



Payroll-Based Journal Public Use Files: Technical Specifications

April 2022

1. Background

The Centers for Medicare & Medicaid Services (CMS) has prepared public data sets, the Payroll Based Journal Public Use Files (herein referred to as “PBJ PUFs”) that have daily data on nursing home staffing levels based on data submitted by nursing homes to CMS through the Payroll Based Journal (PBJ) system and daily resident census information based on Minimum Data Set (MDS) records. Long-term care (LTC) facilities must electronically submit to CMS through the PBJ system complete and accurate direct care staffing information, including information for agency and contract staff, based on payroll and other verifiable and auditable data in a uniform format according to specifications established by CMS. Facilities are required to submit the number of hours that each staff member (including agency and contract staff) is paid to deliver services for each day worked.

There are two PBJ PUFs, one for nursing staff (RNs, LPN/LVNs and nurse aides) and one for all other categories of non-nursing staff for which PBJ data submission is mandatory. Both quarterly files are available beginning with data from the first calendar quarter of 2017¹. New data files will be uploaded to data.cms.gov on a quarterly basis. The public use files report information on staffing hours for each day in the quarter, along with resident census information derived from the Minimum Data Set (MDS). Note that the PBJ PUFs include only data that were received by the reporting deadline (which is 45 days after the last day in the quarter) and exclude data from facilities that submitted incomplete or erroneous data (see details below).

More details on the PBJ system, including the PBJ Policy Manual, can be found here:

<https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/NursingHomeQualityInits/Staffing-Data-Submission-PBJ.html>

2. Aggregation

The staffing data in the PBJ PUFs are aggregated to the facility-day. This means that all included facilities will have one record (or row of data) for each day in the quarter. This will range from 90-92 records depending on the number of days in the quarter but for a given quarter will be the same for all included facilities. Thus, the total number of records in the file will be equal to the number of included facilities multiplied by the number of calendar days in the quarter.

3. Facility Inclusion and Exclusion Criteria

All inclusion and exclusion criteria are applied at the level of the facility. No additional editing or exclusions were applied for individual days or staff types. These criteria are based on the nurse staffing submitted; however, if a facility meets the inclusion criteria for the nursing PUF for a given quarter, that facility will also be included in the non-nursing PUF for the same quarter.

a. Facility Inclusion Criteria

In order to be included in the PBJ PUF, nursing homes must have been active on last day of the quarterly submission period.

b. Facility Exclusion Criteria

A nursing home is excluded from the PBJ PUFs if the nursing home had aggregate PBJ nurse staffing levels for the quarter that are considered aberrant, based on rules previously applied for the

¹ The specifications for generating these files changed starting with 2018Q4 (posted in April 2019) and again for 2020Q3 (posted in January 2022). The differences in the specifications for the earlier files are described in the Appendix to this document.

exclusion of Certification and Survey Provider Enhanced Reports (CASPER)-based staffing measures reported on Nursing Home Compare and used in the Five-Star Quality Rating System. Specifically, nursing homes are excluded from the PBJ PUFs if:

- (1) Total nurse staffing = 0 hours ² per resident per day; or
- (2) Total nurse staffing > 12 hours per resident per day; or
- (3) Nurse aide staffing > 5.25 hours per resident per day ; or
- (4) Total nurse staffing on weekends = 0 hours per resident per day³; or
- (5) Total nurse staffing on weekends >12 hours per resident per day³; or
- (6) Total nurse aide staffing on weekends > 5.25 hours per resident per day³

For this purpose, total nurse staffing includes the following specific job categories: Registered Nurse (RN) Director of Nursing (DON), RN with administrative duties, RN, Licensed Practical Nurse (LPN) with administrative duties, LPN, Certified Nursing Assistant (CNA), Medication Aide/Technician, and Nurse Aide in Training. Aggregate staffing for the quarter is calculated by summing all relevant staffing hours across the quarter and dividing this by the sum of the daily MDS census across the quarter. For weekend staffing, both staffing hours and census are aggregated across all weekend days (Saturdays and Sundays) in the quarter. These aggregates include only days for which there was a non-zero census. Note also that if a nursing home is excluded due to aberrant weekend staffing (i.e., exclusions 4, 5, or 6 above), no data for this provider will be included in the PUFs, regardless of staffing levels on other days in the quarter. In other words, the PUFs will include all days in the quarter for a nursing home or the nursing home will be excluded entirely.

