Appendix e-1. System for Classification of Evidence of Risk Factors in ALS: Rating of Analytic Epidemiological Articles

Reproduced from Armon C. An evidence-based medicine approach to the evaluation of the role of exogenous risk factors in sporadic amyotrophic lateral sclerosis. Neuroepidemiology 2003; 22:217-228, with permission from S. Karger AG, Basel.

Class I: One of the following:

1. Prospective or retrospective cohort study with parallel controls.

    a. Exposure, and hence assignment to the "exposed" cohort, established before knowledge of diagnostic status, or without knowledge of diagnostic status, or confirmable independently of the knowledge of diagnostic status. Consideration of and accounting for possible misclassification.
    b. Unexposed cohort is appropriate to the risk factor in question, is well-matched to the exposed cohort on factors other that the exposure, and is otherwise representative of the general population.
    c. Diagnosis of ALS made applying uniform efforts and criteria to exposed and unexposed cohorts.
    d. Loss to follow-up low, and comparable in exposed and unexposed cohorts. Possible roles of competing causes of mortality accounted for. Preferably – all mortality data available for both cohorts.
    e. Exposure quantified, where possible, to permit assessment of dose-response relationships.
    f. Sources of biases and confounding identified and accounted for.
    g. Conclusion based on large numbers. Appropriate statistical analysis.

2. Population-based case-control studies.

    a. Putative risk factor or exposure occurred before probable biologic onset of disease.
    b. Demonstration that ascertainment of patients is complete in the given population.
    c. Appropriate choice of controls, to assure they are matched to the patients and are also representative of the general population. (Assure adequate matching, avoid "overmatching").
    d. High response rates from patients and controls.
    e. Uniform effort to gather information equally from affected and unaffected individuals.
    f. Blinding of information-gathering method to individuals' disease status ideal; if not done – adequate justification as to why this does not affect the assessment of the risk factor in question.
    g. Blinding of subjects and individuals gathering the data as to the hypotheses being tested. If not done – adequate justification as to why this does not affect the assessment of the risk factor in question.

h. Meticulous attention to avoiding recall bias or, if not possible, to evaluating its impact, estimating the magnitude of its impact and controlling for it.
i. Diagnosis of ALS made applying established criteria.
j. Exposure quantified, where possible, to permit assessment of dose-response relationships.
k. Sources of biases and confounding identified and accounted for.
l. Conclusions based on large numbers. Appropriate statistical analysis. Methods state if hypotheses were selected a priori for confirmatory analysis. If more than one exposure considered in exploratory analysis, statistical significance is established with correction for multiple comparisons.

Class II

1. Cohort studies with parallel controls meeting most of criteria b-g, where the findings may be considered valid for the risk factor in question. This requires justification. (Criterion a is mandatory).

2. Population-based case-control studies meeting most of criteria c-l, where the findings may be considered valid for the risk factor in question. This requires justification. (Criteria a and b are mandatory).

3. Well-designed case-control studies that are not population-based, meeting most of criteria c-l. Criterion a is mandatory. Justification is necessary, why the findings may be considered valid for the risk factor in question, with initial attention to referral bias.

Class III.

1. Cohort studies with parallel controls where not all of the criteria b-g have been met, and consequently bias or confounding may account for the findings with regards to the risk factor in question, but not to an extent that would invalidate the findings completely.

2. Case-control studies where not all the criteria b-l have been met, and consequently bias and confounding may account for the findings with regards to the risk factor in question, but not to an extent that would invalidate the findings completely.

Findings which result from otherwise unbiased exploratory analysis, or encountered through the performance of multiple comparisons, may belong in this class, provided no sources of bias or other material limitations are present. If there are additional limitations, then the evidence is Class IV or V. Assignment to Class III requires justification.

Class IV.

All other studies with controls, where the risk factor occurred before biological disease onset. Results that do not attain statistical significance. Results of post hoc analyses, uncorrected for implicit multiple comparisons.

Class V.

1. Studies with controls where the risk factor studied most likely occurred after biological disease onset. Assignment to this class is specific to that risk factor.
2. Uncontrolled data (case series, case reports, chance observations, expert opinions that are not based on verifiable data).

Comment: Epidemiologic studies are designed to look for presence and magnitudes of associations. Findings of absence of association require special consideration. *Absence* of *evidence* (of association) is not equivalent to *evidence* of *absence*. In general, power calculations are needed to provide an estimate of the Type II error (the likelihood of missing a true association where one is present). This information is needed in order to know how likely it is that absence of association is due to chance or to a sample size too small to detect a true effect of a predetermined magnitude. If power calculations are available, failure to find an association may be construed as class I or class II evidence in support of a conclusion that there is no association if the power is sufficiently large, typically >80%. In the absence of power calculations, failure to find an association might be considered at most as class III evidence in support of a conclusion that there is no association. However, power calculations are usually not performed when designing epidemiologic studies, and do not guide sample size. In fact, sample size is usually determined by practical constraints of resources (time and budget). Power calculations are not provided in published reports, and may be difficult to derive in retrospect.

Therefore, I am proposing an additional criterion that might permit considering failure to find an association as (at most) class II evidence for lack of association. It has two elements: (a) the finding is based on a large sample of patients and a large proportion of patients in the sample and (b) the 95% confidence interval (CI) around the odds ratio of 1.0 is "tight." There are no universally accepted definitions of "large number of patients" or of "tight 95%CI." I propose that the number of patients may be considered "large" if the actual number of patients in the sample is greater than 50 and the proportion of patients be considered "large" if all, or close to all, the patients in the sample inform the conclusion. A 95% confidence interval round 1.0 may be considered "tight" if its upper limit (UL) is less than 2.0 (to conclude "no increased risk") or if its lower limit (LL) is greater than 0.5 (to conclude "no protective effect"). If this criterion is to be applied, then the number and proportion of patients on whom this conclusion is based should be provided, as well as the actual 95% confidence interval, so that readers may decide how robust they consider the conclusion.