Changing Modes on the Fly: Transitioning a Complex Longitudinal Survey from In-Person to Phone due to COVID-19

Becky Reimer[1], Kylie Carpenter[1], Ann Bisognano[1], Liz Kantor[1]
[1]NORC at the University of Chicago, 55 E. Monroe Street, Chicago, IL 60603

Abstract
Typically, when long-standing longitudinal survey projects are transitioned from one data collection mode to another, researchers devote substantial time and resources to planning, field testing, and evaluating the change before implementation. With the arrival of COVID-19 in 2020, many field survey projects were forced to change course without any such research plan. Among them was the Medicare Current Beneficiary Survey (MCBS), an in-person survey of a nationally representative sample of the Medicare population, conducted by the Centers for Medicare & Medicaid Services (CMS) through a contract with NORC at the University of Chicago. The MCBS transitioned to phone administration in March 2020 after a brief pilot test phase in production. The present study aims to measure and understand the impact of the mode transition on the quality of the data collected across a wide scope of MCBS questionnaire variables. Many techniques included in the body of literature on mode analyses (de Leeuw & Berzelak, 2016) do not apply to this work due to the lack of experimental design and synchronous data collection across multiple modes. Instead, we assess the stability of response patterns across several years of data collected in-person prior to the pandemic using a model-based approach to determine the degree to which the data collected via phone in 2020 continued or broke those trends. The analysis encompasses all three rounds of data collection within the 2020 administration of the MCBS, including both the Community and Facility components. This analysis focuses on changes in reporting of health care events and their associated costs, which were typically collected by interviewers manually reviewing documents (e.g., statements) and abstracting information from them before the transition as compared to interviewers guiding respondents on reporting this information during the phone interview. We discuss difficulties separating the impact of the mode change from the impact of the pandemic itself on MCBS beneficiaries.

Key Words: survey mode, mode analysis, in-person interviewing, telephone interviewing, data quality, response patterns, COVID-19

1. Introduction

When making survey mode decisions for new projects, researchers typically consider a range of factors, including characteristics of the survey population of interest, complexity of the data being collected, context or interpretation of specific questionnaire items, cost implications, timing considerations, interviewing resources, etc. (de Leeuw, 2008; de Leeuw & Berzelak, 2016; Blair, et al., 2014). Those contemplating changes in mode for longitudinal survey projects face the same considerations, along with potential breaks in trends and changes in data quality, and typically devote substantial time and effort to planning, conducting field tests or methodological experiments, and evaluating the potential change before implementation (Dillman, 2009). The emergence of the COVID-19 pandemic in 2020 led many large-scale survey operations to rapidly transition away

from in-person data collection, which did not allow for careful methodological considerations. The present study focuses the impact on data quality of one such rapid mode change implemented due to the pandemic: transitioning the Medicare Current Beneficiary Survey (MCBS) from in-person to phone interviewing. In particular, we focus on the impact on the quality of data collected related to health care utilization and the costs associated with that care, which is particularly vulnerable to mode changes as it often is supported by the review of records by interviewers.

To protect the health and safety of respondents and field interviewers, the MCBS transitioned from in-person to phone data collection in March 2020 after a brief pilot test phase. Due to the nature of the pandemic combined with the urgent need to continue data collection, this rapid change did not allow for in-depth field testing, the incorporation of an experimental design, or an evaluation before implementation. Staged analyses were conducted during and after data collection to evaluate the data quality impact of this unplanned mode change. This work has implications both for future MCBS mode decisions as well as other complex surveys with the potential to switch data collection modes.

1.1 Introduction to the MCBS

The Medicare Current Beneficiary Survey (MCBS) is a continuous, multipurpose survey of a nationally representative sample of the Medicare population, conducted by the Centers for Medicare and Medicaid Services (CMS) through a contract with NORC at the University of Chicago (NORC). Since 1991, the survey has typically been conducted in-person and serves as the leading source of information on the Medicare program and its impact on beneficiaries. The MCBS uses a round-based rotating panel design to collect data for beneficiaries at three points (e.g., rounds) per year (referred to as winter, summer, and fall interviews) over four years for beneficiaries living both in community and facility settings. The MCBS covers a wide range of topics including beneficiary health status, health related behaviors, health outcomes, access to care, satisfaction with care, insurance coverage, cost of health care services, and sources of payment for those costs. The present study focuses on two key topics: health care utilization and the costs associated with that utilization.

While utilization and cost are among the most important collected in the MCBS, they are also the most complex and burdensome to implement. The continuous nature of health care among the Medicare population supported the initial justification for the longitudinal design of the MCBS, because it is necessary to collect information at regular intervals throughout the year to maximize recall of health care events that have occurred since the prior interview. In addition, the survey has traditionally been conducted in-person because it relies heavily on documentation to support the collection of the complex cost data associated with care, and it is typically easier for trained interviewers to extract details from this documentation than to ask respondents to do so. Given that health care cost information and documentation is not always available in a timely fashion, the MCBS allows for follow-up on cost information across multiple interviews during the year, and respondents are asked to save their documentation between interviews.

The MCBS Community and Facility components have several important differences in their administration and content regarding utilization and cost. While the Community interview takes place with either the beneficiary or a knowledgeable proxy respondent (often a family member, caregiver, or friend), the Facility interview is conducted with facility staff. Due to these differences and the ongoing nature of care within long-term care facilities, the structure of the utilization and cost data collected, as well as the supporting

documentation requested by the field interviewers, differs between the two settings (as shown in Table 1). In both settings, interviewers verbally ask survey questions for many items, and are trained to locate specific details about utilization and costs within documentation in a process referred to as "abstraction". For the Community interview, this primarily occurs within the cost sections of the survey, where interviewers sort through statements, bills, receipts, and other documentation to locate dollar amounts, claim numbers, and other details. They also abstract information about prescription medicine utilization from medicine bottles, pharmacy receipts, and other documentation. In facility settings, interviewers are frequently provided with medical records and billing records by facility staff and rely on abstraction for a larger portion of the data collected than in the Community interview.

Table 1: Data Collection Details for Community and Facility Interviews

Domain	Community Interview	Facility Interview
Structure of data collection for utilization	Exhaustive collection: ▪ Individual health care events	Aggregated collection: ▪ Use of care/services ▪ Frequency of visits
Structure of data collection for cost data	For each health care event: ▪ Charges ▪ Payments	For each billing period (e.g., month): ▪ Charges ▪ Payments
Types of documentation requested	▪ Medicine bottles ▪ Insurance statements ▪ Bills and receipts	▪ Facility medical records ▪ Billing records

1.2 MCBS 2020 Mode Transition Details

MCBS in-person data collection was paused in mid-March of 2020 due to the COVID-19 pandemic. Interviewing resumed via phone in both the community and facility settings following a relatively short pilot testing period with a new protocol involving field interviewers contacting respondents via phone. Phone data collection continued in both settings through the end of 2020 and into 2021. The main goal of the present study is to assess the impact of the mode change on the quality of the health care utilization and cost data collected, including the impact of the increased respondent burden associated with the transition.

