Kidney PREM 2020 Technical Report

About the Kidney PREM 2020 Data

The Kidney PREM consisted of 39 questions covering 13 themes plus an overall experience question, using the same questionnaire as in 2018 and 2019. Patients responded to each question on a scale from 1 to 7. All questions had the option of "don't know" and "not applicable", with the exception of question 39 "Your Overall Experience". The themes and related questions with the response scale can be seen in table 1.

Table 1: Themes in the 2020 Kidney PREM, with the response scale.

Section	Theme	Questions	Response scale
1	Access to the Renal Team	Q1-Q3	1 Never – 7 Always
2	Support	Q4-Q6	
3	Communication	Q7-Q11	
4	Patient Information	Q12-Q13	
5	Fluid Intake and Diet	Q14-Q15	
6	Needling	Q16[a]	
7	Tests	Q17-Q19	
8	Sharing Decisions About Your Care	Q20-Q22	
9	Privacy and Dignity	Q23-Q24	
10	Scheduling and Planning	Q25-Q26, Q27[a]	
11	How the Renal Team Treats You	Q28-Q30	
12	Transport	Q31-Q33[a]	
13	The Environment	Q34-Q38	1 Poor – 7 Excellent
14	Your Overall Experience	Q39[b]	1 Worst it can be – 7 Best it can be

Those questions marked (a) referred to filtered questions, where only a subset of patients was required to answer depending on their treatment. Question 39 (b) was a question about patients' overall experiences. In addition to the survey data, information about patient characteristics was collected for current treatment (including CKD and transplant, and location if haemodialysis), age category, gender, ethnicity and use of PatientView.

Data Collection Process

Online data collection was over a period of six weeks from 1st October to 12th November 2020. Paper surveys were not made available, to reduce the burden on staff and in accordance with infection control measure in place during the COVID-19 pandemic.

Each centre across England, Wales, Scotland and Northern Ireland was invited to take part in the 2020 survey. The Kidney PREM survey was made available online via the Kidney Care UK website (https://www.kidneycareuk.org/prem-2020), with patients able to select their centre from a drop-down list. Space was given to type a centre/unit name if patients were unable to find the correct unit.

Renal centres were encouraged to promote the Kidney PREM as much as possible using their existing communication tools, such as email, text messaging or by post. Business cards and posters were provided, highlighting the website address. In addition, some patients may have found the survey online without being told it was available. As well as English, the questionnaire was available in Welsh, Gujarati and Urdu.
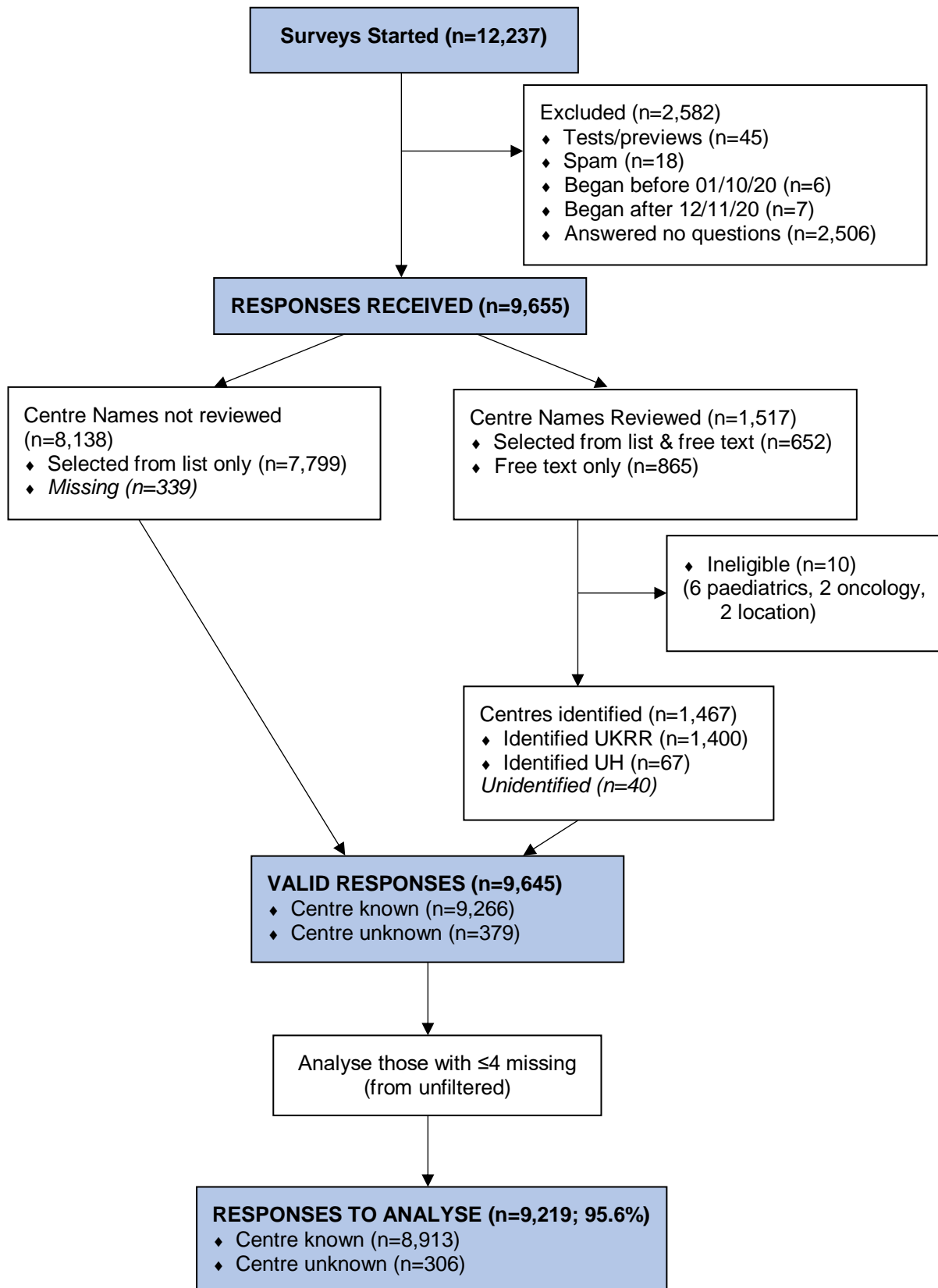
Data Cleansing

