Assessing the Utility of 2015 Medicare Advantage Encounter Data to Improve MCBS Estimates

Nicholas Davis, Michael Trierweiler, Holly Hagerty

NORC at the University of Chicago, 55 E. Monroe Street, Chicago, IL 60603

Abstract

Medicare Advantage (MA) encounter data, first available to researchers in April 2018, have information about services provided to beneficiaries enrolled in MA for calendar year 2015. These data have the potential to improve the accuracy of estimates derived from the Medicare Current Beneficiary Survey (MCBS). MCBS data on the usage and cost of Medicare covered services are matched with administrative data from traditional Fee-For-Service (FFS) claims in order to fill in and correct survey-reported payment amounts with more accurate FFS billing information. In 2015, however, 33% of MCBS respondents were enrolled in MA plans. The result is a substantial portion of healthcare utilization by MCBS respondents without FFS claims. MA encounter data could possibly be used to supplement the claims match for MCBS beneficiaries enrolled in MA plans, which would increase the precision of survey-reported information for MA enrollees. We assess the quality and representativeness of the MA encounter data, examining their internal consistency as well as their consistency with data collected from the MCBS in 2015, in order to assess the viability of matching survey events to MA encounter data.

Key Words: Medicare, Medicare Advantage, Matching, Medicare Claims Data, Data Validation, Imputation

1. Introduction

Approximately 61 million Americans are enrolled in Medicare, the federal health insurance program for individuals 65 years of age and older and for individuals younger than 65 years with a qualifying disability, such as end stage renal disease. As baby boomers continue to age into eligibility for Medicare over the next ten years, enrollment and national Medicare spending are predicted to rise (CMS Office of the Actuary, 2019).

Medicare benefits can be obtained in two main ways. Approximately two-thirds of beneficiaries are enrolled in traditional or Fee-For-Service (FFS) Medicare, wherein Medicare pays for benefits at the time of service for enrolled beneficiaries (Medicare Enrollment Dashboard, 2019). Traditional Medicare covers Medicare Part A (hospital insurance) and Part B (medical insurance) benefits.

Over a third of Medicare beneficiaries receive Medicare benefits through a Medicare Advantage (MA) plan, also known as Medicare Part C. MA plans, including health maintenance organization (HMO) and preferred provider organization (PPO) plans, are administered by private insurance companies, or Medicare Advantage Organizations (MAOs), which have been approved by Medicare. Medicare pays MAOs a fixed amount per beneficiary per month to cover their Medicare benefits. MA plans typically include prescription drug coverage in addition to Medicare Part A and Part B benefits, and can often include other benefits such as medical help hotlines, dental and vision care, and non-emergency transportation to doctor's offices. They are

sometimes described as a "one stop" alternative to traditional Medicare (Centers for Medicare and Medicaid Services, 2018).

Enrollment in MA plans has nearly doubled since 2009, with over 22 million beneficiaries enrolled in MA plans as of May 2019 (Kaiser Family Foundation, 2019). In addition, the Congressional Budget Office (CBO) projects that government spending on MA plans will more than double by 2028, at which time an estimated 42% of Medicare enrollees will be covered by an MA plan (Congressional Budget Office, 2019).

The Medicare Current Beneficiary Survey (MCBS) is one of the most important sources of data on Medicare beneficiaries. The MCBS is a continuous, multipurpose 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. MCBS beneficiaries are interviewed at three points per year for four consecutive years.

Survey data from the MCBS cover a broad range of topics, including health status of the beneficiary, household characteristics, and knowledge of the Medicare system. Additionally, a large amount of data is collected on each beneficiary's interactions with the healthcare system in the form of access to medical care, utilization of care, and cost of all drugs and services obtained by a beneficiary.

Survey-reported data on healthcare costs and utilization are characterized by a significant amount of missingness. To fill in some of the gaps, survey-reported data on the usage and cost of Medicare covered services are matched with administrative data from traditional FFS claims in order to correct survey-reported payment amounts with more accurate FFS billing information. This process results in more accurate estimates of healthcare costs for MCBS beneficiaries who are enrolled in traditional FFS Medicare, and thus less imputation for missing data (MCBS 2015 Methodology Report).