Along with the impact of the mode change on data collection for health care utilization and costs, we need to acknowledge the potential impact of the pandemic itself, as well as some intentional content changes made to the Community interview after the mode change. The content of the Facility interview remained as planned throughout 2020, but several important changes were made to the collection of cost information within the Community interview. Modifications were made based on qualitative feedback from field staff indicating that the transition to phone was imposing new types of respondent burden as interviewers were not physically present during the interview – some respondents were finding it particularly difficult to sort through billing and insurance statements and other documentation to locate and report specific details about the costs associated with health care events. Considering this increased burden, the Summer 2020 Community interview

was shortened and excluded all collection of utilization and cost information.[1] For the Fall 2020 data collection round, we elected to include the utilization and cost sections again[2], but included an "escape hatch" mechanism that allowed interviewers to easily skip over some or all data collection of cost information in situations where respondents were particularly fatigued or frustrated. These content changes complicate any analysis of the mode change as it intentionally creates missing data attributable to interviewer and respondent behavior.

1.3 Analysis Background

Like many other large-scale surveys in 2020, the MCBS followed a much different trajectory when switching from in-person to phone interviewing due to the COVID-19 pandemic. In particular, we were not able to conduct a field test or incorporate an experimental design comparing in-person to phone administration (Schräpler, et al., 2010; Jäckle, et al., 2006; Watson & Wilkins, 2011). As a result, many of the techniques included in the body of literature on mode analyses do not apply to this work, and analytic techniques are constrained (de Leeuw & Berzelak, 2016; Olson, et al., 2019).

This analysis focuses on quantifying changes in response patterns by comparing the data collected via phone against previous rounds of in-person data collection. The purpose of this is to measure and understand the impact of the transition from in-person to phone data collection on the quality of data collected. Specifically, we sought to determine 1) how the transition from in-person to phone data collection affected respondent recall and willingness to provide data, and 2) whether the data collected by phone departed from historical trends in several areas: quantity of health care utilization and costs reported, assisted recall via documentation, and item level response. We also considered the extent to which we can attribute these departures to the data collection mode change.

2. Methods

2.1 Data Sources

The majority of the data used for these analyses come from the Community and Facility MCBS questionnaire data. These data were merged with paradata including case management details and length of time since prior interview (reference period). For the sake of performing these analyses in a timely fashion, we analyzed data on a round basis as each data collection round unfolded during 2020, using raw data for all comparisons, rather than waiting for annualized, cleaned data. This means that we worked with data that had not yet undergone cleaning or filtering to remove cases that would ultimately not be eligible for MCBS annual data products. As a result, our analytic datasets included more "noise" than the final, cleaned data released by CMS after considerable data editing and

[1] CMS and NORC made this decision during April 2020, when the trajectory of the pandemic was not yet clear. At the time, our focus was on minimizing the attrition of the longitudinal panels while collecting additional data in a new series of Rapid Response COVID-19 Supplements using the MCBS sample.

[2] Because utilization and cost information had not been collected during Summer 2020 interviews, this change involved updating the reference periods, such that respondents were asked to report all utilization and costs going back to the date of their Winter 2020 interviews. In prior fall rounds of data collection conducted in-person, respondents typically report on utilization and costs that occurred during the past 4-7 months, but due to this change, respondents in the Fall 2020 interview were asked to report this information during the past 8-10 months.

processing. However, the types of errors present (e.g., skip errors) tended to be randomly distributed, the number of cases that are filtered out as ineligible are typically quite low, and this decision should not have biased these analyses. By using raw data for the entire analysis, we put all rounds of data collection on relatively equal footing when performing comparisons. These analyses were conducted separately for the Community interview and Facility interview. Figure 1 illustrates the timing of each round of data collection in 2020. Within the figure, each of the six colored bars with phone icons represent segments of the analysis comparing the phone data against prior years of in-person data.

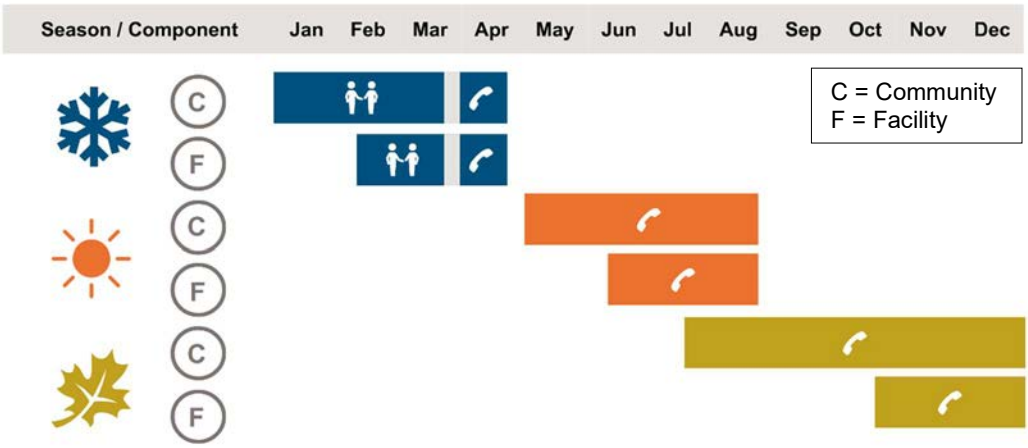


Figure 1: 2020 timeline of data collection mode transition

2.2 Process

In the absence of an experimental design, we crafted a framework for evaluating changes in response patterns over time and between data collection modes for Winter 2020, Summer 2020, and Fall 2020. These response patterns included a wide variety of health care utilization and cost-related metrics, as shown in Appendices A and B.

We evaluated the stability of response patterns in the years prior to the mode transition using regression models. These models predicted specific questionnaire metrics (e.g., the total number of new health care events reported within an interview and whether a particular item was answered or not), using a variety of socio-demographic variables and paradata variables as predictors. To assess changes over time, we also included dummy variables for each specific prior year and determined whether these indicators were or were not statistically significant predictors. Assessing this stability provided historical context for any changes observed after the transition. Once this context was established, we compared the newly collected data from the phone interviews against the combined pool of data from prior rounds using pooled T-tests.

Table 2 summarizes the rounds of data collection used for each analysis. Historical data from four prior years (2016-2019) of the appropriate season were compared to the 2020 phone data to assess trends over time.

Table 2: Rounds of data collection included in the mode analysis

Year	Winter	Summer	Fall
2016-2019	In-person	In-person	In-person
2020	In-person and phone	Phone	Phone

The Winter 2020 data collection round presented unique challenges for this analysis because it included both in-person interviews prior to the introduction of pandemic restrictions and phone interviews after the mode transition had occurred. In an attempt to cleanly assess the impact of phone interviewing within Winter 2020, we decided to separate interviews based on this timing and ignore the in-person interviews. For Facility interviews, data collection was in-person for weeks 1-10 and via phone for weeks 11-16. For Community interviews, the timing of the transition was slightly different, with data collection occurring in-person for weeks 1-11 and via phone for weeks 12-16. To avoid unfair comparisons against prior years of data, we decided to limit our analytic universes for prior round winter data to the same time frames (meaning week 11 through the end of the round for the Facility interview and week 12 through the end of the round for the Community interview). The reasoning for this is that interviews completed in the early weeks of a data collection round often differ in important ways from those completed later on, both due to strategic decisions regarding case release and to factors intrinsic to the ease of contacting individual respondents. In general, we tend to observe that interviews completed later in the data collection round are associated with respondents who are harder to reach or retain in the survey; anecdotal evidence suggests this may be related to poorer beneficiary health and/or higher health care utilization. Appendix C shows the N sizes for the Community and Facility interviews, including this limitation.

