data
- understand the methods used to create the data

# 3 . Important points

- The Health Insight Survey (HIS) is commissioned by NHS England and delivered by the Office for National Statistics (ONS).
- The HIS was launched in July 2024 in England to understand participants' experience of their GP practice and other NHS services, including dental care and pharmacy services.
- Approximately 244,500 CIS participants were identified as being eligible for inclusion in the HIS and invited to participate.
- Achieved sample size in Wave 1: 104,109 participants.
- Only private residential households and their residents are included in the survey; people in hospitals, care homes and other communal settings are not included.
- We will be publishing weighted data every four weeks to summarise the responses received for each question in the survey.

# 4 . Quality summary

## Overview of the Health Insight Survey

The Health Insight Survey (HIS) was launched in July 2024, in England.

The initial HIS sample is drawn from participants in our Coronavirus (COVID-19) Infection Survey (CIS).

Participants were identified as eligible for the HIS if they:

- took part in the digital version of the CIS
- had agreed to be approached about other ethically approved research studies
- were aged 16 years and over on 1 May 2024
- had an address in England

More information on the CIS sampling method we used to select our participant pool is available in our CIS quality report: December 2022, which was last updated on 30 March 2023.

## Participant recruitment

All eligible CIS participants were sent an invitation to the HIS by email. GOV.UK Notify, a service offered by the Cabinet Office, sent email communications on our behalf.

In many cases, more than one participant from a single household participated in the HIS. This is because the CIS was based initially on a random sample of households to provide a nationally representative survey.

Unlike in the CIS, no financial incentives were offered to participants for taking part. This could be a factor in the lower sign-up rates to the HIS, particularly among younger age groups.

During the survey, we might invite additional participants who were not part of the CIS to obtain a sample that is as representative as possible. Recruitment of new sample members will be carried out by initially contacting selected households by post. Those who wish to take part will be able to sign up online or by telephone. More information will be provided under "How we collect the data" in Section 6: Methods used to produce the data, when this activity occurs.

## Participant engagement

From the 244,500 individuals initially approached, 104,109 took part in the first wave of the HIS (a response rate of 43%).

Completion of the questionnaire was online only (paper copies were not posted to participants). However, if a participant required help with completing the survey online, a telephone number was provided so the dedicated contact centre could assist the participant and, if needed, complete the questionnaire on their behalf over the phone.

## Uses of the HIS

The HIS provides important information about participants' experiences of accessing care at their local GP practice, and their experience of other NHS services in England, including dental care and pharmacy services.

The statistics produced will provide valuable information to help the NHS to:

- improve local health services
- monitor their Delivery plan for recovering access to primary care, which aims to make it easier and quicker for patients to get the help they need from primary care

## Interpreting the HIS data

Users should be aware that data collected during some time periods may be affected by specific wider events, for example, the Microsoft Windows outage on 19 July 2024 and GP work to rule, effective from 1 August 2024. We do not expect this will affect the data quality, but it might cause unexpected trends in the outcomes reported.

Changes over time should be interpreted with caution. The estimates are weighted, but fluctuations can still occur because of factors such as seasonality, changes in sample composition, especially when new participants are added, adjustments to weighting methodology and changes in NHS geography (for example, at the organisational level or because of patients moving between GP practices). Trends may reflect genuine changes in access to NHS healthcare services or be influenced by these factors. Data should be examined over multiple time points to determine if a real trend exists.

## Strengths and limitations

- a large sample of participants, which allows for the production of estimates for smaller geographic areas
- a questionnaire which regularly examines various aspects of NHS primary health care and allows for changes in participant experiences to be monitored over time
- questions found to be unreliable in testing are improved and re-tested before being released into the field
- questions that have been cognitively tested to ensure they capture the intent of the questions, are easy to understand and reduce participant burden
- questions that have been developed in consultation with NHS England topic experts to meet user needs
- a quick turnaround, as datasets are published on our website four weeks after each wave has closed

The HIS sample, as with all surveys, is subject to possible bias. For example, participants need access to an email address to participate. For more information see ”Representativeness” in Section 5: Quality characteristics of the Health Insight Survey.

All estimates in our publications contain uncertainty. Although the statistics produced as outputs from the survey data are our best estimates, they will not be identical to the unknown true numbers we are trying to measure.

