APPENDIX A

Evidence Table

Author/Year	Study Design	Demographics	Diagnostic Accuracy/Results	Methodological Comments
Hogl, Kiechl, Willeit, Saletu, Frauscher, Seppi, Muller, Rungger, Gasperi, Wenning, Poewe; 2005	Cross-sectional study of age-sex-stratified random sample; Compared iron-study lab test in patients with RLS to non-RLS subjects	N= 74 (50 to 89 years of age)	1. Study results reveal that free serum iron, transferrin, and ferritin levels were similar between both groups (no statistical difference), soluble transferrin receptor (Str) concentration were different in subjects with and without RLS (1.48 vs 1.34 $p<0.001$); Female gender and high Str levels independently predicted risk of RLS.	1. Small sample size 2. No measures of accuracy used (sensitivity, specificity), Positive predictor values (PPV) or negative predictor values (NPV)
Berger, von Eckardstein, Trenkwalder Rothdach, Junker, Weiland; 2002	Cross-sectional design	N= 365 (65 to 83 years of age); of this number 36 were found to have RLS	1. OR associated with RLS for iron = 3.08 (CI 1.02-9.29); OR associated with RLS for transferrin = 5.68 (CI 1.18-27.26 for transferrin saturation. 2. No association with ferritin and soluble transferring receptor found 3. Researchers found no evidence that iron or ferritin deficiency were a major cause of RLS	1. No measures of accuracy used (sensitivity, specificity), Positive predictor values (PPV) or negative predictor values (NPV) 2. Cross-sectional study

Mizuno, Mihara, Miyaoka, Inagaki, Horiguchi; 2004	Cross-sectional design	N= 20 (10 in RLS group, [mean age 71.3] and in 10 non-RLS group [mean age 70.5]),	1. For RLS group serum Fe 121.4, ferritin level 138.1, transferrin level 243; CSF Fe 1.5, CSF ferritin 4.06, CSF transferrin 2.18. For non-RLS subjects serum Fe 114.8, ferritin level 111.7, transferrin level 243.1; CSF Fe 3.00, CSF ferritin 6.68, CSF transferrin 1.60 2. Serum iron studies similar between both groups, but CSF iron and ferritin levels lower in RLS, and transferring levels higher in RLS subjects compared to non-RLS subjects 3. Correlation between the serum and CSF ferritin levels in the RLS group was r=0.652 (p=0.039), while serum and CSF correlation between non-RLS subject is 0.887 (p=0.002	1. No mention of measures of accuracy (e.g., sensitivity, specificity, PPV, NPV) 2. Small sample size (N=20)
Earley, Connor, Beard, Clardy, Allen; 2005	Cross-sectional design	N= 55 (30 subjects with RLS [15 early onset and 15 late onset]), and 22 age and sex matched controls	1. Study revealed a strong correlation between age of symptom onset and CSF values (r = .64); the earlier the age, the lower the ferritin level. Regression model showed that both gender and RLS subtype had significant effect on the CSF level 2. Night-time CSF ferritin	1. No mention of measures of accuracy (e.g., sensitivity, specificity, PPV, NPV) 2. Small sample size 3. Though correlation was moderate to high, it is felt to be the consequence of sex-based bias in the data

			levels were lower in RLS group compared to controls; Subjects < 45 years of age (early onset) had significantly lower CSF ferritin levels than controls	4. Correlation does not mean causation
Clardy, Earley, Allen, Beard, Connor; 2006	Cross-sectional study	N=39 (25 subjects had RLS-12 were early-onset, while 13 were late-onset; 14 subjects were used as controls). CSF specimens of ferritin and iron studies were obtained concomitantly with serum specimens	1. CSF H and L-ferritin subunits were decreased in early onset of RLS. Also total protein amounts in RLS CSF were normal	1 .Small sample size 2. No mention of measures of accuracy (e.g., sensitivity, specificity, PPV, NPV)
O'Keeffe; 2005	Case study	83 year old male with 2 year history of severe RLS based on IRLSSG	1. Serum ferritin level was 93 mcg/L though Hgb. 12.7 g%, MCV 89%, and transferrin saturation 25%;	1. Case study 2. No measures of accuracy used (sensitivity, specificity), Positive predictor values (PPV) or negative predictor values (NPV)
Davis, Rajput, Rajput, Aul, Eichhorn; 2000	Randomized, double-blind, placebo-controlled trial	28 patients met IRLSG criteria, half received Fe, the other half received placebo.	1. Studied revealed no significant difference for both groups for primary outcome measures: improvement vs. no improvement in quality of sleep over a 2-week period; comparing a pretreatment 2-week baseline to weeks 13 to 14; and secondary outcome measures: comparison of the quality of sleep; the effect of RLS on quality of life ;and the	1. Small sample size 2. The use of VAS as a measurement tool for measuring severity, instead of using validated tools (e.g., IRLSSG rating scale). 3. No mention of measures of accuracy (e.g., sensitivity, specificity, PPV, NPV)