2.3 Analytic Domains and Metrics

While the MCBS covers a wide variety of content areas, the present study focuses on changes in response patterns specifically related to health care utilization and costs. The Community interview collection of this complex information is structured with the goal of minimizing respondent burden in mind (where respondents are beneficiaries or knowledgeable family members or friends). It relies fairly heavily on respondent recall, particularly for the collection of exhaustive details about individual health care events (i.e. provider visits, prescription medicines, medical equipment rentals). It also allows for documentation to supplement respondent recall, primarily in the collection of cost information, which relies on statements, bills, and receipts. The Facility interview is structured to accommodate facility operation, with the expectation that facility staff are likely to provide interviewers with facility records rather than rely on their own recall. For this reason, the Facility interview aggregates the collection of utilization and cost information by asking about health care usage and frequency and collects charges and payments by billing period.

As noted above, a detailed list of outcome metrics included for the Community and Facility analyses are included in Appendices A and B, respectively. For Community utilization and cost, a number of outcomes focus on counts of newly entered events and costs within the current round's interview. These represent new care that occurred during the reference period since the last interview. While the MCBS does allow for follow-up on cost information from one interview to the next (e.g., in situations where a beneficiary is expecting an insurance statement but has not yet received it in a given interview), the comparisons of cost counts in this analysis are limited to new costs created within a particular interview, not follow-up information on costs initially reported in prior rounds.

It is important to note that the count-based variables are likely to vary based on the timing between interviews.[3] To standardize these measures, we computed the number reported per 100 days in the reference period.

3. Results

For the sake of brevity, we have elected not to include regression model results in this section. Nearly all models comparing response patterns between prior rounds of in-person data collection showed some statistically significant differences. Most response patterns have some level of fluctuation over time, but even very slight changes between rounds were often associated with high levels of statistical significance in these models due to the large sample sizes on the MCBS. Rather than impose an arbitrary threshold to determine which statistically significant results in prior rounds represent stable trends with some expected fluctuation and which represent unstable trends, we have elected to visually display these trends in comparison to the 2020 estimates. The results reported below, including notes about statistical significance, focus on the pooled T-tests comparing prior years of data collected via in-person interviewing (2016-2019) against the 2020 data collected via phone.

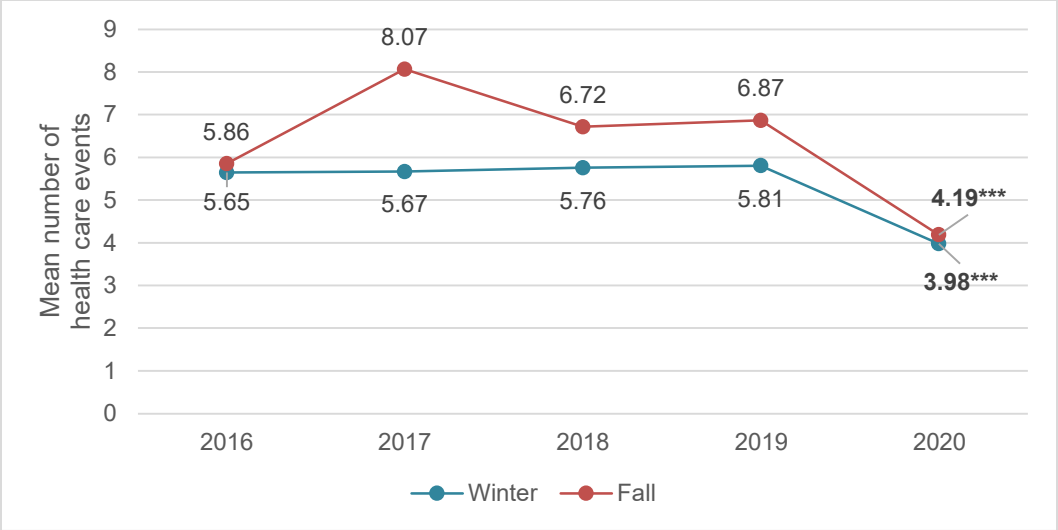
3.1 Community Health Care Utilization

Significant declines in utilization were seen in both Winter 2020 phone interviews and Fall 2020 interviews as compared to previous winter and fall rounds, respectively. We did not collect utilization information in the Summer 2020 interview, but the Fall 2020 interview included a longer reference period to capture summer utilization (and we standardized the event count outcomes to account for the longer reference period in our comparisons). It should be noted that reporting of utilization is typically slightly lower in winter rounds of the MCBS than fall or summer rounds, which is likely due in part to the holiday season and colder temperatures.

3.1.1 Number of health care events reported

As shown in Figure 2, there were considerable, statistically significant decreases in event reporting (adjusted for the length of the reference period) in the Winter 2020 phone interviews and Fall 2020 interviews as compared to the four prior rounds of each corresponding season. We believe much of these decreases were due to actual declines in utilization during the COVID-19 pandemic (Cox, et al., 2021).

[3] This is particularly true for the Fall 2020 data collection round of the Community interview, which had a longer reference period than is typical for fall rounds. This was due to the decision to pause collection of all utilization and cost information during Summer 2020 and resume in Fall 2020.

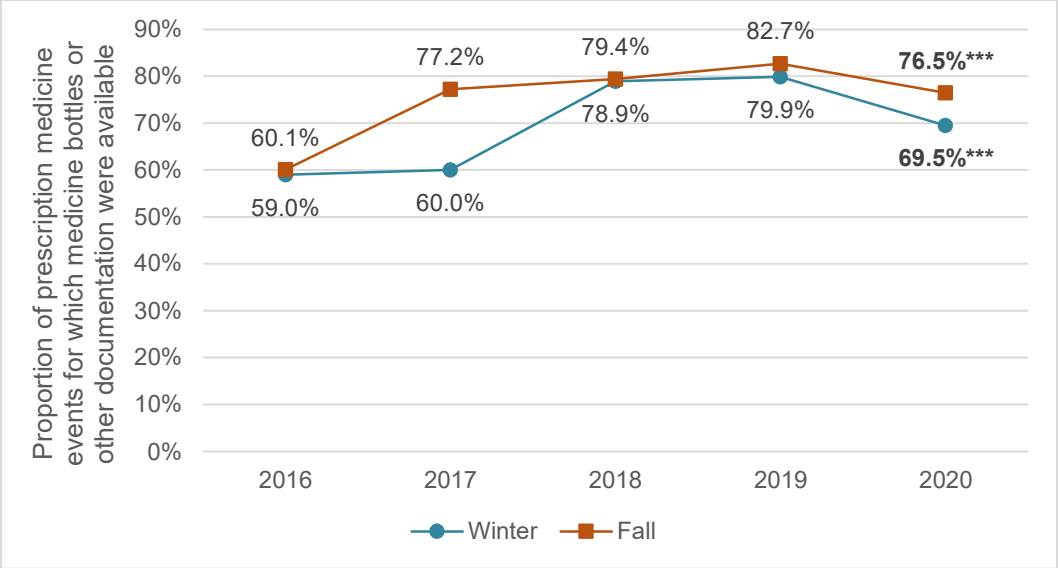


* p < 0.05, ** p < 0.01, *** p < 0.001

Figure 2: Mean number of health care events reported per 100 reference period days