## How we quality assure and validate the data

We quality assure our HIS survey data at each stage of the research process, including when we draw the sample, develop our questionnaire, and publish our data and estimates. These processes include:

- updating contact information in the sample daily through the HIS contact centre to reflect withdrawals or changes to personal details
- cognitively testing survey guidance and instructions to ensure the questions collect appropriate data
- iteratively testing questions until no further changes are made
- following a rigorous methodology for producing the weights used in the analysis
- testing the questionnaire, using test scenarios, to sense check questions and question routing
- conducting a series of checks on the data after collection to identify inconsistencies and invalid responses
- ensuring that a researcher reviews inconsistent or invalid responses before a decision is made about how to deal with them
- dual running of processes and checking estimates at multiple stages between analysis and publication (including the publication of final reference tables)
- applying statistical disclosure control and suppressing estimates to avoid any disclosure of personal information
- at appropriate points in time monitoring and analysing feedback on the questionnaire to improve question wording and response options

# 5 . Quality characteristics of the Health Insight Survey

This section describes the quality characteristics of the data and identifies issues that should be considered when using the statistics.

### Relevance

The Health Insight Survey (HIS) seeks to understand how access to NHS services is changing over time, both across England as a whole and across different regions and demographics.

We collect data through self-completed online questionnaires. Data are analysed to understand how users access different NHS services and their perception of how these services are offered. This study analyses patients' experiences of local NHS services, including GPs, dental care, and pharmacy services.

Once a time series is available and an evaluation of the statistics is complete they can be used to highlight potential pressures for the NHS and to help support wider services. For example, assisting governments with informed decisions on important policies, such as service planning, and understanding the impact of systems on different groups in the population.

### Accuracy and reliability

HIS estimates in our publications contain some uncertainty. Potential sources of uncertainty include:

- the quality of the data collected in the questionnaire
- statistical uncertainty in the estimates, as the data are based on a sample of people rather than the whole population
- bias, which may not be fully addressed by weighting methods

To quantify uncertainty in our analyses, we present 95% confidence intervals in our data. Confidence intervals give an indication of the degree of uncertainty in an estimate. A wider interval indicates more uncertainty in the estimate. Overlapping intervals indicate that there may not be a true difference between two estimates.

Routing based on responses means that some questions in the survey are only asked to a specific subset of the sample. As a result, some questions have very small sample sizes and may not provide reliable trends.

We minimise errors during data collection, for example, missing data, by only enabling participants to submit the questionnaire after they have answered all the questions required.

## Representativeness

Ensuring a representative sample of the general population is important for producing survey-based estimates broken down by characteristics such as age, sex, and NHS region. In the HIS, ensuring a representative sample helps us to understand trends in different population sub-groups across England.

Aspects of the study design may affect the representativeness of the study population compared with the wider population. For example, participants needed access to an email address to participate (this did not necessarily need to be their own and the survey could be completed by others on the participant's behalf). This excludes a small proportion of the population from participating. The latest data from Ofcom's Online Nation 2023 report (PDF, 4.5MB) reports that 92% of households in the UK have internet access. We also know that 89% of households responded to Census 2021 online (for more information, see our Population and household estimates, England and Wales: Census 2021 bulletin, published in June 2022).

Unlike the CIS, no incentives were offered to take part in the HIS. This may affect response rates in some groups more than others.

In this section, we compare those that responded to the HIS with the target population. Note that the weighting described in Section 6: Methods used to produce the data aims to reduce the effect of these differences in the survey.

Within the HIS sample:

- females are slightly over-represented, while males are slightly under-represented (England 52% female and 48% male; HIS 57% female and 43% male)
- younger age groups (aged 16 to 24 years, 25 to 34 years and 35 to 44 years) are under-represented when compared with older age groups (aged 55 to 64 years, 65 to 74 years and 75 years and over), which are over-represented
- those reporting white ethnicity are largely over-represented (England 83%; HIS 96%)

Table 1: Actual relevant population and Health Insight Survey response population by sex, England

| | Actual relevant population profile | | HIS unadjusted response profile | | Difference | |
|---|---|---|---|---|---|---|
| **Category** | **Number of people** | **Proportion** | **Number of people** | **Proportion** | **Absolute difference to actual population** | **Relative difference to actual population** |
| **Female** | 23,719,832 | 0.52 | 59,393 | 0.57 | 0.05 | 11% |
| **Male** | 22,287,125 | 0.48 | 44,716 | 0.43 | -0.05 | -11% |