Note that the exclusions for aberrant nurse staffing levels are applied to facilities' data aggregated across all days (or all weekend days for exclusions 4, 5, and 6) in the entire quarter, not for individual days. As a result, some extremely high and low daily values exist in the data, and in particular, none of the inclusion or exclusion criteria require that facilities reported RN or LPN hours. Analysts may wish to apply additional edits to the data and should be cautious in the interpretation of extreme values. It should also be noted that facilities are instructed to report only paid hours; yet, some staff, particularly salaried workers, may work additional hours that are unpaid and unreported.

4. Data Contents

Table 1 shows the variables (columns) that are included in the PBJ nursing PUF, and Table 2 shows the variables (columns) that are included in the non-nursing PUF. For additional description of the job categories, see the PBJ Policy Manual (URL in Section 1). For additional methodological detail regarding the MDS census, see Section 5.

The file contains no missing data (i.e. no null or blank values). If there are no hours reported for a given job category for a given day, the hours variable for this job category will have a value of 0 in the PUF. This will be the case if there were no hours submitted by the quarterly reporting deadline for that facility for that job category for that day.

Additionally, as noted above, beyond the facility-level inclusion and exclusion criteria, no additional editing or exclusions were applied for individual days or hours. Thus, some erroneously high (or low) values may remain in the data.

² Beginning with 2021Q3 data (posted in January 2022), the low staffing exclusion criteria was changed from 'Total Nurse Staffing < 1.5 hours per resident day' to 'Total Nurse Staffing = 0 hours per resident day'.

³ The exclusion criteria related to weekend staffing were added beginning with the 2021Q3 data (posted in January 2022).

Table 1. Columns included in the PBJ Nursing Public Use File

| Column Name | Description | Specification notes |
|--------------------|---|----------------------------|
| PROVNUM | Medicare provider number | 6 characters |
| PROVNAME | Provider name | character |
| CITY | Provider City | character |
| STATE | Postal abbreviation for State | 2 characters |
| COUNTY_NAME | Name of Provider County, unique within state | character |
| COUNTY_FIPS | FIPS Code for Provider County, unique within state | 3 digits |
| CY_Qtr | Calendar Quarter (yyyyQq, e.g. 2018Q4) | 6 characters |
| WorkDate | Day for Reported Hours (yyyymmdd) | 8 characters |
| MDS Census | Resident Census from MDS | integer |
| Hrs_RNDON | Total Hours for RN Director of Nursing (PBJ job code 5) | numeric |
| Hrs_RNDON_emp | Employee Hours for RN Director of Nursing (5) | numeric |
| Hrs_RNDON_ctr | Contract Hours for RN Director of Nursing (5) | numeric |
| Hrs_RNadmin | Hours for RN with administrative duties (6) | numeric |
| Hrs_RNadmin_emp | Employee Hours for RN with administrative duties (6) | numeric |
| Hrs_RNadmin_ctr | Contract Hours for RN with administrative duties (6) | numeric |
| Hrs_RN | Total Hours for RN (7) | numeric |
| Hrs_RN_emp | Employee Hours for RN (7) | numeric |
| Hrs_RN_ctr | Contract Hours for RN (7) | numeric |
| Hrs_LPNadmin | Total Hours for LPN w/ admin duties (8) | numeric |
| Hrs_LPNadmin_emp | Employee Hours for LPN w/ admin duties (8) | numeric |
| Hrs_LPNadmin_ctr | Contract Hours for LPN w/ admin duties (8) | numeric |
| Hrs_LPN | Total Hours for LPN (9) | numeric |
| Hrs_LPN_emp | Employee Hours for LPN (9) | numeric |
| Hrs_LPN_ctr | Contract Hours for LPN (9) | numeric |
| Hrs_CNA | Total Hours for CNA (10) | numeric |
| Hrs_CNA_emp | Employee Hours for CNA (10) | numeric |
| Hrs_CNA_ctr | Contract Hours for CNA (10) | numeric |
| Hrs_NAtrn | Total Hours for Nurse aide in training (11) | numeric |
| Hrs_NAtrn_emp | Employee Hours for Nurse aide in training (11) | numeric |
| Hrs_NAtrn_ctr | Contract Hours for Nurse aide in training (11) | numeric |
| Hrs_MedAide | Total Hours for Med Aide/Technician (12) | numeric |