3.1.2 Prescription medicine events for which medicine bottles or other documentation were available to help guide data entry

We observed several statistically significant differences in prescription medicine reporting which indicate that shifts occurred with phone interviewing. As an example, the proportion of prescription medicine events for which interviewers indicated availability of documentation (e.g., prescription bottle, container, bag) to help guide data entry showed an increase over the past four winter and fall rounds, but then a substantial drop in the Winter 2020 phone interviews and Fall 2020 interviews, respectively (shown in Figure 3). Importantly, this change could be related to the adjustment in protocol associated with collecting prescription medicine information over the phone. During in-person interviews, interviewers ask respondents to provide them with access to medicine documentation and then sort through it while entering details about each medicine. During phone interviews, interviewers rely on respondents to report whether they are able to locate bottles or other documentation. It seems plausible that some respondents could be more willing to locate this documentation when interviewers are physically present than they might be during a phone interview.



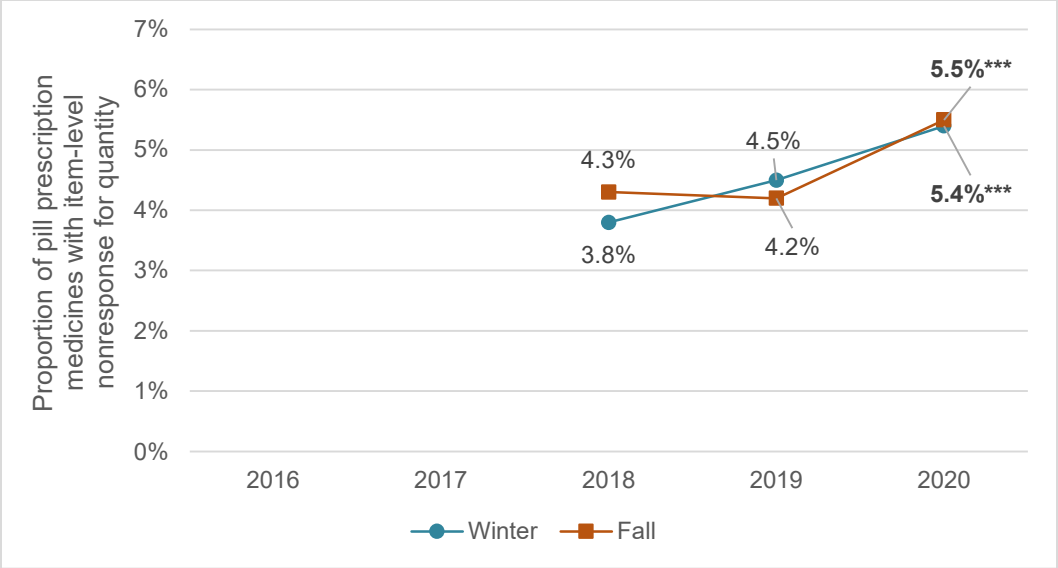
* p < 0.05, ** p < 0.01, *** p < 0.001

Figure 3: Proportion of prescription medicine events for which medicine bottles or other documentation were available to help guide data entry

3.1.3 Prescription pill/non-pill medicine quantity

Respondents' ability to provide information about the quantities of medicine obtained for prescription medicine events was lower than usual in the Winter 2020 phone interviews and Fall 2020 interviews. Item-level nonresponse increased over time and continued to climb in the Winter 2020 phone interviews and Fall 2020 interviews for the amount and unit indicators for non-pill medicines. There was also a slight, but statistically significant increase in item-level nonresponse for the number of tablets included in the bottle for both rounds (shown in Figure 4).[4]

[4] For each of these prescription medicine quantity measures, we were only able to compare 2020 data to the prior two rounds of data collection due to a questionnaire change in universe between 2017 and 2018.



* p < 0.05, ** p < 0.01, *** p < 0.001

Figure 4: Item-level nonresponse for prescription pill medicine quantity

3.1.4 Additional Metrics

We also assessed the proportion of interviews reporting zero health care events for the entire reference period. Using this metric, we uncovered several existing trends, which continue with the Winter 2020 phone interviews and Fall 2020 interviews. The proportion of beneficiaries with zero events reported is slightly higher in Winter 2020 phone interviews and Fall 2020 interviews than prior rounds. For Winter 2020 phone interviews, similar trends were observed for medical provider events and prescription medicine events, with a larger proportion of beneficiaries reporting zero of these types of events than in prior winter rounds. For Fall 2020 interviews, we observed a larger proportion of beneficiaries reporting zero dental, other medical, and vision events than in prior fall rounds. For the sake of brevity, we have elected to not include charts for these results given the lack of major changes due to the data collection mode switch or the COVID-19 pandemic.

Consistent with the prior results showing decreased levels of utilization in the Winter 2020 phone interviews and Fall 2020 interviews, we observed a downward trend in the proportion of beneficiaries who reported seeing any medical doctors during the reference period covered by Winter 2020 and Fall 2020 interviews as compared with prior rounds.

3.2 Community Health Care Costs

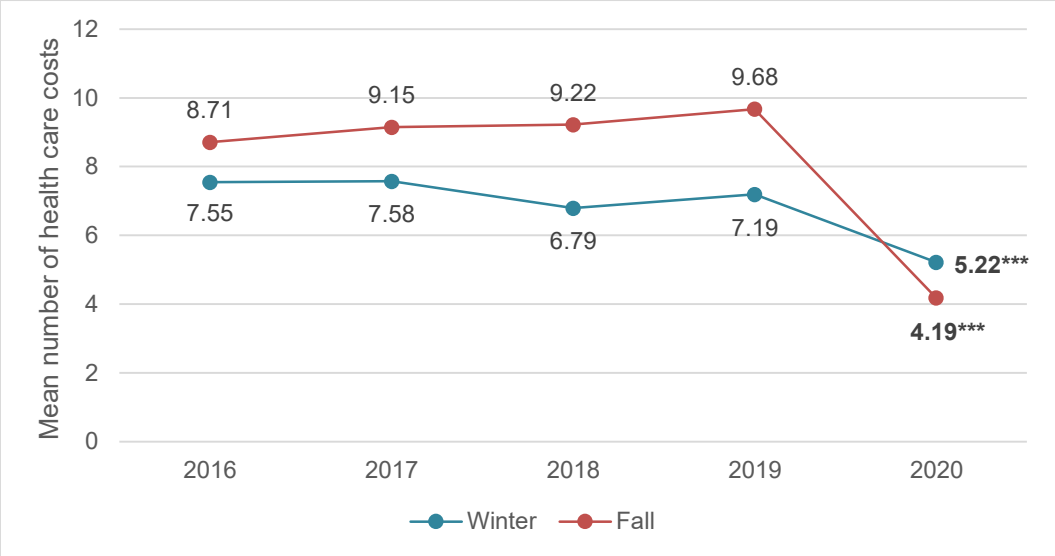
Our results related to health care cost outcomes revealed similar trends to those we observed for utilization. We were unable to examine cost outcomes for Summer 2020, as we did not collect cost information that round.[5]