Source: Health Insight Survey from the Office for National Statistics

Notes

1. Population figures used for the "actual relevant population" are drawn from Census 2021 for England.
2. Response figures are unadjusted and represent the actual response numbers in Wave 1.
3. Absolute difference represents the difference in the unweighted proportion in our response population to the actual population.
4. Relative difference represents the proportion of difference attributed to each group. This is calculated by taking the absolute difference and dividing by the proportion of the actual population.

Table 2: Actual relevant population and Health Insight Survey response population by age, England

| | Actual relevant population profile | | HIS unadjusted response profile | | Difference | |
|---|---|---|---|---|---|---|
| **Category** | **Number of people** | **Proportion** | **Number of people** | **Proportion** | **Absolute difference to actual population** | **Relative difference to actual population** |
| **16 to 24 years** | 5,989,233 | 0.13 | 1,764 | 0.02 | -0.11 | -87% |
| **25 to 34 years** | 7,667,865 | 0.17 | 2,443 | 0.02 | -0.14 | -86% |
| **35 to 44 years** | 7,375,794 | 0.16 | 7,438 | 0.07 | -0.09 | -55% |
| **45 to 54 years** | 7,510,397 | 0.16 | 14,426 | 0.14 | -0.02 | -15% |
| **55 to 64 years** | 7,062,367 | 0.15 | 25,913 | 0.25 | 0.10 | 62% |
| **65 to 74 years** | 5,564,143 | 0.12 | 32,569 | 0.31 | 0.19 | 159% |
| **75 years and over** | 4,837,156 | 0.11 | 19,556 | 0.19 | 0.08 | 79% |

Source: Health Insight Survey from the Office for National Statistics

Notes

1. Population figures used for the “actual relevant population” are drawn from Census 2021 for England.
2. Response figures are unadjusted and represent the actual response numbers in Wave 1.
3. Absolute difference represents the difference in the unweighted proportion in our sample to the actual population.
4. Relative difference represents the proportion of difference attributed to each group. This is calculated by taking the absolute difference and dividing by the proportion of the actual population.

Table 3: Actual population and Health Insight Survey response population by ethnicity, England

| | Actual relevant population profile | | HIS unadjusted response profile | | Difference | |
|---|---|---|---|---|---|---|
| **Category** | **Number of people** | **Proportion** | **Number of people** | **Proportion** | **Absolute difference to actual population** | **Relative difference to actual population** |
| **White** | 38,191,425 | 0.83 | 99,779 | 0.958 | 0.13 | 15% |
| **Asian, Asian British** | 4,134,389 | 0.09 | 2,360 | 0.023 | -0.07 | -75% |
| **Black, African, Caribbean, Black British** | 1,792,283 | 0.04 | 598 | 0.006 | -0.03 | -85% |
| **Mixed, Multiple ethnic groups** | 938,207 | 0.02 | 861 | 0.008 | -0.01 | -59% |
| **Other ethnic groups** | 950,648 | 0.02 | 511 | 0.005 | -0.02 | -76% |

Source: Health Insight Survey from the Office for National Statistics

Notes

1. Population figures used for the “actual relevant population” are drawn from Census 2021 for England.
2. Response figures are unadjusted and represent the actual response numbers in Wave 1.
3. Absolute difference represents the difference in the unweighted proportion in our sample to the actual population.
4. Relative difference represents the proportion of difference attributed to each group. This is calculated by taking the absolute difference and dividing by the proportion of the actual population.

## Characteristics of the sample

Participants are asked for their employment status, and whether they work exclusively from home, another location or a combination of both. Further demographic information, such as age and ethnicity, is taken from the participants’ response to the CIS. This demographic data will be collected as part of the registration process for participants who join the sample at a later date.

## Accessibility and clarity

The Office for National Statistics’ (ONS’s) recommended format for accessible content is a combination of HTML web pages for narrative, charts, and graphs, with data provided in usable formats, such as Excel spreadsheets. Our outputs conform to the ONS Website accessibility policy in terms of formats, font sizes, and the presentation of tables and charts.