			percentage of nights patients were symptomatic.	
Earley, Heckler, Allen; 2004	Open-label design	10 subjects (mean age 62.4) with RLS using JHRLS scale. Half had moderate scores, the the other half with severe symptoms	1. Though IV therapy improved the mean global RLS symptom severity, total sleep time, hours with RLS symptoms and periodic leg movement in 7 subjects, 3 subjects who were fully treated failed to produce any response. Brain iron concentration increased (though not statistically different from non-responders); serum ferritin levels showed a greater predicted rapid decrease. 60% of subjects showed complete remission of RLS symptoms	1. Limitations of the study include the small sample size as well as the open label design. 2. No control or blinding during treatment phase 3. No mention of measures of accuracy (e.g., sensitivity, specificity, PPV, NPV)
Earley, Heckler, Allen; 2005	Open-label design	3 subjects completed the 2-year study, receiving between 2 to 4 courses of supplemental iron	1. Ferritin levels declined at a rate higher than the predicted value. The study noted that the slower the rate of ferritin decline, the more prolonged the symptom improvement	1. Limitations of the study include the small sample size as well as the open label design. 2. No control or blinding during treatment phase. 3. No mention of measures of accuracy (e.g., sensitivity, specificity, PPV, NPV)
Connor, Wang, Patton, Menzies, Troncoso, Earley; 2004	Cross-sectional study	8 subjects	1. Ferritin, divalent metal transporter 1, ferroportin and transferring receptor were decreased, while transferring levels were increased in RLS subjects	1. Limitation of study was low sample size 2. No mention of measures of accuracy (e.g., sensitivity,

			compared to controls; though the total iron regulatory protein (IRP1 and IRP2) activity were decreased in RLS, total IRP2 protein levels were not decreased.	specificity, PPV, NPV) 3. All subjects were female Patients - may be difficult to generalize to male subjects.
Connor, Boyer, Menzies, Dellinger, Allen, Ondo, Earley; 2003	Cross-sectional study	12 subjects - 7 with RLS, 5 served as controls	1. Though no histopathological abnormalities were noted between both groups, iron staining as well as H-ferritin staining was decreased in RLS subjects; though H-ferritin levels were difficult to detect in the substantia nigra of RLS subjects, L-ferritin staining was strong.	1. Small sample size 2. No mention of measures of accuracy (e.g., sensitivity, specificity, PPV, NPV)
Earley, Connor, Beard, Malecki, Epstein, Allen; 2000	Cross-sectional study	16 subjects with RLS, matched with 8 age-matched controls	1. Subjects with RLS had lower CSF ferritin levels (1.11 vs. 3.50 ng/ml) and higher CSF transferrin levels (26.4 vs. 6.71 mg/L) compared to controls, respectively. There were no differences in serum ferritin and transferring levels between both groups.	1. Small sample size 2. No mention of measures of accuracy (e.g., sensitivity, specificity, PPV, NPV)
Sun, Chen, Ho, Earley, Allen; 1998	Blinded retrospective review of patients with RLS	27 subjects included in the study (18 females and 9 males, ranging in age from 29 to 81).	1 Study revealed that lower ferritin level correlated significantly with greater RLS severity and decreased sleep efficiency; Also showed that patients with lower ferritin levels (<50	1. Retrospective design 2. Small sample size 3. No mention of measures of accuracy (e.g., sensitivity,