3.2.1 Number of health care costs reported

Cost reporting in the Winter 2020 phone interviews and Fall 2020 interviews was lower than in prior winter and fall rounds, respectively, which was likely related to several factors

[5] As noted in Table 1, in the Community interview, costs are collected as charges for each health care event. These include total amounts charged and payments made by various sources such as Medicare, out of pocket, etc.

including lower utilization during the pandemic and potential respondent difficulty with reporting costs by phone (shown in Figure 5). Without being physically present, interviewers do not have direct access to sort and abstract from documentation. Per qualitative feedback from field staff, interviewers were relatively successful in their requests to respondents to locate, organize, and report information from billing documents, but respondents are far less experienced with this process than highly trained MCBS interviewers.

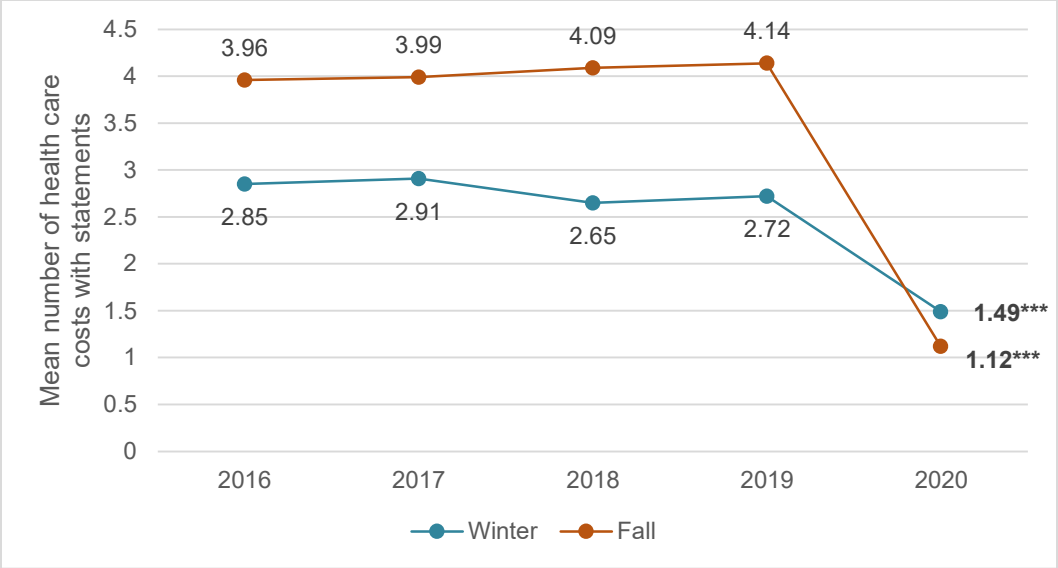


* p < 0.05, ** p < 0.01, *** p < 0.001

Figure 5: Mean number of health care costs reported per 100 reference period days

3.2.2 Number of statements reported

As shown in Figure 6, the mean number of costs associated with statements (in which cost information is sourced from documentation, such as a Medicare Summary Notice) reported in the Winter 2020 phone interviews and Fall 2020 interviews was also substantially lower than in prior rounds. Statement costs are a subset of all costs reported, and the steeper declines shown in Figure 6 compared with Figure 5 could reflect respondents' difficulty with locating and extracting relevant details from complicated documentation.



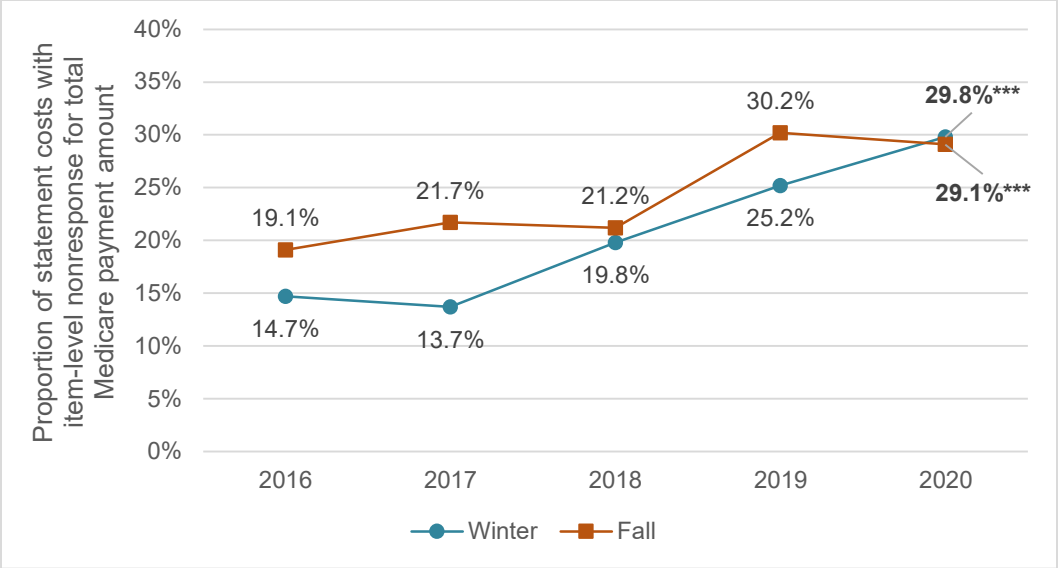
* p < 0.05, ** p < 0.01, *** p < 0.001

Figure 6: Mean number of health care costs with statements reported per 100 reference period days

Along with these mean differences, we assessed changes in the proportion of beneficiaries with zero costs for the entire reference period. We found no significant differences for the proportions with zero costs overall or zero non-statement costs, but a substantial increase for the proportion with zero statement costs in Winter 2020 phone interviews. Conversely, we found a significant increase in zero costs overall, statement costs, and non-statement costs in Fall 2020 interviews.

3.2.3 Medicare payment amount

For statement costs, we also observed that the Winter 2020 phone interviews and Fall 2020 interviews had significantly higher rates of item-level nonresponse for several items collecting dollar amounts, including total Medicare payment amount (shown in Figure 7), as compared to typical rates from prior rounds (with the exception of Fall 2019). Similar patterns were observed for total Medicare approved amount and total charge of cost/copay.