All data analysis was performed using Excel or Stata/IC version 15.1. Figure 1 illustrates the initial data cleaning process followed to define the number of "valid responses". The criteria used to define these were: completed between 1/10/20 and 12/11/20, answering at least one survey question (from Q1-Q39) and from an eligible renal service. Patients could select their centre from a drop-down list or enter text into the "Unit Free Text" box, either in addition to or instead of selecting a unit. Patients answering about paediatric centres, oncology services and units which were not based in the UK were removed from the dataset. All survey responses with free text entered were independently reviewed by staff at both The Renal Association and the University of Hertfordshire. Out of 1,517 with free text entries, 1,032 changes were made to centre names. Correct centres had already been selected for 485 responses with entered text. Once the review process was complete, UK Renal Registry (UKRR) codes were added to the data.

After the initial data cleaning, 9,645 valid responses were identified, of which 9,266 centres were known and 379 could not be determined.

Patient characteristics variables (treatment, treatment location, age, sex, ethnicity and PatientView use) were encoded to allow for analysis. The patient summary characteristics table was produced based on all patients who provided valid responses and percentages were calculated across each characteristic (age, gender, ethnicity, treatment and haemodialysis location).

Figure 1: Flowchart detailing the Kidney PREM 2020 data cleaning and analysis process

Surveys Started (n=12,237)

Excluded (n=2,582)
- Tests/previews (n=45)
- Spam (n=18)
- Began before 01/10/20 (n=6)
- Began after 12/11/20 (n=7)
- Answered no questions (n=2,506)

RESPONSES RECEIVED (n=9,655)

Centre Names not reviewed (n=8,138)
- Selected from list only (n=7,799)
- *Missing (n=339)*

Centre Names Reviewed (n=1,517)
- Selected from list & free text (n=652)
- Free text only (n=865)

- Ineligible (n=10)
(6 paediatrics, 2 oncology, 2 location)

Centres identified (n=1,467)
- Identified UKRR (n=1,400)
- Identified UH (n=67)
Unidentified (n=40)

VALID RESPONSES (n=9,645)
- Centre known (n=9,266)
- Centre unknown (n=379)

Analyse those with ≤4 missing (from unfiltered)

RESPONSES TO ANALYSE (n=9,219; 95.6%)
- Centre known (n=8,913)
- Centre unknown (n=306)

www.renal.org/kidney-patient-reported-experience-measure

Estimation of Scale/Sub-Scale Scores

Sub-scale scores were estimated for each theme, for each patient. This produced a score for each theme, with equal weight given to every question. If questions were not answered then theme sub-scale scores could still be estimated, so long as there was no more than 1 question missed. For themes containing just one question (Needling and Overall) a score could not be estimated if missed. Questions for themes 6 (Needling) and 12 (Transport) were only asked of in-centre and in-satellite haemodialysis patients. Question 27 (within theme 10) related to blood tests and so was asked of all other patients.