			mcg/l) showed more PLMS with arousal than did RLS with higher ferritin levels.	specificity, PPV, NPV)
O'Keeffe, Gavin, Lavan; 1994	Prospective study	36 elderly patients, 13 female, 5 male in each group. Median age was 81 (age range 70-87);	1. Serum ferritin levels were lower in the RLS group than in the control group (median 33μg/l vs. 59μg/l p< 0.01), while serum iron, B12, and folate values were similar. inverse correlation between ferritin levels and RLS symptoms (-0.53, p<0.01). 2. For RLS patients with ferritin levels <100 who were prescribed ferrous sulfate 200 mg three times a day and completed 12-week treatment, all showed an increase in serum ferritin levels, and RLS median scores also improved.	1. Severity of illness scale used not validated; 2. Criteria used to make diagnosis of RLS is less specific than criteria currently being used. 3. Small sample size; 4. No control during treatment phase 5. No blinding during treatment phase 6. No mention of measures of accuracy (e.g., sensitivity, specificity, PPV, NPV)
Sloand, Shelly, Feigin, Bernstein, Monk; 2004	Double-blind, placebo-control study	25 subject (11 in the treatment group, 14 in the placebo group); age range-36 to 74.	1. Improvement in RLS symptoms occurred only in the treated group during week 1, but was greatest at week 2 of the study. Salutary effects of iron persisted at 4 week, but were not statistically significant. Significant increases in serum ferritin and iron saturation were	1. Small sample size 2. Tool used to assess severity of RLS-no mention of it being validated; 3. Short duration of study 4. No mention of measures of accuracy (e.g., sensitivity, specificity, PPV, NPV)

			noted in the treated group, but not noted in the placebo-treated group	
Aul, Davis, Rodnitzky; 1998	Retrospective review	113 subjects, 72 (64%) female, 41 (36%) males. Ages ranged from 24 to 90 years (mean age 64 +/- 14 years).	1. Of the 80 patients that had a CBC, 21% were anemic. Of the 48 patients with iron studies, 62.5% had low serum iron levels, and 77% had low iron saturation. Of the 20 subjects that had ferritin studies, 25% had low levels. Of the 47 subjects that had both CBCs and iron studies, serum iron was low in 64%. Of these 30 subjects, 33% were anemic. Iron saturation was low in 78%; Of these 37 subjects with low iron saturation, 27% were anemic.	1. No mention of measures of accuracy (e.g., sensitivity, specificity, PPV, NPV)
Siddiqui, Kavanagh, Traynor, Mak, Deighan, Geddes; 2005	Multicenter cross-sectional study	277 patient records were reviewed, 127 subjects had RLS, 150 subjects had no RLS;	1. Logistic regression MVA revealed that gender, duration since first dialysis and increasing body weight are the only statistically significant variables predicting RLS; anemia and ferritin levels not associated with risk of RLS	1. Potential limitations include: risk that other unidentified factors might better explain risk for RLS (R^2 =0.13); potential for confounding of the data related to not differentiating between symptoms that only occurred during dialysis and symptoms that occurred at other times 2. No mention of measures of accuracy (e.g., sensitivity, specificity, PPV, NPV)

O'Keeffe, Noel, Lavan; 1993	Cross-sectional study	15 patients (11 women and 4 men) with a diagnosis of RLS.	1. Serum ferritin was lower in the RLS group compared to the non-RLS group ($p<0.025$). Of the 3 patients that completed a course of ferrous sulfate (200 mg three times a day), all reported substantial improvement in leg symptoms.	1. Small sample size 2. Ferritin levels marking deficiency is not consistent with levels used in other studies; 3. Diagnostic criteria used to make diagnosis is not the same as criteria currently used; 4. Subjects receiving ferrous sulfate not blinded, no control during treatment phase 5. No mention of measures of accuracy (e.g., sensitivity, specificity, PPV, NPV)
Akyol, Kiylioglu, Kadikoylu, Bolaman, Ozgel; 2003	Cross-sectional study	Of 34 patients with iron deficiency anemia (IDA), only 14 had symptoms of RLS; age and sex adjusted	1. Neurological examination (e.g, electrophysiological examination including motor and sensory nerve conduction, F-responses, H-reflexes, blink reflexes, and mixed nerve silent periods) were normal for both groups. The study noted that IDA did not seem to cause electrophysiological changes in the peripheral nerves, brainstem, spinal cord, so measurement of these parameters in IDA patients does not seem to be effective in confirming	1. Small sample size 2. No mention of measures of accuracy (e.g., sensitivity, specificity, PPV, NPV)