This matching process, however, is not available for MCBS beneficiaries who are enrolled in MA plans as they do not have FFS claims. MA encounter data, first made available to researchers in 2018, represent a potential opportunity to supplement survey-reported data from MA enrollees in a manner similar to the claims match process for FFS enrollees. This could increase the overall precision of estimates derived from MCBS data for MA enrollees.

2. Assessment of Encounter Data

MAOs submit encounter data to CMS in order to characterize a beneficiary's unique interactions, or "encounters," with healthcare providers. MA encounter data are roughly analogous in structure to FFS claims data (Medicare Advantage Encounter Data User's Guide, 2019). In this paper, we outline our approach to assessing the viability of matching survey events to MA encounter data in a way that would approximate the FFS claims match process.

Encounter data files are available for six settings: practitioner/provider services (Carrier), durable medical equipment (DME), home health (HHA), inpatient (IP), institutional outpatient (OP), and skilled nursing facility (SNF) (Medicare Advantage Encounter Data User's Guide, 2019). Data files for each setting are partitioned into base and line-level records. Base records are associated with the header portion of a claim, with line-level records comprising associated line items that fall under the claim.

For this analysis, we had access to a subset of MA encounter data from 2015, for MCBS beneficiaries from the 2012-2017 panels. We also received unique access to information on payments and costs, which are not typically publicly available due to the proprietary nature of the data. Since payments are not made at the time of service for MA-covered services, the exact meaning of payment and cost variables within the MA encounter data can be ambiguous (Medicare Advantage Encounter Data User's Guide, 2019). Payment and cost variables are provided by the plan providers. The data in these fields vary depending upon the MA plan; the fields might contain placeholder data, imputed data, or values that are close to the true costs for an encounter.

Table 1 shows the distribution of claims between the six settings in the raw MA encounter data. Overall, we had access to a starting set of approximately 550,000 claims and their associated line-level information. Most of these claims fall into the Carrier setting, which reflects beneficiary encounters with a broad set of professional providers, including physicians, physician assistants, nurse practitioners, and clinical social workers. This setting also reflects some organizational providers such as independent clinical laboratories (ResDAC Carrier Data Documentation, 2019).

Table 1. Distribution of claims between MA encounter data settings

Setting	Number of records in base file
Provider/Practitioner Services (Carrier)	456,618
Durable Medical Equipment (DME)	21,974
Home Health (HHA)	4,731
Inpatient (IP)	5,464
Institutional Outpatient (OP)	59,069
Skilled Nursing Facility (SNF)	1,714

2.1 De-duplication of MA encounter data

In the MA encounter data files, multiple records can be associated with a single service. Record de-duplication is necessary to identify unique services for beneficiaries. Different methods of de-duplication can be used depending on the research goal; therefore, it is necessary to be thoughtful about the type of de-duplication used (Medicare Advantage Encounter Data User's Guide, 2019). We conduct de-duplication within the base files of each of the six settings.

Chart review records present a unique challenge when dealing with encounter data. These records allow providers to add or remove diagnoses codes to an existing "parent" claim. Chart review records are linked to a "parent" claim—the original claim which the chart review record is intended to modify or substantiate with additional context—by a claim control number. However, some chart review records are missing this claim control number. We call these "independent" chart review records, as we do not have a record of the original claim the chart review record is meant to modify. Conversely, we define "dependent" chart review records when an original parent claim can be identified. We remove these dependent chart review records during de-duplication, as our purpose is not to evaluate the specific diagnoses or procedures conducted, but to evaluate the costs associated with the medical encounter. We retain the independent chart review records as they are our only source of information about the given encounter. However, we do remove independent chart review records if they have an identical beneficiary ID value, claim start date, claim end date, National Provider Identifier (NPI) number, and type of bill value, when considered against other independent chart review records.