More details on our related releases can be found in Section 7: Related links. If there are any changes to our pre-announced release schedule, public attention will be drawn to the change and the reasons for it will be explained fully.

HIS data will be available in our Trusted Research Environment (TRE). As this provides access to microdata and more detailed data that have the potential to identify individuals, access to these data requires approved researcher accreditation.

### Timeliness and punctuality

This publication provides timely and punctual information from the HIS, detailing our analysis on patients' experiences of local NHS services, including GPs, dental care, and pharmacy services. These data are collected, processed, and published approximately four weeks after data collection ends.

### Why you can trust our data

The ONS is the UK's largest independent producer of official statistics and its recognised national statistical institute. Our Data Policies and Information Charter details how data are collected, secured and used in the publication of statistics. We treat the data that we hold with respect, keeping the data secure and confidential. We use statistical methods that are professional, ethical, and transparent. More information about our data protection policies is available.

### Comparison with other data sources

The results produced using HIS data may differ from those reported in other surveys, such as the GP Patient Survey (GPPS) and the Opinions and Lifestyle Survey (OPN). This may be explained by differences between the questionnaires, such as sample characteristics that have not been fully adjusted for in the weighting methodology, timeframes the questions refer to, wording of questions, and the order of when questions were asked.

This survey also collects information that aims to provide an understanding of people's experience of being on an NHS waiting list. The data collected should not be used to monitor the number or proportion of the population in England who are currently on a hospital waiting list; this information can be found in NHS England's Waiting List Minimum Data Set (WLMDS).

### Provisional estimates and revisions

When any HIS data are found to be in error, both the data and any associated analysis that has been published by the ONS will be updated in line with our Revisions and corrections policy.

There are several reasons why we may wish to update the survey estimates or datasets once they have been published; for example, if errors are discovered in raw or derived variables.

Every effort is made to check the data thoroughly before they are published. However, errors do occasionally occur. When errors occur, corrections are made in a timely manner, announced, and clearly explained to users in line with the ONS guide to statistical revisions. Work is also undertaken to prevent the same error from occurring again, for example, by reviewing and improving code.

# 6 . Methods used to produce the data

The data collected by the Health Insight Survey (HIS) enable us to estimate how people access NHS GP services, dentists, and pharmacies as well as their perception of these services.

## How we collect the data

### Sampling method

Our overview of the HIS in Section 3: Important points outlines the participation criteria for the survey. Coronavirus (COVID-19) Infection Survey (CIS) participants were considered eligible to be invited to HIS if they:

- took part in the digital version of the CIS
- had agreed to be approached about other ethically approved research studies
- were aged 16 years and over on 1 May 2024
- had an address in England

More information on the initial CIS sample from which we selected our participant pool is available in the CIS quality report: December 2022 which was last updated on 30 March 2023.

## Study design

Participants have been divided into nine equal groups. Each household has been assigned randomly to a group, and each group had a similar distribution of main characteristics (for example, age).

Each week, for three consecutive weeks, participants in three groups are sent an email invitation (on Tuesday, Wednesday, and Thursday) to complete the online questionnaire. Participants have a seven-day window, including the day their email was sent, to complete the questionnaire. In the fourth week, no new email invitations are sent out. This provides a short period between participants completing the questionnaire and new invitations being sent, for updates to our systems and processes to be implemented. This study design means that each participant is emailed approximately every four weeks with an invitation to participate. The study follows a panel design which means that the full sample is emailed every four weeks.

Each four-week period is referred to as a “wave”. The seven-day questionnaire windows overlap, as this helps to smooth out the number of responses received each day. Participants were more likely to return the questionnaire at the beginning of their window.

Participants who did not complete the HIS questionnaire for three waves in a row were removed from the survey sample. From the end of Wave 7 (7 January to 29 January 2025) onwards this rule was no longer applied to try and reduce the further removal of under-represented groups.

## Sample boost

At the end of Wave 4 a small sample boost was carried out, these new participants will be included in the analysis from Wave 6 onwards. The boost was targeted in 22 Primary Care Networks (PCNs) in order to increase the sample size to aid potential policy evaluation in these areas.