| Column Name | Description | Specification notes |
|-----------------|---|---------------------|
| Hrs_MedAide_emp | Employee Hours for Med Aide/Technician (12) | numeric |
| Hrs_MedAide_ctr | Contract Hours for Med Aide/Technician (12) | numeric |
| Incomplete* | (0/1) Provider has incomplete staffing data | numeric |

Note: More information on labor codes and descriptions can be found in Table 1 of the PBJ Policy Manual, which can be found on the web link listed in section 1 of this document.

* In 2021Q4, the staffing data for nearly 900 nursing homes was affected by a ransomware attack on a large, nationwide payroll provider. For the 2021Q4 PUF only, the variable "incomplete" is used to designate which providers are among the affected nursing homes (incomplete = 1). In the 2021Q4 PUF, these providers have their data truncated and include data only for October 1 – December 10, 2021. Additional information about these changes is in the Appendix.

Table 2. Columns included in the PBJ Non-Nursing Public Use File

| Column Name | Description | Specification notes |
|----------------|--|---------------------|
| PROVNUM | Medicare provider number | 6 characters |
| PROVNAME | Provider Name | character |
| CITY | Provider City | character |
| STATE | Postal abbreviation for State | 2 characters |
| COUNTY_NAME | Name of Provider County, unique within state | character |
| COUNTY_FIPS | FIPS Code for Provider County, unique within state | 3 digits |
| CY_QTR | Calendar quarter (yyyyQq, e.g. 2020Q4) | 6 characters |
| WorkDate | Day for Reported Hours (yyyymmdd) | 8 characters |
| MDS Census | Resident Census from MDS | integer |
| Hrs_Admin | Total Hours for Administrator (PBJ job code 1) | numeric |
| Hrs_Admin_emp | Employee Hours for Administrator (1) | numeric |
| Hrs_Admin_ctr | Contract Hours for Administrator (1) | numeric |
| Hrs_MedDir | Total Hours for Medical Director (2) | numeric |
| Hrs_MedDir_emp | Employee Hours for Medical Director (2) | numeric |
| Hrs_MedDir_ctr | Contract Hours for Medical Director (2) | numeric |
| Hrs_OthMD | Total Hours for Other Physician (3) | numeric |
| Hrs_OthMD_emp | Employee Hours for Other Physician (3) | numeric |
| Hrs_OthMD_ctr | Contract Hours for Other Physician (3) | numeric |
| Hrs_PA | Total Hours for Physician Assistant (4) | numeric |
| Hrs_PA_emp | Employee Hours for Physician Assistant (4) | numeric |
| Hrs_PA_ctr | Contract Hours for Physician Assistant (4) | numeric |