MAOs submit encounter records iteratively, such that records can be submitted and then later voided or replaced by the submission of another record. These voided and replaced records need to be eliminated during de-duplication, along with the dependent chart review records. To keep these records would be to over-count single encounters. Retained chart review records as a percentage of the final de-duplicated base files range from 0% (Home Health) to 11.1% (Inpatient). Within Home Health, there were a very low number of chart review records to start out with (41), none of which were retained using these rules.

We de-duplicate the remaining data files by beneficiary ID, claim start date, claim end date, NPI number, and type of bill value—the same as the independent chart review records—along with all diagnosis code values.

A small set of records with default NPI values of 1999999976, 1999999984, or 1999999992 are de-duplicated on the same variable set minus NPI number.

After de-duplication, the data files primarily consist of claim records that identify unique services. Table 2 illustrates that approximately 80-90% of records are retained among base files for each of the six settings. The Inpatient setting requires the most de-duplication.

Table 2. Percentage of original records that are kept after de-duplication by setting

Carrier	DME	HHA	IP	OP	SNF
82%	84%	87%	77%	88%	86%

3. Internal Quality Metrics

It is important to assess the internal quality of a new data source such as the MA encounter data to identify unexpected patterns or incorrect values that may limit their usefulness. To that end, we developed internal consistency checks for the correspondence of claim-level variables on the base files and the sum or range of the associated individual items on the line-level files. The main categories of variables evaluated for internal consistency concern claim dates, costs, and payments.

3.1 Validity of date variables

Claims on claim-level files have start and end dates. Line-level records associated with these claims have associated dates. We checked whether the range of dates covered by line-level items associated with a claim corresponded with the start and end date from the claim-level file. No significant discrepancies were identified. Small deviations can be permissible—a line-level record can be submitted with a line-level end date after the claim-level end date. This is rare. For example, there are 29 instances of claim start and end dates occurring on the same day in the base file (which is often the case for encounters such as physician visits that take place on a single day), and line-level expenses being submitted 1-9 days later. This could be due to the amount of time it takes for invoices to be processed. We do not consider it a threat to internal data quality.

We also checked for pre-defined measures of data quality related to date variables, such as claim end dates occurring before claim start dates, which would be logically impossible. No such cases were found. Claim dates were verified for consistency with the expected year (2015).

We also examined the length of claim by setting, in number of days. Table 3 contains these data. The number of records in Table 3 are limited to members of the 2015 MCBS Survey File. For the Carrier, Durable Medical Equipment, and Outpatient settings, it is most common for encounters to occur within the space of a day, although there are outlying observations with long claim lengths. It makes intuitive sense that claims related to Inpatient stays, Home Health encounters, or stays in Skilled Nursing Facilities would be associated with a longer median claim length.

Table 3. Length of claim by setting

Base file	N	Min	Mean	Median	Max
Carrier	107,237	1	1.3	1	364
DME	5,815	1	3.6	1	124
HHA	1,473	1	15.1	7	181
IP	1,235	1	5.8	4	65
OP	15,455	1	2.3	1	86
SNF	506	1	12.3	10	78

3.2 Validity of cost and payment variables

Consistency between claim- and line-level data on costs and payments is also important. We expect the sum of line-level payments to correspond to the claim-level total payment value. Likewise, we expect the sum of line-level charges to correspond to the claim-level total charge value. For most records, we find that this expectation is met. Table 4 shows the percentage of claims by setting where the difference between claim- and line-level values is less than $1.00, for both payment and charge related variables. These differences are less than $1.00 for 95-99% of claims overall.

Table 4. Consistency of claim- and line- level costs and payments across setting

Measure	Carrier	DME	HHA	IP	OP	SNF	All
Absolute payment difference < $1	95%	96%	88%	94%	96%	94%	95%
Absolute charge difference < $1	99%	99%	85%	96%	96%	99%	99%

4. External Comparisons

In addition to verifying the internal consistency of variables related to claim dates, costs, and payments, we investigated whether the information on healthcare costs found in the MA data is consistent with benchmarks from the MCBS population. To understand whether we could use MA encounter data in an analog claims match process, we also examined the structure of an MA encounter record relative to a FFS claim.