			RLS	
Allen, Barker, Wehrl, Song, Earley; 2001	Cross-sectional study	N=10 (5 subjects with RLS, 5 serving as controls)	1. Regional brain iron concentration (R_2) was significantly decreased in the substantia nigra and the putamen in patients suffering the RLS compared to control group	1. Small sample size 2. No measures of accuracy used (sensitivity, specificity), Positive predictor values (PPV) or negative predictor values (NPV)
O'Keeffe; 2005	Prospective study	80 consecutive subjects (mean age 71.2) seen over a 5-year period by physicians who have a special interest in patients with RLS	1. Serum ferritin levels <50 ng/ml) were present in 22% of subjects with onset before age 50, 39% of those with onset at 50 to 64 years, and 58% of those with onset after 64 years (p=0.009).	1. Authors relied on self-reporting of symptoms by subjects 2. Findings might not be generalizable to target population 3. No measures of accuracy used (sensitivity, specificity), Positive predictor values (PPV) or negative predictor values (NPV)
Rich; 2000	Between group comparison	N=24 (8 subjects with familial RLS and RLS onset before age 45 years, eight with non-familial on onset after age 45 years, and 8 age-matched controls)	1. Study found no difference between those with familial and those without familial disease; serum iron of patients were higher than controls (yet both were within normal ranges); no differences in either ferritin or transferrin in serum; CSF ferritin levels were significantly lower and CSF transferrin levels were higher in RLS subjects versus controls; CSF iron levels did not show a difference; no correlation between serum and CSF	1. Small sample size 2. No measures of accuracy used (sensitivity, specificity), Positive predictor values (PPV) or negative predictor values (NPV)

			iron levels.	
Collado-Seidel, Kohnen, Samtleben, et al. 1998	Prospective study, involved patients with uremia	N=136 (32 participants had definitive RLS, 12 participants had questionable RLS, 12 patients reported RLS in the past, and 88 participants reported no symptoms of RLS)	Study revealed that ironstudies revealed no statistical difference between group with definitive RLS and group without RLS, only parathormone levels were statistically different.	1. Uremia may act as a confounder in this relationship 2. No measures of accuracy used (sensitivity, specificity), Positive predictor values (PPV) or negative predictor values (NPV)
Lee, Zaffke, Baratee-Beebe; 2001	A secondary analysis of a previous longitudinal study of pregnant patients	N=45 (though 45 subjects were recruited for the study, only 30 subjects completed all 3 trimesters and the postpartum time points of 3-4 weeks and 11-12 weeks.	1. Patients with RLS at preconception had low serum ferritin levels, and significantly lower folate levels at each trimester point 2. Rather than indicators of iron deficiency anemia or pernicious anemia, it was reduced serum folate levels that was associated with RLS	1. A convenience sample was recruited for study 2. Selection bias could be introduced 3. Did not follow intention to treat Approach 4. No measures of accuracy used (sensitivity, specificity), Positive predictor values (PPV) or negative predictor values (NPV)
O'Keeffe; 2005	Case study	83 year old patient with RLS	1. Subject's severity score (IRLSSG rating) was 29 out of 40 2. Serum ferritin was 93 mcg/L 3. Other iron studies (Hgb., mean corpuscular volume, transferrin) were within normal limits	1. Case study 2. No measures of accuracy used (sensitivity, specificity), Positive predictor values (PPV) or negative predictor values (NPV)
Ondo, Tan, Mansoor; 2000	Cross Sectional study	N=68 (all with a diagnosis of RLS based on IRLSSG criteria)	1. Subjects without a family history of RLS were more likely to have lower ferritin levels, more cases of	1. Cross Sectional study 2. No measures of accuracy used (sensitivity, specificity), Positive predictor values (PPV)