The boost sample was drawn fromAddressBase, which is maintained by the Ordnance Survey, and contains details for 40 million residential and business addresses in the UK. As PCNs are not strictly geographically bound, a list of postcodes with a high likelihood of addresses and households within the postcode being registered with a relevant PCN was used to select the sample. 95,800 addresses in these postcodes were selected, with the households sent a letter inviting them to register their details via SmartSurvey. The details provided were then cleaned (for example, those with invalid dates of birth or email addresses were removed), and following this 6,650 participants were added to one of the nine groups of the HIS sample.

## Sample boosts June 2025 onwards

Regular sample boosts are being carried out from June 2025, with the aim of recruiting more participants in groups underrepresented in the survey.

We use the NHS Personal Demographic Service (PDS) to draw the sample. We select NHS numbers from the required categories and these are then passed through the NHS National Data Opt-Out to remove those who have specified that their records are not used for healthcare research or planning. The remaining NHS numbers are then passed to NHS Notify to send survey invites to individual participants by email. If no valid email address is available or the email is undelivered, a text message is sent.

This approach was initially trialled with half of the boost sample in June 2025, because using text messages incurs a cost. Based on analysis of the sign-up rate, the proportion of the sample that NHS Notify were able to contact increased from 64% to 87% when sending text messages after email failure or non-contact. Text messages were also more effective in contacting the younger age groups, where survey participation is lower. Text messages are being sent to all those sampled after a failed email, or where no email address is available, from August 2025.

Those interested in taking part in the survey are invited to register their details through SmartSurvey. Following data cleaning, just over 15,000 individuals were added to the sample for the June 2025 boost. This represents 5% of the sample drawn.

## Data we collect

We collect data from each participant by using an online questionnaire. Participants are asked a series of questions on their ability to access NHS GP services, dentistry services and pharmaceutical care. They are also asked questions about their subjective perception of these services.

At the beginning of Wave 4 new questions were added to the survey to enable the collection of data relating to being offered a face-to-face appointment and an appointment with a preferred healthcare professional.

## Response rates

The response rate for Wave 1 of the HIS was 43%, a further breakdown of the response rates by participant characteristics are included in Table 40 of the data tables accompanying our Experiences of NHS healthcare services in England: September 2024 bulletin. Response rates for subsequent waves will be added to this table in each publication.

The eligible population used to calculate response rates includes anyone who has been invited to participate in at least one wave of the survey and has not moved out of England or died. Participants who have withdrawn from the survey for any other reason are not contacted again to take part after their withdrawal, but they remain part of the eligible population. This definition has been used to calculate response rates from Wave 12 (27 May to 18 June 2025) onwards. We have included response rates from the previous 11 waves that have been recalculated using this definition in our Health Insight Survey, England: Wave 12 dataset (XLSX, 1.3MB).

Until the end of Wave 6 (10 December 2024 to 1 January 2025), participants who had not responded for three consecutive waves were removed from the sample. This means they no longer receive further invitations to take part in the HIS. By the end of Wave 6, around 137,000 non-active participants had been removed from the sample in this way. However, they remain in the denominator of published response rates as they are still considered part of the eligible population, unless they meet the ineligibility criteria outlined in this section.

## How we analyse the data

### Allocating NHS geography

A participant's registered GP surgery is identified using the Patient Demographic Service. A lookup file is then used to attach the primary care network (PCN) to which the GP surgery belongs. Higher level health geographies are then mapped on through the PCN. This is used to ensure participants nest correctly in the hierarchy (for example, the number of participants in each integrated care board (ICB) is the same as the sum of all participants in the PCNs that are located in each ICB).

In some cases, a participant may be allocated to an ICB that is different to the one that covers the area where they are recorded as living. We carried out analysis to determine whether this makes a difference to the number of participants in each ICB. We found that most participants were allocated to the same ICB if they had been mapped using their home postcode.

There are a small number of participants that we are unable to link to a GP surgery (under 1%). Questionnaire responses from these participants are included in the overall figures and the demographic breakdowns but not the geographic breakdowns.

### Weighting

We apply weighting to our HIS survey results to ensure they are representative of the target population. This adjusts for differential inclusion or response rates of demographic groups, which could otherwise result in under-representation (for example, of younger age groups).