4.1 MA encounter data versus MCBS survey-reported data

For this comparison, we measured the consistency of MA encounter data with survey-reported data on medical costs and events from the MCBS. To do this, we identified 2,845 MA enrollees present in both the 2015 MA encounter data and the MCBS 2015 Cost Supplement (CS) ever-enrolled population.

The MCBS 2015 Cost Supplement ever-enrolled population is comprised of MCBS sample members who represent the population of beneficiaries enrolled at any time in Medicare in 2015 (Medicare Current Beneficiary Survey Cost Supplement 2015 Data User's Guide). We linked the Cost Supplement ever-enrolled population to the MA encounter data using a unique Medicare beneficiary identifier.

We identified members of the Cost Supplement ever-enrolled population who were also MA enrollees. We defined MA enrollee status as at least one month of enrollment in an MA plan, per administrative data that classifies monthly enrollment status.

Table 5 contains summary information for these same 2,845 beneficiaries in both the MA encounter data and the MCBS survey-reported data.

Table 5. Record volume for MA enrollees in MCBS data versus MA encounter data

	MCBS Data	MA Data
Number of MA enrollees	2,845	2,845
Number of MA enrollees with MA payments	1,243	2,600
Percentage of MA enrollees with MA payments	43.7	91.4
Number of MA enrollee records	51,103	82,605
Number of MA enrollee records with MA payments	10,585	62,003
Percentage of MA enrollee records with MA payments	20.7	75.1

Several insights emerge from analysis of these beneficiaries and their associated data. For the same 2,845 individuals, the MA encounter data has more beneficiaries with MA payments (91% versus 44%). The MA encounter data contain more records overall for this group (approximately 83,000 versus 51,000 records). In the MA data, three quarters of beneficiary records have an MA payment present. In comparison, in MCBS survey-reported data for these same beneficiaries, just over one fifth of records are associated with a MA payment. Given that MA encounter data should reflect all of a beneficiary's MA-covered services, it is reasonable that the MA encounter data would have a higher volume of data on MA records and payments for the same beneficiaries, relative to the MCBS survey-reported data.

Despite the larger volume of records that exists in the MA encounter data, Table 6 shows that there is consistency between the MA data and the MCBS survey-reported data using measures of central tendency for cost and payment related variables. For claim total charge and claim payment amount, the median value from the MA data is within ten percent of the median value of MCBS survey-reported data. Medians are most useful in this instance as means are likely to be influenced by large outliers in both data sources. This table reflects cases with positive claim payment amounts.

Table 6. Comparison of select variables for MCBS data versus MA encounter data

	MCBS Data	MA Data	Ratio (MA / MCBS)
Mean Claim Total Charge	$896	$1,209	1.35
Mean Claim Payment Amount	464	299	.64
Mean Beneficiary Payment	22	12	.55
Median Claim Total Charge	189	204	1.08
Median Claim Payment Amount	82	79	.96
Median Beneficiary Payment	0	0	-

4.2 MA encounter records versus FFS claims

We also investigated whether MA encounter records are similar to FFS claims in terms of their structure and volume per beneficiary. To that end, we looked at the body of MA encounter records versus FFS claims associated with members of the 2015 Cost Supplement ever-enrolled population. As Table 7 shows, FFS claims are associated with a higher proportion of records with payments—almost 99% of FFS records are associated with any payment, compared to 75% of MA encounter records. Within FFS claims data, there are typically more records per beneficiary than in MA encounter data.

Table 7. Record volume for the MCBS 2015 Cost Supplement Sample versus MA encounter data

	FFS Claims Data	MA Data
Number of beneficiaries with claims data	6,582	2,936
Number of beneficiaries with payment amounts > 0	6,214	2,674
Percentage of beneficiaries with payments	94.4	91.1
Number of records	242,630	84,276
Number of records with payments	239,091	63,364
Percentage of records with payments	98.5	75.2
Number of records per beneficiary	37	29
Number of records with payments per beneficiary with payments	38	24

In Table 8, we see that the median claim charge and payment values are similar when comparing FFS claims data to MA encounter data. This table includes cases with positive claim payment amounts.