| Column Name | Description | Specification notes |
|---------------------|--|---------------------|
| Hrs_NP | Total Hours for Nurse Practitioner (13) | numeric |
| Hrs_NP_emp | Employee Hours for Nurse Practitioner (13) | numeric |
| Hrs_NP_ctr | Contract Hours for Nurse Practitioner (13) | numeric |
| Hrs_ClinNrsSpec | Total Hours for Clinical Nurse Specialist (14) | numeric |
| Hrs_ClinNrsSpec_emp | Employee Hours for Clinical Nurse Specialist (14) | numeric |
| Hrs_ClinNrsSpec_ctr | Contract Hours for Clinical Nurse Specialist (14) | numeric |
| Hrs_Pharmacist | Total Hours for Pharmacist (15) | numeric |
| Hrs_Pharmacist_emp | Employee Hours for Pharmacist (15) | numeric |
| Hrs_Pharmacist_ctr | Contract Hours for Pharmacist (15) | numeric |
| Hrs_Dietician | Total Hours for Dietician (16) | numeric |
| Hrs_Dietician_emp | Employee Hours for Dietician (16) | numeric |
| Hrs_Dietician_ctr | Contract Hours for Dietician (16) | numeric |
| Hrs_FeedAsst | Total Hours for Feeding Assistant (17) | numeric |
| Hrs_FeedAsst_emp | Employee Hours for Feeding Assistant (17) | numeric |
| Hrs_FeedAsst_ctr | Contract Hours for Feeding Assistant (17) | numeric |
| Hrs_OT | Total Hours for Occupational Therapist (18) | numeric |
| Hrs_OT_emp | Employee Hours for Occupational Therapist (18) | numeric |
| Hrs_OT_ctr | Contract Hours for Occupational Therapist (18) | numeric |
| Hrs_Otasst | Total Hours for Occupational Therapy Assistant (19) | numeric |
| Hrs_Otasst_emp | Employee Hours for Occupational Therapy Assistant (19) | numeric |
| Hrs_Otasst_ctr | Contract Hours for Occupational Therapy Assistant (19) | numeric |
| Hrs_Otaide | Total Hours for Occupational Therapy Aide (20) | numeric |
| Hrs_Otaide_emp | Employee Hours for Occupational Therapy Aide (20) | numeric |
| Hrs_Otaide_ctr | Contract Hours for Occupational Therapy Aide (20) | numeric |
| Hrs_PT | Total Hours for Physical Therapist (21) | numeric |
| Hrs_PT_emp | Employee Hours for Physical Therapist (21) | numeric |
| Hrs_PT_ctr | Contract Hours for Physical Therapist (21) | numeric |
| Hrs_Ptasst | Total Hours for Physical Therapy Assistant (22) | numeric |
| Hrs_Ptasst_emp | Employee Hours for Physical Therapy Assistant (22) | numeric |
| Hrs_Ptasst_ctr | Contract Hours for Physical Therapy Assistant (22) | numeric |
| Hrs_Ptaide | Total Hours for Physical Therapy Aide (23) | numeric |
| Hrs_Ptaide_emp | Employee Hours for Physical Therapy Aide (23) | numeric |

| Column Name | Description | Specification notes |
|----------------------|---|---------------------|
| Hrs_Ptaide_ctr | Contract Hours for Physical Therapy Aide (23) | numeric |
| Hrs_RespTher | Total Hours for Respiratory Therapist (24) | numeric |
| Hrs_RespTher_emp | Employee Hours for Respiratory Therapist (24) | numeric |
| Hrs_RespTher_ctr | Contract Hours for Respiratory Therapist (24) | numeric |
| Hrs_RespTech | Total Hours for Respiratory Therapy Technician (25) | numeric |
| Hrs_RespTech_emp | Employee Hours for Respiratory Therapy Technician (25) | numeric |
| Hrs_RespTech_ctr | Contract Hours for Respiratory Therapy Technician (25) | numeric |
| Hrs_SpcLangPath | Total Hours for Speech/Language Pathologist (26) | numeric |
| Hrs_SpcLangPath_emp | Employee Hours for Speech/Language Pathologist (26) | numeric |
| Hrs_SpcLangPath_ctr | Contract Hours for Speech/Language Pathologist (26) | numeric |
| Hrs_TherRecSpec | Total Hours for Therapeutic Recreation Specialist (27) | numeric |
| Hrs_TherRecSpec_emp | Employee Hours for Therapeutic Recreation Specialist (27) | numeric |
| Hrs_TherRecSpec_ctr | Contract Hours for Therapeutic Recreation Specialist (27) | numeric |
| Hrs_QualActvProf | Total Hours for Qualified Activities Professional (28) | numeric |
| Hrs_QualActvProf_emp | Employee Hours for Qualified Activities Professional (28) | numeric |
| Hrs_QualActvProf_ctr | Contract Hours for Qualified Activities Professional (28) | numeric |
| Hrs_OthActv | Total Hours for Other Activities Staff (29) | numeric |
| Hrs_OthActv_emp | Employee Hours for Other Activities Staff (29) | numeric |
| Hrs_OthActv_ctr | Contract Hours for Other Activities Staff (29) | numeric |
| Hrs_QualSocWrk | Total Hours for Qualified Social Worker (30) | numeric |
| Hrs_QualSocWrk_emp | Employee Hours for Qualified Social Worker (30) | numeric |
| Hrs_QualSocWrk_ctr | Contract Hours for Qualified Social Worker (30) | numeric |
| Hrs_OthSocWrk | Total Hours for Other Social Worker (31) | numeric |
| Hrs_OthSocWrk_emp | Employee Hours for Other Social Worker (31) | numeric |
| Hrs_OthSocWrk_ctr | Contract Hours for Other Social Worker (31) | numeric |
| Hrs_MHSvc | Total Hours for Mental Health Service Worker (34) | numeric |
| Hrs_MHSvc_emp | Employee Hours for Mental Health Service Worker (34) | numeric |
| Hrs_MHSvc_ctr | Contract Hours for Mental Health Service Worker (34) | numeric |
| Incomplete* | (0/1) Provider has incomplete staffing data | numeric |