* p < 0.05, ** p < 0.01, *** p < 0.001

Figure 7: Item-level nonresponse for total Medicare payment amount

3.2.4 Additional Metrics

For more detailed cost items, we observed a number of changes that appeared to be related to difficulties with documentation during phone interviews. MCBS interviewers receive intensive specialized training for the complicated task of sorting and matching billing statements associated with the same events before cost entry in the questionnaire. With phone administration, interviewers are now relying on respondents to tackle this task, and we saw that the proportion of reported charge bundles that contain multiple statements was substantially lower in the Winter 2020 phone interviews and Fall 2020 interviews than in prior rounds. This likely indicates that respondents were less able to identify instances where multiple statements from multiple sources (e.g., Medicare and private insurance) corresponded to the same health care event/cost bundle than our trained interviewers.

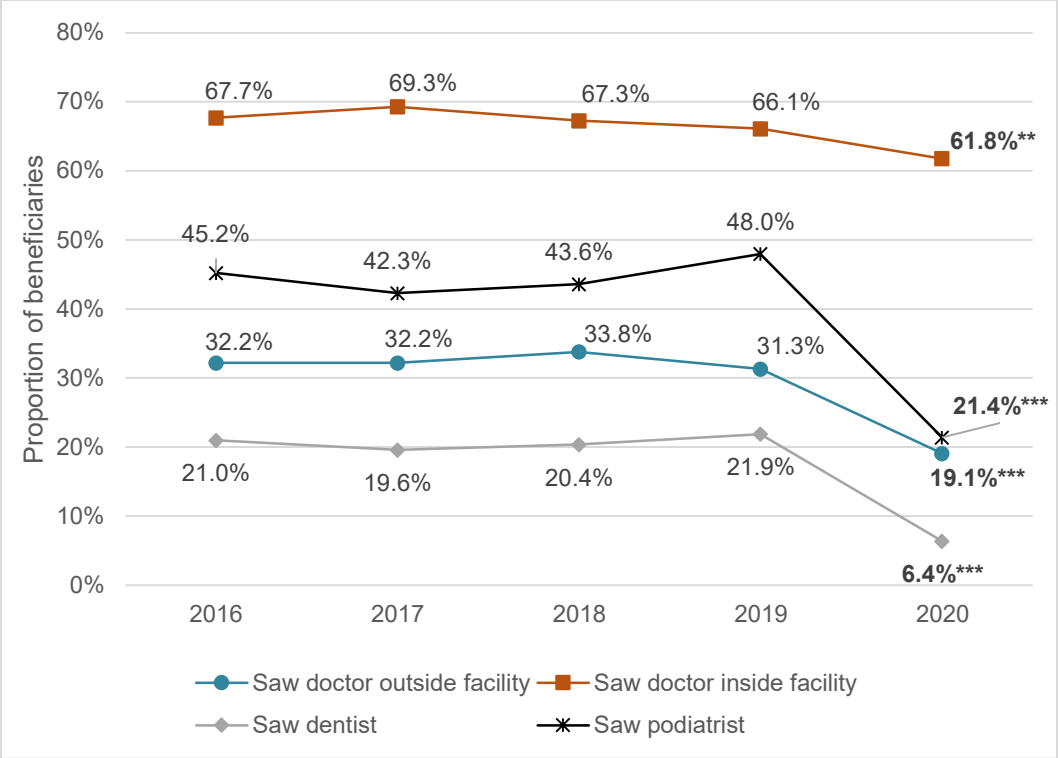
3.4 Facility Health Care Utilization

Substantial declines in utilization were revealed in Summer 2020 as compared to previous summer rounds but were not observed in Winter 2020 phone interviews or Fall 2020 interviews as compared to previous winter and fall rounds, respectively.

3.4.1 Doctor visits

The Summer 2020 Facility interview data showed significant declines in several utilization categories compared with prior summer rounds, including the proportion of beneficiaries who saw doctors inside facilities, doctors outside of facilities, dentists, and podiatrists, as shown in Figure 8. In contrast, data from the Winter 2020 Facility interviews completed by phone and Fall 2020 Facility interviews did not show substantial declines in these types of utilization compared to prior winter and fall rounds, with the exception of seeing podiatrists in Winter 2020 interviews completed by phone. This finding could be related to differences in the data collection protocols and data structure for utilization between the Community and Facility interviews. Changes may have been more difficult to detect in the facility setting because utilization information is not collected on an individual health care event basis like the Community interview and is instead measured through items asking facility staff members to report whether

beneficiaries had particular types of care during the reference period or not (with some follow-up questions regarding frequency of care). This may be particularly true for Winter 2020 interviews completed by phone, since the majority of the reference period covered time before pandemic-related factors may have impacted care.



* p < 0.05, ** p < 0.01, *** p < 0.001

Figure 8: Summer round health care utilization for beneficiaries living in facilities

An investigation of utilization items related to frequency of visits, supplies, equipment, and devices revealed few substantial changes between the in-person data and phone data. There were also no statistically significant increases in rates of item-level nonresponse for utilization items.

3.5 Facility Health Care Costs

In the Facility cost section, modest declines in data quality were observed via item-level nonresponse in Winter 2020. This decline in data quality appeared to resolve by Summer and Fall 2020 rounds.

3.5.1 Basic and ancillary charges and payments

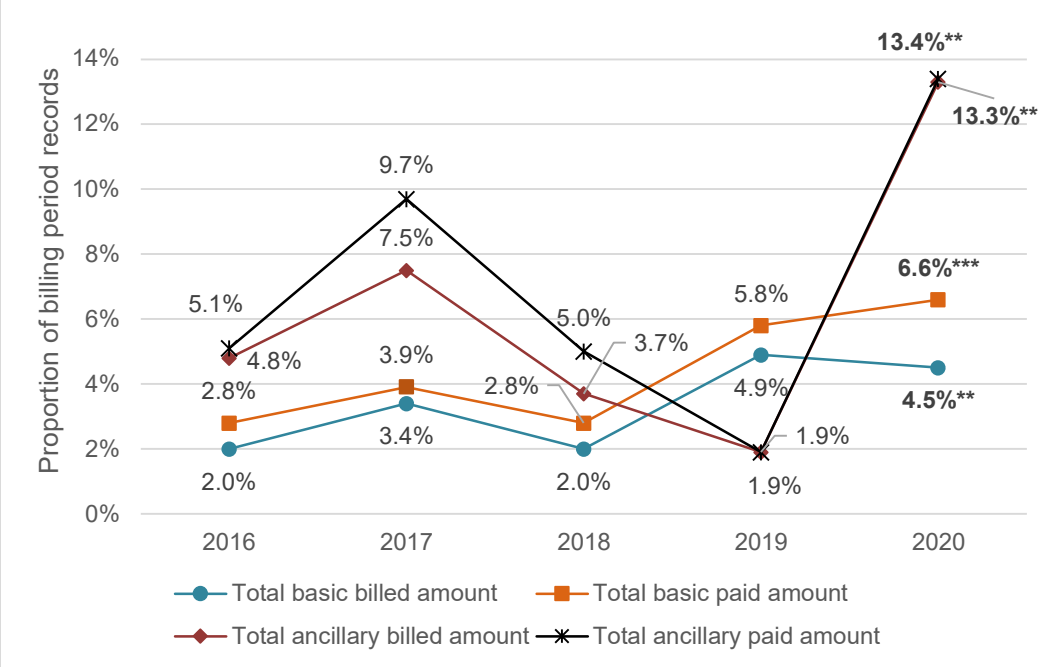
Health care cost data are collected in two main categories for beneficiaries living in facility settings: basic charge amounts, which include basic care and room and board, and ancillary charges, which are billed separately in some facilities and may cover services such as radiology, medications, lab work, or therapy.[6] For each of these categories, the MCBS collects both billed amounts and paid amounts. In prior rounds with in-person data

[6] Note that the MCBS collects facility cost data for only those ancillary services that are health-related. Ancillary charges for non-health-related services, such as haircuts, beautician services, therapeutic massage, laundry, telephone, television, etc. are excluded.

collection, facility staff often provided interviewers with physical billing records and interviewers entered the appropriate amounts into the questionnaire, but during phone administration, interviewers asked facility staff to provide this information directly.