Table 8. Comparison of select variables for FFS claims data versus MA encounter data

	FFS Claims Data	MA Data	Ratio (MA / MCBS)
Mean Claim Total Charge	$934	$1,220	1.31
Mean Claim Payment Amount	303	305	1.01
Mean Beneficiary Payment	3	12	3.55
Median Claim Total Charge	160	204	1.28
Median Claim Payment Amount	78	79	1.02
Median Beneficiary Payment	0	0	-

5. Discussion

The findings of this paper illustrate some of the initial exploratory steps that should be taken when evaluating the utility of a new auxiliary data source. These exploratory steps include the evaluation of internal consistency measures and verification of the data source against existing benchmarks.

We have studied the internal consistency of the MA encounter data, primarily related to variables related to payments, costs, and dates. Most of the MA encounter data seem to be consistent and of good internal quality, although minor payment inconsistencies exist.

We also explored several external comparisons relating to the quality of MA encounter data. When benchmarked against MCBS sample members, MA encounter data are consistent with MCBS survey-reported data from MA enrollees. We found that MA encounter records are similar to FFS claims in terms of payments overall and by setting.

We believe that the MA encounter data would be suitable for use in MCBS data processing and imputation for beneficiaries enrolled in MA plans. If timing enabled direct matching of MA encounter data to MCBS survey-reported data, we could employ probabilistic record linkage to match specific MA encounter records to survey-reported claims.

Under the current structure of processing and release for MA encounter data, direct matching of this type would not be possible during the window of time that MCBS imputation is conducted. Release may become faster with future iterations of MA encounter data. Currently, we will explore other ways in which MA encounter data could benefit the MCBS imputation process for MA enrollees. These include statistical matching of current MA enrollees within the MCBS to donors within MA encounter data from past years who are similar on observable characteristics or have similar medical claims. Furthermore, if further study of MA encounter data continues to confirm a pattern of consistently higher event volume, we may also consider employing a ratio adjustment to scale upward the number of survey-reported events for MA enrollees in the MCBS.

References

"A Dozen Facts About Medicare Advantage in 2019", Kaiser Family Foundation.
https://www.kff.org/medicare/issue-brief/a-dozen-facts-about-medicare-advantage-in-2019/

Centers for Medicare and Medicaid Services. "Understanding Medicare Advantage Plans."
October 2018. https://www.medicare.gov/Pubs/pdf/12026-Understanding-Medicare-Advantage-Plans.pdf

CMS Office of the Actuary Releases 2018-2027 Projections of National Health Expenditures.
February 20, 2019. https://www.cms.gov/newsroom/press-releases/cms-office-actuary-releases-2018-2027-projections-national-health-expenditures

Medicare Enrollment Dashboard. July 2019. https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Dashboard/Medicare-Enrollment/Enrollment%20Dashboard.html

Medicare Current Beneficiary Survey Cost Supplement 2015 Data User's Guide.
https://www.cms.gov/Research-Statistics-Data-and-Systems/Research/MCBS/Codebooks-Items/2015_Cost_Supplement.html?DLPage=1&DLEntries=10&DLSort=0&DLSortDir=descending

Medicare Current Beneficiary Survey 2015 Methodology Report. February 2018.
https://www.cms.gov/Research-Statistics-Data-and-Systems/Research/MCBS/Downloads/MCBS2015MethodReport508.pdf

Medicare Advantage Encounter Data User's Guide. May 2019.
https://www.ccwdata.org/web/guest/user-documentation

Medicare – Congressional Budget Office's May 2019 Baseline. May 2, 2019.
https://www.cbo.gov/system/files?file=2019-05/51302-2019-05-medicare.pdf

Research Data Assistance Center (ResDAC). Carrier (Encounter) Data Documentation.
https://www.resdac.org/cms-data/files/carrier-encounter-final