

APPENDIX F
RECOMMENDED DATA QUALITY CHECKS

APPENDIX F
RECOMMENDED DATA QUALITY CHECKS

Physician and Other Provider:

Data Element	Expectation	Validity Criteria
Beneficiary ID	Should be valid ID as found in eligibility file. Can use State's ID unless States also accepts SSN.	100% valid.
Beneficiary Name	Should be captured in such a way that makes separating pieces of name easy. There may be some confidentiality issues that make this difficult to obtain. If collectable, expect data to be present and of good quality.	85% present. Lengths should vary and there should be at least some last names > 8 digits and some first names > 8 digits. This will validate that fields have not been truncated. Also verify that a high percentage have at least a middle initial.
Beneficiary DOB	Should not be missing and should be a valid date.	< 2% missing or invalid.
Plan ID	Critical Data Element.	100% valid.
Provider ID	Should be an enrolled provider listed in provider enrollment file.	95% valid.
Attending Provider ID	Should be an enrolled provider listed in provider enrollment file (also accept the MD License number if listed in provider enrollment file).	> 85 % match with provider file using either provider ID or MD license number.
Provider Location	Minimal requirement is county with zip code being strongly advised.	>=95% w/valid county. >=95% w/valid zip code if available.
Place of Service	Should be routinely coded, especially for physicians.	>=95% valid for physicians. >=80% valid across all providers.
Specialty Code	Coded mostly on physician and other practitioner, optional on other types of providers.	Expect > 80% non-missing and valid on physician or other applicable provider type claims (e.g., other practitioners).
Principal Diagnosis	Well coded except by ancillary type providers.	> 90% non-missing and valid codes (using ICD-9-CM lookup tables) for practitioner providers (not including transportation, lab and other ancillary providers).

**APPENDIX F
RECOMMENDED DATA QUALITY CHECKS**

Physician and Other Provider:

Data Element	Expectation	Validity Criteria
Other Diagnosis	This is not expected to be coded on all claims even with applicable provider types but should be coded with a fairly high frequency.	90% valid when present (see primary diagnosis for applicable provider types). Expect 60% non-missing.
Date of Service	Dates should be evenly distributed across time.	5-7% per month if looking at a full year of data.
Unit of Service (Quantity)	The number should be routinely coded.	98% non-zero. > 70% should be one if CPT code in range 99200-99215, 99241-99291.
Procedure Code	This is a critical data element and should always be coded.	99% present (not zero, blank, 8 or 9-filled). All should be valid codes in State's formulary. The frequency of codes at the three digit level should be the same as the distribution prior to capitation. That is, there should be a wide range of procedures with the same frequency as previously encountered.
Procedure Code Modifier	This is not in the McDATA list, but is important to pick up to separate out surgical procedures/anesthesia/asst. surgeon. It is not applicable for all procedure codes.	> 20% non-missing. Expect a variety of modifiers both numeric (CPT) and Alpha (HCPCS). The more common codes which should appear with at least a minimal frequency are: 47 (anesthesia), 80 (asst. surgeon).
EPSDT Indicator	If this indicator is turned on, the recipient should be < 21 years of age and the provider is certified to administer EPSDT screens.	95% recipients < 21; 85% providers certified as EPSDT.

**APPENDIX F
RECOMMENDED DATA QUALITY CHECKS**

Hospital (Inpatient and Outpatient):

Data Element	Expectation	Validity Criteria
Header Data Elements		
Beneficiary ID	Should be valid ID as found in eligibility file. Can use State's ID unless States also accepts SSN.	100% valid.
Beneficiary Name	Should be captured in such a way that makes separating pieces of name easy. There may be some confidentiality issues that make this difficult to obtain. If collectable, expect data to be present and of good quality.	85% present. Lengths should vary and there should be at least some last names > 8 digits and some first names > 8 digits. This will validate that fields have not been truncated. Also verify that a high percentage have at least a middle initial.
Beneficiary DOB	Should not be missing and should be a valid date.	< 2% missing or invalid.
Plan ID	Critical Data Element.	100% valid.
Provider ID	Should be an enrolled provider listed in provider enrollment file.	95% valid.
Attending Provider ID	Should be an enrolled provider listed in provider enrollment file (also accept the MD License number if listed in provider enrollment file).	> 85 % match with provider file using either provider ID or MD license number.
Provider Location	Minimal requirement is county with zip code being strongly advised.	>=95% w/valid county. >=95% w/valid zip code if available.
Place of Service	Should be routinely coded, especially for physicians.	>=95% valid for physicians. >=80% valid across all providers.
Date(s) of Service	Dates should be evenly distributed across time.	7-8% per month if looking at full year of data.
Principal Diagnosis	Well coded by all providers.	100% non-missing (zero, blank or 9-filled).
Other Diagnosis	Should be present for most but not all claims.	> 60% non-missing (zero, blank or 9-filled).
Date of Service	Dates should be evenly distributed across time.	5-7 per month if looking at a full year of data.