			neuropathy, and older age at onset of symptoms. 2. Subjects in the study had no clinical or laboratory evidence of rheumatologic disorder	or negative predictor values (NPV)
Ondo, Vuong, Jajkovic; 2002	Cross Sectional study	N=303 (all patients had Parkinson's Disease), queried about RLS symptoms	1.Low serum ferritin levels were associated with RLS symptoms in patients with Parkinson's Disease (P=0.01)	1. Study used patients with Parkinson's Disease which could be a confounder. 2. Unable to generalize to non-Parkinson Medicare population 3. No measures of accuracy used (sensitivity, specificity), Positive predictor values (PPV) or negative predictor values (NPV)
Sahil, Mills, Webley; 1994	Cross Sectional study	N=175 patients with Rheumatoid arthritis based on ARA 1987 criteria; RLS based on Gibb and Lee's criteria	1. RLS symptoms more frequent in RA patients than non-RA patients 2. Ferritin and Hgb levels lower in RLS patients than in RA controls	1. Gibb and Lee's criteria used to establish RLS (other studies use IRLSSG criteria 2. No measures of accuracy used (sensitivity, specificity), Positive predictor values (PPV) or negative predictive values (NPV)
Siddiqui, Kavanagh, Traynor, Mak, Deighan, Geddes; 2005	Cross Sectional study	N=277 (127 with RLS and 150 without RLS).	1. Iron studies including ferritin failed to show statistical difference between both groups 2. Female gender, duration of dialysis and increased body were statistically associated with RLS	1. Patients receiving dialysis only generalizable to Medicare patients receiving dialysis 2. No measures of accuracy used (sensitivity, specificity), Positive predictor values (PPV) or negative predictive values
Silber, Richardson; 2003	Prospective study	Though 245 with RLS were enrolled in the study, only 8 met	1. In 75% of subjects, RLS started at about the time of or after blood	1. Small sample size 2. No measures of accuracy used (sensitivity, specificity),

		inclusion criteria	donations. 2. In 25% of subjects, symptoms related to RLS resolved correction of iron stores	Positive predictor values (PPV) or negative predictive
Winkelman, Chertow, Lazarus; 1996	Cross Sectional study	N=333 (204 subjects with ESRD, and 129 controls with heart disease)	1. Low association between RLS symptoms and Hct (-0.21) and Hgb (0.22)	1. No ferritin studies involving the use of ferritin 2. No measures of accuracy used (sensitivity, specificity), Positive predictor values (PPV) or negative predictive

APPENDIX B

General Methodological Principles of Study Design
(Section VI of the Decision Memorandum)

In general, when making national coverage determinations, CMS evaluates relevant clinical evidence to determine whether or not the evidence is of sufficient quality to support a finding that an item or service is reasonable and necessary. The overall objective for the critical appraisal of the evidence is to determine to what degree we are confident that: 1) the specific assessment questions can be answered conclusively; and 2) the intervention will improve net health outcomes for patients.

We divide the assessment of clinical evidence into three stages: 1) the quality of the individual studies; 2) the generalizability of findings from individual studies to the Medicare population; and 3) overarching conclusions that can be drawn from the body of the evidence on the direction and magnitude of the intervention's potential risks and benefits.

The methodological principles described below represent a broad discussion of the issues we consider when reviewing clinical evidence. However, it should be noted that each coverage determination has its unique methodological aspects.

Assessing Individual Studies

Methodologists have developed criteria to determine weaknesses and strengths of clinical research. Strength of evidence generally refers to: 1) the scientific validity underlying study findings regarding causal relationships between health care interventions and health outcomes; and 2) the reduction of bias. In general, some of the methodological attributes associated with stronger evidence include those listed below:

- Use of randomization (allocation of patients to either intervention or control group) in order to minimize bias.
- Use of contemporaneous control groups (rather than historical controls) in order to ensure comparability between the intervention and control groups.
- Prospective (rather than retrospective) studies to ensure a more thorough and systematical assessment of factors related to outcomes.
- Larger sample sizes in studies to demonstrate both statistically significant as well as clinically significant outcomes that can be extrapolated to the Medicare population. Sample size should be large enough to make chance an unlikely explanation for what was found.
- Masking (blinding) to ensure patients and investigators do not know to which group patients were assigned (intervention or control). This is important especially in subjective outcomes, such as pain or quality of life, where enthusiasm and psychological factors may lead to an improved perceived outcome by either the patient or assessor.

Regardless of whether the design of a study is a randomized controlled trial, a non-randomized controlled trial, a cohort study or a case-control study, the primary criterion for methodological

strength or quality is the extent to which differences between intervention and control groups can be attributed to the intervention studied. This is known as internal validity. Various types of bias can undermine internal validity. These include:

- Different characteristics between patients participating and those theoretically eligible for study but not participating (selection bias).
- Co-interventions or provision of care apart from the intervention under evaluation (performance bias).
- Differential assessment of outcome (detection bias).
- Occurrence and reporting of patients who do not complete the study (attrition bias).