Theme 10 (Scheduling and Planning) contained only one filtered question out of three, so scores were estimated using the unfiltered questions if applicable.

The overall scale score was estimated excluding question 39 (Your Overall Experience). This left 38 questions, of which five were filtered, leaving 33 which were applicable to all patients. Scale scores were estimated if there were less than four missed questions from the 33. Those where "Don't Know" or "N/A" was selected (approx. 9%) were not considered missing but could not be used to calculate scale scores.

In previous years, sub-scale and scale scores were not estimated for patients with missing centres because centre staff completed these themselves and so very few were missing. This year, scores were still calculated for those with missing centres because it was felt to be important to specifically analyse this group of patients in case any systematic difference between them and others could be identified.

Responses were deemed "Responses to Analyse" if the number of missed questions was ≤1 for each theme or ≤4 for the total scale score. Of the 9,645 valid responses, 9,219 (95.6%) qualified as Responses to Analyse.

Sub-scale and scale scores were each calculated using the following algorithm:

1. "Don't Know" and "N/A" responses were recoded as missing
2. The number of missed responses from each theme and the overall score were calculated (from unfiltered questions) (M)
3. The total score for each theme and the overall scale was calculated (R)
4. Each scale or subscale score was calculated: $= \frac{R}{Q-M}$ where Q is the number of questions being evaluated

Mean Scores by Centre

Mean sub-scale and scale scores were calculated using patient scores across each centre. Scores were only reported if there were at least seven responses per centre. All centres had enough responses to be reported throughout, although several centres had insufficient in-centre and in-satellite haemodialysis responses for the Needling and Transport theme scores to be estimated.

The centre scores estimated in this way were used for reporting in the following report tables:
- Appendix D, a summary of the highest and lowest mean scores by centre, with the range in centre scores, for 2020, 2019 and 2018;
- Graphs of mean 2020 theme scores by centre; and

- Graphs showing 2019/20 theme comparisons (available online).

The caterpillar plots in the Kidney PREM report (under the heading Patient Experience Scores According to Treatment Type and by Centre) provide a visual guide to variation between centres across the 13 themes and the overall question. Each plot (one per theme) shows the median, lower quartile and upper quartile for all centres as a vertical line. For each theme, the data was sorted in descending order by centre, and the mean value for each centre and 95% confidence intervals are shown. Centres with less than seven responses for any theme were excluded from the graphs.

For the waterfall plots (2019/20 theme comparisons), mean scores across centres were calculated for each theme. 2019 mean scores were previously calculated (excluding centres with fewer than 7 responses). The data was sorted in descending order of the 2020 means and by centre.   The plots show the 2020 and 2019 means for each centre and theme, with the overall mean values for 2019 and 2020 as vertical lines.

Scores by Treatment Type

To explore differences in scores by treatment modality, medians and interquartile ranges were calculated for each theme. Unlike with centre scores, patient scores were left-skewed, so the small number of treatment types compared to the number of centres (6 v 71) meant that medians were a useful way to measure changes in scores across most patients within each group. Medians and interquartile ranges were also calculated for 2019 data so that comparisons could be made. Scores for Needling and Transport could only be estimated for in-centre and in-satellite haemodialysis patients.

Although means of left-skewed data are influenced by low scores, they are still a useful measure for comparisons across treatment groups. Therefore, means and 95% confidence intervals were also calculated and are shown in appendix E. These values were also used to produce plots for each theme, again under the heading Patient Experience Scores According to Treatment Type and by Centre in the main report.

In addition to the data reported in the main report, tables were produced displaying for each centre:
- Means and 95% confidence intervals of each theme by treatment modality
- Means of each theme by treatment modality for 2020, 2019 and 2018.

For each item under consideration, data was removed if there were fewer than 7 responses, to limit the potential of patient identification.