APPENDIX F
RECOMMENDED DATA QUALITY CHECKS

Hospital (Inpatient and Outpatient):

Data Element	Expectation	Validity Criteria
Patient Status Code	Should be valid codes for inpatient claims with most common code to be "Discharged to Home." For outpatient claims it can be coded as not applicable.	For inpatient claims, expect > 90% "Discharged to Home." Expect 1-5% in all other values (except not applicable or unknown). For outpatient claims, not applicable data elements so no requirement except that 98% are valid (including N/A and Unknown).
Line Item Detail		
Date(s) of Service	Must fall within the dates on the header part of the record.	100% valid.
Unit of Service (Quantity)		
Procedure Code	For IP claims, only needs to be coded for surgical procs or other critical procedures. For OPD Claims, should be required for all except the most specific revenue codes.	IP: >=40% non-missing. OT: 100% non-missing unless revenue code is very specific.
Revenue Code	If facility uses UB92 claim form, should always be present.	100% valid.

**APPENDIX F
RECOMMENDED DATA QUALITY CHECKS**

LTC Facility:

Data Element	Expectation	Validity Criteria
Header Data Elements		
Beneficiary ID	Should be valid ID as found in eligibility file. Can use State's ID unless States also accepts SSN.	100% valid.
Beneficiary Name	Should be captured in such a way that makes separating pieces of name easy. There may be some confidentiality issues that make this difficult to obtain. If collectable, expect data to be present and of good quality.	85% present. Lengths should vary and there should be at least some last names > 8 digits and some first names > 8 digits. This will validate that fields have not been truncated. Also verify that a high percentage have at least a middle initial.
Beneficiary DOB	Should not be missing and should be a valid date.	< 2% missing or invalid.
Plan ID	Critical Data Element.	100% valid.
Provider ID	Should be an enrolled provider listed in provider enrollment file.	95% valid.
Attending Provider ID	Should be an enrolled provider listed in provider enrollment file (also accept the MD License number if listed in provider enrollment file).	> 85 % match with provider file using either provider ID or MD license number. The standard may need to be lowered if it proves difficult to collect this from LTC facilities.
Provider Location	Minimal requirement is county with zip code being strongly advised.	>=95% w/valid county. >=95% w/valid zip code if available.
Place of Service	The main place of service here should be the LTC Facility.	> 80% should be coded as "LTC Facility."
Date(s) of Service	Dates should be evenly distributed across time.	7-8% per month if looking at full year of data.
Principal Diagnosis	Not routinely coded. If coded, it should be accurate. (Not part of McDATA for LTC; included here to be picked up if available).	> 30% coded. > 90% accurate if coded.

**APPENDIX F
RECOMMENDED DATA QUALITY CHECKS**

LTC Facility:

Data Element	Expectation	Validity Criteria
Date of Service	Dates should be evenly distributed across time.	5-7% per month if looking at a full year of data.
Patient Status Code	Should be most commonly coded as "Still a Patient".	Expect > 70% "Still a Patient". Expect 1-15% as "Transferred to Home"; 1-5% coded at "Died" and < 2% all other values (except not applicable or unknown).
Line Item Detail		
Specialty Code	This relates to facility specialty. The following institution specialties should be represented if covered by the state: Nursing Facility (SNF/ICF) ICF/MR Psychiatric (both Long and Short Term)	All facility specialties should be represented with the same frequency level as was present prior to capitation (unless program change other than capitation occurred).
Date(s) of Service	Must fall within the dates on the header part of the record.	100% valid.
Unit of Service (Quantity)	Should represent days of care if claim represents daily care information (as opposed to ancillary only services).	95% > zero. 80% should be equal to length of stay (span of dates); 0-filled if no accommodation charge (or related information).
Procedure Code	This will only be coded if recipient receives a procedure while in the facility. This will be rare unless the LTC offers an acute level of care as well.	>=20% non-missing if provider is licensed to perform acute care in addition to LTC.
Revenue Code	If facility uses UB92 claim form, should always be present.	100% valid if facility uses this claim form. If they use standard LTC filling form, this data element will not be available.