* In 2021Q4, the staffing data for nearly 900 nursing homes were affected by a ransomware attack on a large, nationwide payroll provider. For the 2021Q4 PUF only, the variable "incomplete" is used to designate which providers are among the affected nursing homes (incomplete = 1). In the 2021Q4 PUF,

these providers have their data truncated and include data only for October 1 – December 10, 2021. Additional information about these changes is in the Appendix.

5. MDS-Based Resident Census

The method used to calculate the daily resident census using MDS data is as follows:

- 1) Identify the reporting period (quarter) for which the census will be calculated (e.g., CY 2021 Q1: January 1 – March 31, 2021).
- 2) Extract MDS assessment data for all residents of a facility beginning one year prior to the reporting period to identify all residents that *may* reside in the facility (i.e., any resident with an MDS assessment may still reside in the facility). For example, for the CY 2021 Q1 reporting period, extract MDS data from January 1, 2020 through March 31, 2021.
- 3) Identify discharged residents using the following criteria:
 - a) If a resident has a MDS discharge assessment, use the discharge date reported on that assessment and assume that the resident no longer resides in the facility as of the date of discharge on the last assessment. If there is a subsequent admission assessment, then assume that the resident re-entered the nursing home on the entry date indicated on the admission assessment.
 - b) For any resident with an interval of 150 days or more with NO assessments, assume the resident no longer resides in the facility as of the 150th day from the last assessment. (This assumption is based on the requirement for facilities to complete MDS assessments on all residents at least quarterly.) If no assessment is present, assume the resident was discharged, but the facility did not transmit a discharge assessment.
- 4) For any particular date, residents whose assessments do not meet the criteria in #3 above prior to that date are assumed to reside in the facility. The count of these residents is the census for that particular day.

6. Data Limitations and Caveats

These Public Use Files are being made available to facilitate analysis and exploration of this richly detailed source of staffing information about US Nursing Homes. However, there are some limitations that are worth noting:

- There may be some instances where data submitted by a facility may not be a full representation of the hours staff actually worked. There may be some erroneous reporting as some providers continue to fine-tune their data submissions.
- The staffing measures previously displayed on the Nursing Home Compare website were derived from data reported on the CMS-671 and 672 forms. We do not expect data from these forms to exactly match the PBJ data for several reasons. (Note: Staffing data is no longer being collected on the CMS-671 form as of June 1, 2018).
 - The staffing data submitted on the CMS-671 captures only a two-week period prior to the facility's annual survey, whereas the PBJ data captures the staffing of a facility for each day in a quarter.
 - The definitions of certain job categories have been altered for PBJ data as compared to the instructions on the CMS-671 form. For example, when using data from the CMS-671 form, the total hours for registered nurses (RN) with administrative duties includes licensed practical/vocational nurses with administrative duties. However, the PBJ data have a separate category for licensed practical/vocational nurses with administrative duties.

- Staffing measures based on the data from the CMS-671 form used the census for a single date submitted by the facility on the CMS-672 form. Staffing measures based on PBJ data include the census for each day in a quarter based on MDS submissions.
- The PBJ system requires facilities to submit only those hours that staff are paid to work (in order for the data to be auditable). Facilities may have submitted unpaid hours on the CMS-671 form.