Question Response Centre/Unit Data

In addition to the main Kidney PREM report, question-level response data was made available via an online portal. Some relabelling or recoding was done to improve the readability of the dataset. Patient answers of any unit or centre with less than 10 responses were removed to preserve the anonymity of individual patients (note the report plots exclude centres with 7 or fewer responses).

Questions answered by any patients included as "Valid Responses" were included in this presentation. Data for previous years' Kidney PREMs remain online and can be compared to 2020 results for each centre.

www.renal.org/kidney-patient-reported-experience-measure

To produce the data in the required format to be used on the portal, each question was tabulated in Stata by main unit and site, then exported to Excel. Percentages were calculated (excluding missing responses, don't know and not applicable) for each table and added to the Excel spreadsheet.

Data in the portal is grouped by geographical location (Country, region, main unit and site). Data presented in the bar-chart and the table can be separately expanded or contracted to amalgamate sites, centres, regions or countries by using the small (+) and (-) symbols which appear when a user hovers over the geography title in the chart or table. In addition, it is possible to restrict the data to regions, centres or satellites using the filters to the right of the table/chart. Individual questions are selected using the panel under the table.

It is not possible to select multiple questions simply because of the volume of data which would potentially be displayed within the panels.

Data in the portal is presented as either numbers of people who gave each response (one to seven, not applicable, don't know or missing), or as a proportion of total group who gave a numerical response (i.e. excluding the NA, don't know and missing responses). In both cases, hovering over the column or the cell displays both the value and proportion in a tool-tip.

It is crucial to consider the number of people making a response before making a judgement on whether the proportion who responded in that way is large enough to allow firm conclusions to be drawn.

Limitations and caveats for interpreting the plots and data tables

Presenting data to allow for meaningful interpretation is always challenging. As in previous years, additional tables have been provided providing means and confidence intervals for each of the 13 themes and for each treatment modality for each centre, which adds to the information provided by the caterpillar and waterfall plots. Any summary of data (means, intervals) leads to loss of information but increases the ability to make sense of trends across different groups.

The distribution of responses across the response options (1-7) in the Kidney PREM 2020 data does not follow a "normal" distribution. Patients tended to score their experience using high (5/6/7) responses rather than 4 or less, referred to as a left-skewed distribution. A common way to deal with a skewed distribution is to use a median with quartiles to display the distribution of the data, as used to demonstrate variation across treatment types. However, most questions and themes have a median of 6 or 7 so considering the median as the central tendency is not as sensitive to variation.

In addition, a very large number of responses from the Kidney PREM means that statistical reasons for reporting the median and quartiles are less important, so centre means and 95% confidence intervals provide a robust picture of the responses for most patients, making it straightforward to compare different groups. If the interval for a particular group falls outside of the interquartile range (above the 25th or below the 75th percentile) for the group as a whole, then most people responding will be within that range, so the centre can be considered to fall below or above the relevant percentile.

Interpreting the number of responses

The choice of a mean and 95% confidence interval means that the confidence interval provides less information about the "tail" of responses in the lower parts of the response scale, which should be kept when interpreting the data. The width of confidence intervals are sensitive to sample size since a confidence interval is inversely proportional to the square root of the sample size (the confidence interval is produced by dividing by the square root of the sample size). Therefore, as the sample size increases the confidence interval generally gets smaller. If everything else remains the same, the confidence interval for a sample size of 15 will be twice as big as a sample size of 60 (e.g. interval/ $\sqrt{4}$). This is critical for small sample sizes of less than 30 where the confidence interval is likely be very large. A large confidence interval indicates uncertainty, reflected in the idea that a small sample size is not representative of the population. On the other hand, a very small confidence interval may simply reflect a very large sample size and give the false impression of difference where the clinical or psychosocial meaning is less obvious.

It is particularly important that where there are a small number of responses that caution is taken in interpreting what this may mean. It is common for numbers to be translated into percentages, but this may be misleading.

Testing a new drug can be considered as an example. If 3 out of 5 patients respond well to the drug, does the drug work for 60% of patients? In these circumstances there is an element of chance. If a different group of 5 patients is considered where only 2, or perhaps 4 patients respond to the drug, does the drug works for 20%, or 80% of patients? The problem is the uncertainty related to small numbers. Where 100 patients are tested, there will be more confidence in the observed numbers of responses, say 56/100. But even then, the chance of another sample of patients differing by as many as ±10 is considerable. Just looking at the percentage of responses in a particular group, without considering the number of responses may be misleading.