Total basic billed amounts, total basic paid amounts, total ancillary billed amounts, and total ancillary paid amounts were analyzed for item-level nonresponse. We hypothesized that factors related to phone administration or the pandemic itself might affect the quality of reporting of cost data, possibly leading to increased item-level nonresponse. Early qualitative feedback collected from interviewers indicated that some facility staff were having difficulty accessing certain information due to the transition to remote work and not yet being fully equipped to access records remotely. We assessed item-level nonresponse for these dollar amount items, which are reported on multiple billing period records per beneficiary.

We observed statistically significant increases in item-level nonresponse between prior rounds and the Winter 2020 phone interviews, particularly for the ancillary charges (shown in Figure 9). These cost data trends from Winter 2020 did not appear to carry through to Summer 2020 or Fall 2020. Item-level nonresponse differences compared to prior rounds were smaller than those observed in Winter 2020, and some actually showed reductions, rather than increases.



* p < 0.05, ** p < 0.01, *** p < 0.001

Figure 9: Winter round item-level nonresponse for health care cost amounts

We hypothesized that we might see increases in rounding of dollar amounts if facility staff were relying more heavily on memory than usual during phone interviews, but we did not see any statistically significant differences in the proportion of dollar amount values that were rounded to the nearest dollar in any 2020 rounds as compared to prior rounds. We also did not observe decreases in the proportion of billing records indicating

that ancillary charges were posted (meaning the facility processed these charges), or in the proportion of beneficiaries who had any ancillary charges.

4. Discussion

The results of these analyses reveal several noteworthy insights concerning the impact of the transition from in-person interviews to phone interviews and the COVID-19 pandemic had on the quality of health care utilization and cost data collected on the MCBS. These observations provide useful insights for researchers looking to understand the challenges and consequences of mode shifts for large-scale studies that involve collection of highly detailed information, particularly longitudinal surveys and those that rely heavily on documentation provided by respondents.

For the MCBS, these results highlight potential tradeoffs between the operational efficiencies afforded by phone interviewing and data quality implications, and they will be used by the research team to inform decisions regarding interview mode going forward. Across all sections of the MCBS Community and Facility interviews included in the broader mode analysis of which the present study was a part, the vast majority of items analyzed had relatively stable response patterns in the transition from in-person to phone interviewing but, as expected, a number of differences emerged within the health care utilization and cost sections. While the lack of an experimental design limits our ability to isolate mode effects from pandemic effects in explaining these patterns, we believe both played substantial roles in the observed differences.

First, we noted declines in reporting of health care utilization relative to prior years. For interviews in the Community setting, these declines were observed in both Winter 2020 and Fall 2020. Declines were less pronounced in Winter 2020 than later in the year because the time period over which respondents recall utilization included multiple months of pre-pandemic utilization. In the Facility setting, utilization was relatively stable compared to prior years in Winter 2020 and Fall 2020 but declined in Summer 2020. These patterns appear to correspond with actual decreases in utilization, and this hypothesis is supported by emerging literature documenting substantial decreases in utilization and spending as a result of pandemic-related restrictions and consumer caution (Cox, et. al., 2021; Bosworth, et al., 2020).

Second, we saw limited evidence of data quality problems with the collection of utilization information in either interview setting. One notable exception was the collection of prescription medicine data in the Community interview. Reduced availability of medicine bottles and other documentation transfers the burden of extracting key information from that documentation from trained interviewers to respondents, which may have led to decreases in data quality for some details about prescription medicines. This has implications for surveys collecting detailed information based on respondent recall supported by documentation.

Third, we observed an array of changes in terms of cost reporting. For the Facility phone interviews in Winter 2020, this included increases in item-level nonresponse on certain items. Fortunately, many of the issues observed in the Winter 2020 Facility interview seem to have been mitigated in Summer 2020 and Fall 2020. This appears to have been a short-term pandemic-related impact. Qualitative feedback from interviewers revealed that many facility administrative staff had shifted abruptly to remote work and therefore, when contacted to respond for the winter interview in March or April of 2020, it often meant they

had initial difficulties accessing billing systems. As facilities adjusted to the pandemic, difficulties with remote access soon eased, and item-level nonresponse for cost questions returned to typical levels.

We also observed changes in cost reporting for the Community interview which do not seem to be as short-lived as those are for Facility. We observed substantial declines in the number of costs reported in the Winter 2020 phone interviews and Fall 2020 interviews compared with prior rounds, particularly costs associated with statements, and increases in item-level nonresponse for certain items. This was likely due to a combination of lower utilization during the pandemic and difficulties associated with increased respondent burden in the cost series of the Community questionnaire. Qualitative feedback from field staff suggested that many respondents found it burdensome to sort through statements and other documentation to provide requested details, even with supportive guidance from interviewers over the phone. New "escape hatch" functionality added to the Community questionnaire programming in Fall 2020 to allow interviewers to bypass part or all of the cost series for highly fatigued or frustrated respondents accounted for some of this decline. A follow-up analysis focused on the Fall 2020 Community interview found that 16 percent of all health care utilization events were reported without any associated costs as a result of this new functionality being used. In addition, some respondents may have intentionally failed to report cost details by passively refusing – indicating that they did not have or could not find documentation because they were already finding the interview burdensome, although we believe this behavior was relatively rare.

In general, these results suggest that lack of physical access to documentation will continue to be a difficult barrier to overcome with phone interviewing for beneficiaries living in the community, particularly for those with high levels of health care utilization and/or complicated insurance statements. It will likely be possible to populate some cost details that went unreported in the survey through MCBS post-processing steps including claims matching and imputation, but this recovery of information will be incomplete and continued analysis is needed to further assess the overall impact. This finding likely generalizes to other surveys collecting highly detailed information about recalled behavior supported by references to documentation.