The weighting strategy accounts for:

- the probability of inclusion in the CIS
- the probability of selection for the HIS
- the probability of responding to the HIS
- known population totals

The HIS sample is a subset of the CIS sample. Therefore, the CIS design weights, which account for an individual's probability of inclusion in the CIS, are used as a basis for the HIS design weights. An individual's CIS design weight is then adjusted for factors correlated to the likelihood of the individual consenting to further research, derived using logistic regression. Logistic regression is also used to adjust for the probability of sampled participants responding to the survey.

Initial weights are then calibrated to population totals. Calibration groups are age group by sex and age group by health region separately.

### Probability of being selected to take part in the HIS

Design weights are used to account for the probability that a person has been selected to take part in the survey. Differences in the probability that a person is selected to take part in the survey can introduce bias to estimates if they are not accounted for.

The HIS sample is a subset of the CIS population. Therefore, the probability of being selected for the HIS is linked to the initial probability of being selected for the CIS. For this reason, the CIS design weights are used as a basis for the HIS design weights.

Participants need to fulfil certain requirements to be selected from the CIS sample for inclusion in the HIS sample (for example, having access to the internet to complete the questionnaire). This means that their probability of selection is one. However, there are factors that influence the likelihood that participants will fit those selection criteria; for example, some participants are more likely to respond online or consent to further research. To account for this, logistic regression is used to identify the probability of a participant being selected for the HIS, given they took part in the CIS. The regression model includes:

- age
- sex (male and female)
- region
- ethnicity (as two categories: “White” and a combined category comprising all other ethnic groups)
- Index of Multiple Deprivation quintiles (IMD)

For each participant, the CIS design weight is multiplied by the reciprocal of the probability of selection for the HIS, given they took part in the CIS, to produce the HIS design weights.

### Probability of responding to the HIS

There is non-response within the sample that were invited to the HIS. We account for this using a logistic regression model. We use information from previous survey responses that is not available as population totals, for example, previous ethnicity. Therefore, a logistic regression model is used to calculate the probability of response. The model includes:

- age
- sex (male and female)
- region
- ethnicity (as two categories: “White” and a combined category comprising all other ethnic groups)

HIS design weights are multiplied by the reciprocal of the probability for responding to create an initial weight. These weights are the basis used for calibration.