The PBJ data cannot be used to indicate noncompliance with other requirements for LTC facilities. For example, 42 CFR §483.35(b) requires facilities to use the services of a registered nurse for at least 8 consecutive hours a day, 7 days a week. Due to the factors stated above, it is possible a registered nurse worked some additional hours that were unpaid and therefore those hours were not reported in PBJ.

Similarly, we note that the number of hours reported for physicians and non-physician practitioners (e.g., nurse practitioners, physician assistants, and clinical nurse specialists) is low because the services they provide in a facility are typically paid by Medicare, Medicaid, or other medical insurance, and not the facility. For example, the services physicians provide through evaluating residents every 60 days are typically billed to Medicare. Therefore, the hours spent providing these services are not reportable.

7. Limitations of Using Microsoft Excel® to open a large comma-separated value (CSV) file

One of the file formats in which the PBJ PUFs can be downloaded from the data.cms.gov website is CSV. On many computers, if a user double-clicks on the filename for a CSV file, Microsoft Excel® will attempt to open the file. However, Excel cannot correctly load more than 1,048,576 rows. Each PBJ PUF has more than 1.2 million rows. Users may follow the Help provided by Excel or can use a statistical software package such as SAS® or Stata® to directly import the files. An alternative is to download the data for individual states or use filters to select a subset of states for download.

Appendix – Changes in PBJ Public Use File Specifications

When CMS makes changes to the staffing calculations on Nursing Home Care Compare, the same changes are applied to the PBJ PUFs. This section describes the changes that have been made since the first PUFs were posted for 2017Q1.

A. Changes made that apply only to the PBJ PUFs released in April 2022 (2021Q4 data)

In 2021Q4, the staffing data for nearly 900 nursing homes was affected by a ransomware attack on a large, nationwide payroll provider. In response to the ransomware attack, the exclusion rules were applied differently for the affected providers. These providers can be identified through a variable that is present only in the 2021Q4 PUFs (incomplete = 1). For these providers,

- days 0 nursing hours reported were not included in calculations of hours per resident per day when determining whether to include the provider in the PUFs, and
- data from December 11 through December 31, 2021 were completely excluded from the calculations and from the PUFs.

B. Changes made starting with PBJ PUFs released in January 2022

For the PBJ PUFs released starting in January 2022 (data covering CY 2021 Q3), the exclusion criteria changed. These changes coincide with changes made to the exclusion rules for reporting nurse staffing levels on Nursing Home Care Compare and adding weekend nurse staffing measures to the website in January 2022.

- Providers submitting nursing data indicating very low nurse staffing (<1.5 hours per resident per day) are no longer excluded. Nursing homes submitting no nursing hours (i.e., 0 hours per resident per day) are still excluded.
- The same exclusion criteria applied to 7-day staffing are now also being applied for weekend staffing. Specifically, the following nursing homes are now excluded from the PUFs:
 - Providers submitting 0 nursing hours on weekends for the entire quarter are excluded, regardless of nursing hours submitted for weekdays.
 - Providers reporting total nurse staffing on weekends >12 hours per resident per day are excluded.
 - Providers reporting nurse aide staffing on weekends > 5.25 hours per resident per day are excluded.

C. Changes made starting with PBJ PUFs released in April 2019

The PBJ PUFs released prior to April 2019 (that is, the files covering the time periods from calendar 2017Q1 – 2018Q3) were constructed using a slightly different methodology. The two changes made at this time were:

- In earlier files, there was an additional exclusion criterion, which was that nursing homes with 5 or more days in the quarter with a non-zero census but no nursing hours (RN, LPN or CNA) were excluded.
- In earlier files, when staffing hours and census were aggregated (summed) across days in the quarter in order to calculate hours per resident per day for the purpose of applying the exclusion rules, only days with non-zero census and non-zero nursing hours were included. For 2018Q4 and later files, all days with non-zero census are included even if zero nursing hours are reported.

Taken together, these two changes allowed more nursing homes to be included in the PUF, while leading to a slight reduction in per resident per day nurse staffing levels. It should be noted that these changes affected the staffing levels ONLY of nursing homes that have any days with no reported nurse staffing, which is a very small number of nursing homes.