Overall, we view these results in a positive light. This analysis has revealed challenges associated with phone interviewing, which we are continuing to monitor, in parallel with efforts to further investigate the ultimate impact on the quality of utilization and cost data as a result of the mode change and difficulties with physical access to documentation. However, the MCBS was highly successful in rapidly transitioning from in-person to phone administration and outperformed many of our expectations when pandemic-related restrictions first presented themselves. We have gained considerable knowledge about phone administration and demonstrated that large portions of the Community and Facility instruments can successfully be administered by phone. Importantly, it does not appear that our panels suffered from significantly higher attrition. Looking ahead, we will investigate questions regarding the extent to which we can recover information about unreported cost and health care events via claims matching and imputation, and how we can best use background information about beneficiaries (characteristics and prior behavior) to predict optimal interviewing mode.

Disclaimer

The opinions and views expressed in this work are those of the authors. No official endorsement by the Department of Health and Human Services or the Centers for Medicare & Medicaid Services is intended or should be inferred.

References

Blair, J., Czaja, R., & Blair, E. A. (2014). *Designing surveys: A guide to decisions and procedures.* SAGE Publications, Inc.

Bosworth A., Ruhter J., Sheingold S., & Zuckerman R. (2020). *The impact of the COVID-19 pandemic on Medicare beneficiary use of health care services and payments to providers: Early data for the first 6 months of 2020* (Issue Brief No. HP-2020-01). Office of the Assistant Secretary for Planning and Evaluation, U.S. Department of Health and Human Services. https://aspe.hhs.gov/sites/default/files/private/pdf/264071/Medicare-FFS-Spending-Utilization.pdf

Cox, C., Amin, K., & Kamal, R. (2021, March 22). *How have health spending and utilization changed during the coronavirus pandemic?* Peterson-KFF Health System Tracker. https://www.healthsystemtracker.org/chart-collection/how-have-healthcare-utilization-and-spending-changed-so-far-during-the-coronavirus-pandemic/

de Leeuw, E.D. (2008). Choosing the method of data collection. In E. D. de Leeuw, J. J. Hox, & D. A. Dillman (Eds.), *International handbook of survey methodology* (pp. 113–135) Taylor & Francis Group/Lawrence Erlbaum Associates.

de Leeuw, E., & Berzelak, N. (2016). Survey Mode or Survey Modes? In *The SAGE Handbook of Survey Methodology* (pp. 142-156). SAGE Publications, Ltd.

Dillman, D.A. (2009). Some Consequences of Survey Mode Changes in Longitudinal Surveys. In P. Lynee (Ed.), *Methodology of longitudinal surveys* (pp. 127-140). John Wiley & Sons, Ltd.

Jäckle, A, Roberts, C.E., & Lynn, P. (2006). *Telephone versus face-to-face interviewing: mode effects on data quality and likely causes (Report on Phase II of the ESS-Gallup Mixed Mode Methodology Project)* (ISER Working Papers). Institute for Social and Economic Research. https://www.researchgate.net/publication/5017899

Olson, K., Smyth, J.D., Horwitz, R., Keeter, S., Lesser, V., Marken, S., Mathiowetz, N., McCarthy, J., O'Brien, E., Opsomer, J., Steiger, D., Sterrett, D., Su, J., Suzer-Gurtekin, Z.T., Turakhia, C., & Wagner, J. (2019). Transitions from telephone surveys to self-administered and mixed-mode surveys: AAPOR Task Force Report. *Journal of Survey Statistics and Methodology*, *9*(3), 381-411.

Schräpler, J, Schupp, J, & Wagner, G.G. (2010). Changing from PAPI to CAPI: Introducing CAPI in a longitudinal study. *Journal of Official Statistics*, *26*(2), 239-269.

Watson, N., & Wilkins, R. (2011). *Experimental change from paper-based interviewing to computer-assisted interviewing in the HILDA survey (*Melbourne Institute Working Paper No. 2/11*).* Melbourne Institute of Applied Economic and Social Research, the University of Melbourne. https://melbourneinstitute.unimelb.edu.au/assets/documents/hilda-bibliography/hilda-discussion-papers/hdps211.pdf

Appendix A

Domain	Metric
Health Care Utilization	■ Total health care events reported per 100 reference period days ■ Total dental, hearing, medical provider (MP), other medical (OM), prescription medicine (PM), and vision events reported per 100 reference period days ■ Proportion of interviews with zero dental, hearing, MP, OM, PM, and vision events reported ■ Types of procedures reported for dental, hearing, and vision events ■ Proportion of MP events that are repeat visits ■ Proportion of new PM events for which respondent had PM bottles/documentation on hand ■ Proportion of new PMs found in the PM lookup tool ■ Proportion of PMs entered in the utilization sections vs. cost series of the questionnaire ■ Item-level nonresponse: PM details regarding quantity of medicine obtained, prescribed per day, taken per day ■ Item-level nonresponse: OM event details regarding specific types, rental status
Health Care Costs	■ Total health care costs reported per 100 reference period days ■ Total statement costs/non-statement costs reported per 100 reference period days ■ Proportion of interviews with zero costs reported ■ Proportion of interviews with zero statement costs/non-statement costs ■ Proportion of statement costs linked to multiple, matched statements ■ Average non-statement charge bundle amounts ■ Precision of payment amounts reported ■ Item-level nonresponse: Statement items, such as total amount charged, Medicare approved amount, maximum may be billed

Appendix A: Community outcome metrics included in the mode analysis

Appendix B

Domain	Metric
Health Care Utilization	■ Receipt and frequency of medical care inside and outside of facility, dental care, mental health care, physical therapy care. educational and habitational services, and emergency/hospital care ■ Receipt of supplies, medical devices, equipment ■ Item-level nonresponse: Receipt and frequency of medical care inside and outside of facility, dental care, mental health care, physical therapy care, educational and habitational services, and emergency/hospital care ■ Item-level nonresponse: Receipt of supplies, medical devices, equipment
Health Care Costs	■ Ancillary charge details ■ Billed amounts ■ Proportion of cases requiring follow-up collection of billing period/amounts in a later round ■ Precision of payment amounts reported ■ Item-level nonresponse: Ancillary charge details ■ Item-level nonresponse: Billed amounts

Appendix B: Facility outcome metrics included in the mode analysis

Appendix C

Analysis	N sizes
Winter 2020 Community	Winter 2020: 1,877 Winter 2019: 3,302 Winter 2018: 3,256 Winter 2017: 3,122 Winter 2016: 3,468
Winter 2020 Facility	Winter 2020: 396 Winter 2019: 414 Winter 2018: 346 Winter 2017: 337 Winter 2016: 864
Summer 2020 Community	Summer 2020: 9,051 Summer 2019: 8,918 Summer 2018: 8,443 Summer 2017: 10,307 Summer 2016: 7,545
Summer 2020 Facility	Summer 2020: 759 Summer 2019: 811 Summer 2018: 768 Summer 2017: 810 Summer 2016: 747
Fall 2020 Community	Fall 2020: 7,936 Fall 2019: 8,394 Fall 2018: 7,293 Fall 2017: 7,216 Fall 2016: 6,505
Fall 2020 Facility	Fall 2020: 639 Fall 2019: 763 Fall 2018: 716 Fall 2017: 740 Fall 2016: 714

Appendix C: N sizes for completed Community and Facility interviews included in each analysis