**APPENDIX F
RECOMMENDED DATA QUALITY CHECKS**

Outpatient Drug File:

Data Element	Expectation	Validity Criteria
Beneficiary ID	Should be valid ID as found in eligibility file. Can use State's ID unless States also accepts SSN.	100% valid.
Beneficiary Name	Should be captured in such a way that makes separating pieces of name easy. There may be some confidentiality issues that make this difficult to obtain. If collectable, expect data to be present and of good quality.	85% present. Lengths should vary and there should be at least some last names > 8 digits and some first names > 8 digits. This will validate that fields have not been truncated. Also verify that a high percentage have at least a middle initial.
Beneficiary DOB	Should not be missing and should be a valid date.	< 2% missing or invalid.
Plan ID	Critical Data Element.	100% valid.
Provider ID	Should be an enrolled provider listed in provider enrollment file.	95% valid.
Prescribing Provider ID	Should be an enrolled provider listed in provider enrollment file (also accept the MD License number if listed in provider enrollment file).	> 85% match with provide file using either provide ID or MD license number.
Provider Location	Minimal requirement is county with zip code being strongly advised.	>=95% w/valid county. >=95% w/valid zip code if available.
Date of Service	Dates should be evenly distributed across time.	5-7% per month if looking at a full year of data.

APPENDIX F
RECOMMENDED DATA QUALITY CHECKS

Outpatient Drug File:

Data Element	Expectation	Validity Criteria
Unit of Service (Quantity)	The number should be routinely coded.	<p>98% non-missing. > 50% should be > 1 and values of 30, 60, 100 should be present at fairly high frequency (verifying that the number reported is units not number of prescriptions).</p> <p>If drug code grouping algorithms are available, more sophisticated edits can be added to check for the expected types of units by type of drug/DME dispensed.</p>
National Drug Code	This is a critical data element and should always be coded.	<p>99% present (not zero, blank, 8 or 9-filled).</p> <p>All should be valid codes in State's formulary or NDC lookup table.</p> <p>The Frequency of codes should be the same as the distribution prior to capitation, at least at the therapeutic classification level.</p>

APPENDIX F
RECOMMENDED DATA QUALITY CHECKS

Dental File:

Data Element	Expectation	Validity Criteria
Beneficiary ID	Should be valid ID as found in eligibility file. Can use State's ID unless States also accepts SSN.	100% valid.
Beneficiary Name	Should be captured in such a way that makes separating pieces of name easy. There may be some confidentiality issues that make this difficult to obtain. If collectable, expect data to be present and of good quality.	85% present. Lengths should vary and there should be at least some last names > 8 digits and some first names > 8 digits. This will validate that fields have not been truncated. Also verify that a high percentage have at least a middle initial.
Beneficiary DOB	Should not be missing and should be a valid date.	< 2% missing or invalid.
Plan ID	Critical Data Element.	100% valid.
Provider ID	Should be an enrolled provider listed in provider enrollment file.	95% valid.
Referring Provider ID	Should be an enrolled provider listed in provider enrollment file (also accept the MD License number if listed in provider enrollment file).	> 85% match with provide file using either provide ID or MD license number.
Provider Location	Minimal requirement is county with zip code being strongly advised.	>=95% w/valid county. >=95% w/valid zip code if available.
Place of Service	This may not be coded as consistently as physicians, but should be coded when possible.	>=80% valid.
Specialty Code	This should give information about the type of dentistry. Other types of providers should rarely bill these claims. This is not specifically mentioned in McDATA but would be useful to capture when possible.	Expect > 80% to be a dentistry specialty (Dentist, Orthodontist, Oral Surgeons).
Date of Service	Dates should be evenly distributed across time.	5-7% per month if looking at a full year of data.

APPENDIX F
RECOMMENDED DATA QUALITY CHECKS

Dental File:

Data Element	Expectation	Validity Criteria
Unit of Service (Quantity)	The number should be routinely coded.	98% non-zero. > 70% should be one if procedure code represents "visits."
Procedure Code	This is a critical data element and should always be coded. This is not specifically listed in McDATA but gives more information than you can obtain from other data elements required.	99% present (not zero, blank, 8 or 9-filled). All should be valid codes in State's formulary. The frequency of codes at the three-digit level should be the same as the distribution prior to capitation. That is, there should be a wide range of procedures with the same frequency as previously encountered.
EPSDT Indicator	If this indicator is turned on, the recipient should be < 21 years of age and the provider is certified to administer EPSDT screens.	95% recipients < 21; 85% providers certified as EPSDT.
Dental Quadrant	Valid values are UR, UA, UL, LR, LA, LL.	85% valid among the non-missing values. 40% of claims should report either a Dental Quadrant or a Tooth Number.
Tooth Number	Valid values are A-T, and 1-32.	85% valid among the non-missing values. 40% of claims should report either a Dental Quadrant or a Tooth Number.