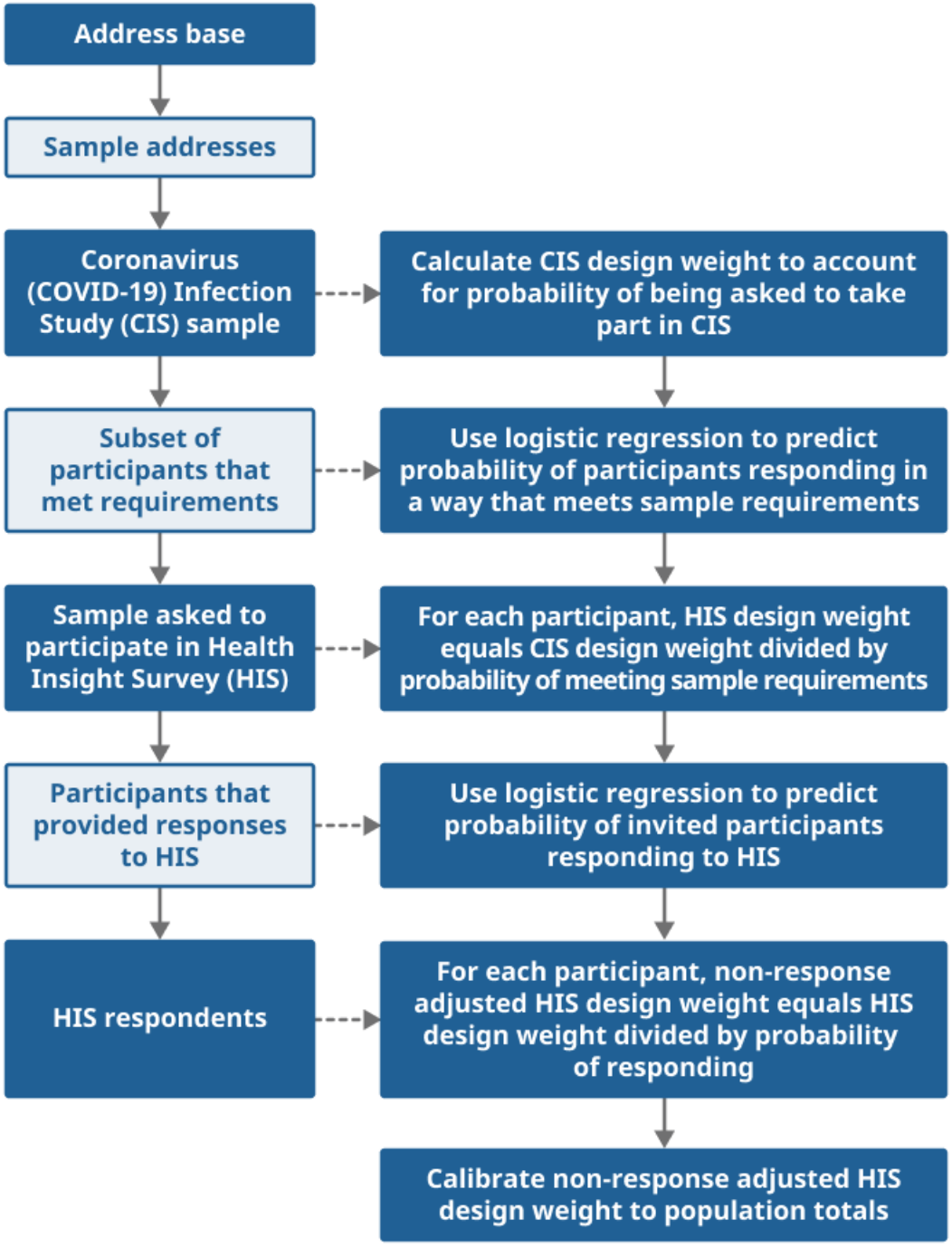
## Calibration and population totals

Calibration is carried out to ensure that the weighted samples for the different population groups sum to the known population totals. Calibration can account for more than one set of population totals, summing to the same number, at the same time (for example, population totals based on age by sex, and population totals based on health region for the same population). This avoids weights being based on small sample sizes, for example, when looking at age by sex by health region.

Here, the calibration groups are age group by sex and health region separately.

To obtain population totals for the period closest to the study dates, ONS population projections were used rather than the GP registered population. The population protection totals will differ to the GP registered population, particularly in areas of the country with higher levels of population movement. The ONS population projections were selected as they have been adjusted to remove those in communal establishments and are more suitable for the questions unrelated to GP access (for example, pharmacy and dentistry questions). The calibration population totals are provided in Table 11 of the data tables accompanying our Experiences of NHS healthcare services in England: September 2024 bulletin.

**Figure 1: Initial Health Insight Survey sample selection and weighting flow diagram**




**Source: Health Insight Survey from the Office for National Statistics**

### Weighting sample boost participants

To calculate design weights for sample boost participants the probability of taking part in HIS for each PCN is the number of addresses selected divided by the number of addresses available.

At this stage the design weights for the boost sample add up to the number of addresses in the PCNs of interest and the design weights for the non-boost sample add up to the number of addresses in England including those PCNs. To allow for the two to be combined a scale factor was applied to the design weights for each PCN of interest and sample type (boost, non-boost). This ensured the design weights across both samples summed to the addresses in England as expected.

After calculating the design weights for sample boost participants, the remaining weighting process follows the same procedure as for the initial sample.

# 7 . Thematic analysis method and quality summary

## Data collection and sample

From Wave 13 (24 June to 16 July 2025), free-text questions concerning reasons for a positive or negative experience of NHS services, specifically participants' experience with their GP surgery or community health service appointments, were added to the Health Insight Survey (HIS).

In each HIS wave since, participants who have tried to contact their GP surgery or had an appointment with a community health service in the last 28 days, are asked to rate their experience. The added free-text questions give participants the option to provide a reason for their rating.

All free-text questions in the HIS have a 250-character limit. As these questions are voluntary to answer, the sample responding to them is only a subset of those who participate in HIS – usually between fewer than a hundred and a few thousand per wave, dependent on the question.

The question wording can be found in the Health Insight Survey questionnaire.

## Analysis method

Thematic analysis is a family of qualitative data analysis methods that result in themes, which are categories of responses.