In principle, rankings of research design have been based on the ability of each study design category to minimize these biases. A randomized controlled trial minimizes systematic bias (in theory) by selecting a sample of participants from a particular population and allocating them randomly to the intervention and control groups. Thus, in general, randomized controlled studies have been typically assigned the greatest strength, followed by non-randomized clinical trials and controlled observational studies. The design, conduct and analysis of trials are important factors as well. For example, a well designed and conducted observational study with a large sample size may provide stronger evidence than a poorly designed and conducted randomized controlled trial with a small sample size. The following is a representative list of study designs (some of which have alternative names) ranked from most to least methodologically rigorous in their potential ability to minimize systematic bias:

- Randomized controlled trials
- Non-randomized controlled trials
- Prospective cohort studies
- Retrospective case control studies
- Cross-sectional studies
- Surveillance studies (e.g., using registries or surveys)
- Consecutive case series
- Single case reports

When there are merely associations but not causal relationships between a study's variables and outcomes, it is important not to draw causal inferences. Confounding refers to independent variables that systematically vary with the causal variable. This distorts measurement of the outcome of interest because its effect size is mixed with the effects of other extraneous factors. For observational, and in some cases randomized controlled trials, the method in which confounding factors are handled (either through stratification or appropriate statistical modeling) are of particular concern. For example, in order to interpret and generalize conclusions to our population of Medicare patients, it may be necessary for studies to match or stratify their intervention and control groups by patient age or co-morbidities.

Methodological strength is, therefore, a multidimensional concept that relates to the design, implementation and analysis of a clinical study. In addition, thorough documentation of the conduct of the research, particularly study selection criteria, rate of attrition and process for data collection, is essential for CMS to adequately assess and consider the evidence.

Generalizability of Clinical Evidence to the Medicare Population

The applicability of the results of a study to other populations, settings, treatment regimens and outcomes assessed is known as external validity. Even well-designed and well-conducted trials may not supply the evidence needed if the results of a study are not applicable to the Medicare population. Evidence that provides accurate information about a population or setting not well represented in the Medicare program would be considered but would suffer from limited generalizability.

The extent to which the results of a trial are applicable to other circumstances is often a matter of judgment that depends on specific study characteristics, primarily the patient population studied (age, sex, severity of disease and presence of co-morbidities) and the care setting (primary to tertiary level of care, as well as the experience and specialization of the care provider). Additional relevant variables are treatment regimens (dosage, timing and route of administration), co-interventions or concomitant therapies, and type of outcome and length of follow-up.

The level of care and the experience of the providers in the study are other crucial elements in assessing a study's external validity. Trial participants in an academic medical center may receive more or different attention than is typically available in non-tertiary settings. For example, an investigator's lengthy and detailed explanations of the potential benefits of the intervention and/or the use of new equipment provided to the academic center by the study sponsor may raise doubts about the applicability of study findings to community practice.

Given the evidence available in the research literature, some degree of generalization about an intervention's potential benefits and harms is invariably required in making coverage determinations for the Medicare population. Conditions that assist us in making reasonable generalizations are biologic plausibility, similarities between the populations studied and Medicare patients (age, sex, ethnicity and clinical presentation) and similarities of the intervention studied to those that would be routinely available in community practice.

A study's selected outcomes are an important consideration in generalizing available clinical evidence to Medicare coverage determinations. One of the goals of our determination process is to assess net health outcomes. These outcomes include resultant risks and benefits such as increased or decreased morbidity and mortality. In order to make this determination, it is often necessary to evaluate whether the strength of the evidence is adequate to draw conclusions about the direction and magnitude of each individual outcome relevant to the intervention under study. In addition, it is important that an intervention's benefits are clinically significant and durable, rather than marginal or short-lived. Generally, an intervention is not reasonable and necessary if its risks outweigh its benefits.

If key health outcomes have not been studied or the direction of clinical effect is inconclusive, we may also evaluate the strength and adequacy of indirect evidence linking intermediate or surrogate outcomes to our outcomes of interest.

Assessing the Relative Magnitude of Risks and Benefits

An intervention is not reasonable and necessary if its risks outweigh its benefits. Net health outcomes is one of several considerations in determining whether an item or service is reasonable and necessary. CMS places greater emphasis on health outcomes actually experienced by patients, such as quality of life, functional status, duration of disability, morbidity and mortality, and less emphasis on outcomes that patients do not directly experience, such as intermediate outcomes, surrogate outcomes, and laboratory or radiographic responses. The direction, magnitude, and consistency of the risks and benefits across studies are also important considerations. Based on the analysis of the strength of the evidence, CMS assesses the relative magnitude of an intervention or technology's benefits and risk of harm to Medicare beneficiaries.