Codebook thematic analysis is a type of thematic analysis where a list of codes (a codebook) is developed initially, and responses are labelled by the codes (topics) that they touch on. These codes, also referred to as subthemes, are grouped into themes, which categorise them into broader topics.

As opposed to another popular type of thematic analysis – reflexive thematic analysis, the codebook approach assumes that findings and conclusions of qualitative analysis can be treated as objective information rather than insight rooted in the researcher's and the sample's subjectivity. This better aligns with our research goals (informing policy based on common patient experiences).

Our approach aims to be inductive, or bottom-up, by developing the codebook based on a subset of responses instead of based on research questions, as well as by iteratively revising the codebook throughout analysis. In addition, we aim to decrease bias by parts of the process being conducted by two researchers, enabling them to moderate each other's subjectivity.

## Our thematic analysis process

1. Familiarisation and initial codebook development: two analysts read and code separately the same 10% (or a minimum of 50, maximum of 200) randomly selected free-text responses, for both the "good" and "poor" experiences (for each service). These initial codes are then compared, and one codebook is agreed on by both analysts.

2. Applying the codebook: the codebook is used for main coding by one analyst, until data saturation is reached. This means that coding is conducted until no new codes are being identified. This is an iterative process, with codes being revised, as well as new codes being added if needed.

3. Finalising the codes and themes: both analysts reach agreement on the final set of codes (which become subthemes) and themes. They categorise the subthemes into categories of interpreted frequencies – few, some, many, and most of the analysed responses. These frequencies indicate how often a topic was mentioned relative to other topics within one question, but they should not be interpreted as numerical insight into exact proportions.

4. Quality assurance: the second analyst quality assures the analysis by checking the greater of either 5% or 10 of the main analyst's coded responses, with a minimum of 80% agreement needed.

For more on thematic analysis and codebook thematic analysis, see the Toward good practice in thematic analysis: Avoiding common problems and be(com)ing a knowing researcher article.

## Strengths and limitations

In addition to the previously described limitations of the overall HIS, there are some strengths and limitations specifically related to the qualitative (thematic) analysis.

- Compared with many qualitative studies that analyse fewer but more in-depth information, the free-text data in the HIS represents a large number of short survey responses.

- By taking the codebook thematic analysis approach and decreasing bias, inferences can be made about experiences of a larger population, where sub-themes relate to most or many respondents.

- The free-text questions provide an inductive insight, which allows respondents to share experiences without being guided by pre-set response options.

- The responses are limited to 250 characters, limiting how much depth of information respondents can provide.

- Despite approaches that decrease bias, qualitative analysis is inherently related to interpretations of the researcher and is never completely objective.

- Responses from certain groups, particularly older people, are more common, leading to their experiences being over-represented.

- By taking the data saturation approach, some individual experiences may be missed because we do not read all the responses provided; however, considering the focus on common patient experiences to inform policy, missing rare experiences is less of a concern.

The reported frequencies (few, some, many, most) indicate how often a topic was mentioned but they do not reflect how important a topic was to a respondent.

# 8 . Related links

Experience of NHS healthcare services in England
Bulletins
Experiences of local GP services, NHS treatment waiting lists, dentistry and pharmacy services, analysing data from the Health Insight Survey commissioned by NHS England. These are official statistics in development.

Experiences of NHS healthcare services in England
Dataset
Experiences of local GP services, NHS treatment waiting lists, dentistry and pharmacy services, analysing data from the Health Insight Survey commissioned by NHS England. These are official statistics in development.

Experiences of GP services in England, thematic analysis: June to September 2025
Technical article
Qualitative analysis of reasons for patients' positive or negative experiences of general practice (GP) services, using data from our Health Insight Survey.

Experiences of NHS healthcare services in England: thematic analysis
Qualitative dataset
Qualitative findings on experiences of NHS services in England, analysing free-text responses from the Health Insight Survey (HIS) commissioned by NHS England. Each dataset presents findings from three waves (12 weeks) of the HIS.

More information about the Health Insight Survey and why it was created is available on our About the Health Insight Survey web page.

# 9 . Cite this methodology

Office for National Statistics (ONS), updated 26 February 2026, ONS website, quality and methodology information report, Experiences of NHS healthcare services in England quality and methodology information (QMI)