Guidance for Industry

Drug Development for Amyotrophic Lateral Sclerosis

ALS Guidance Participants

The following individuals were leading participants in the patient-centered, community-led ALS Guidance initiative.

Steering Committee:

- Barb Newhouse, *Steering Committee Chair*; President and CEO, The ALS Association
- Lucie Bruijn, *Steering Committee Vice Chair*; Chief Scientist, The ALS Association
- Hiroshi Mitsumoto, *Steering Committee Vice Chair*; Columbia MDA/ALS Research Center
- Rick Bedlack, *chair, Public Policy working group*; Duke University Medical Center
- James Berry, *chair, Biomarkers working group*; Massachusetts General Hospital
- Benjamin Brooks, *chair, Benefit/Risk working group*; Carolinas Health System
- Merit Cudkowicz, *co-chair, Clinical Trials and Outcome Measures working group*; Massachusetts General Hospital
- Valerie Cwik, Muscular Dystrophy Association, EVP and Chief Medical and Scientific Officer
- Ted Harada, *co-chair, Patient & Caregiver Advisory Committee*; Person with ALS, ALS Association National Trustee
- Don Johns, *Steering Committee Member*; Vice President, Head ALS Neuroscience Discovery & Development, Biogen
- Catherine Lomen-Hoerth, *chair, FTD/ALS working group*; UCSF
- Nicholas Maragakis, *chair, Diagnosis working group*; Johns Hopkins Hospital
- Tim Miller, *chair, Natural History working group*; Washington University School of Medicine
- Elizabeth Ottinger, *Steering Committee Member*; NIH/NCATS
- Steve Perrin, *Steering Committee Member*; ALS Therapy Development Institute, Chief Scientific Officer
- Jeremy Shefner, *co-chair, Clinical Trials and Outcome Measures working group*; Barrow Neurological Institute
- Stephen Winthrop, *co-chair, Patient & Caregiver Advisory Committee*; Person with ALS, ALS Association National Trustee, Research Ambassador
Benefit/Risk Working Group:

- **Chair**, Ben Brooks, Carolinas Neuromuscular/ALS-MDA Center
- William Cho, Genentech
- Stephen Finan, Person with ALS
- Stephen Finger, HopeNow4ALS, Person with ALS
- Eric Gascho, National Health Council
- Kimberly L Goslin, Providence ALS Center
- Stephen Aaron Goutman, University of Michigan Health System
- Laurie Gutmann, University of Iowa Carver College of Medicine
- Christopher Lee, Vanderbilt University Medical Center
- Joseph Lewcock, Genentech
- Björn Erik Oskarsson, University of California, Davis School of Medicine
- Stephen N Scelsa, Icahn School of Medicine at Mount Sinai
- Katalin Scherer, University of Arizona School of Medicine & Banner University Medical Center
- Ericka P Simpson, Houston Methodist Neurological Institute & Weill Cornell Medical College
- Andrew Wolff, Cytokinetics, Chief Medical Officer

Natural History Working Group:

- **Chair**, Tim Miller, Washington University School of Medicine
- Mary Kay Floeter, NINDS, Chief NINDS EMG Section, Deputy Clinical Director
- Matt Harms, Columbia University
- Eric Macklin, Massachusetts General Hospital
- Bernard Ravina, Voyager Therapeutics, VP Clinical Development
- John Ravits, UC San Diego
- Stacy Rudnicki, University of Arkansas
- Steven Ziegler, Person with ALS

Diagnosis Working Group:

- **Chair**, Nick Maragakis, Johns Hopkins University, Professor, Dept. of Neurology
- Nazem Atassi, Massachusetts General Hospital MDA ALS Clinic
- Kevin Boylan, Mayo Clinic - Jacksonville
- David Ennist, Origent Data Sciences, Chief Science Officer
- Amanda Haidet-Phillips, MDA, Chief Science Officer
- Carlayne Jackson, UT San Antonio
- Sabrina Paganoni, Spaulding Rehabilitation Hospital and Massachusetts General Hospital
- Karen Shideleff, Person with ALS
- Bryan Traynor, NINDS, Chief, Neuromuscular Disease Research Section
Biomarkers Working Group:

- **Chair**, James Berry, Massachusetts General Hospital
- Michael Benatar, University of Miami
- Gil Block, Neuraltus Pharmaceuticals, Chief Medical Officer
- Bob Bowser, Dignity Health
- Terry Heiman-Patterson, Drexel University and ALS Hope Foundation
- Andreas Jeromin, Iron Horse Diagnostics
- Melanie Leitner, Biogen Idec, Associate Director, Clinical Research
- Terry Ann McNearney, Eli Lilly & Co, Senior Medical Advisor
- Lyle Ostrow, Johns Hopkins School of Medicine
- Erik Pioro, Cleveland Clinic
- Wes Rose, Person with ALS
- Seward Rutkove, Beth Israel Deaconess Medical Center

Clinical Trials & Outcome Measures Working Group:

- **Co-Chair**, Merit Cudkowicz, Massachusetts General Hospital MDA ALS Clinic
- **Co-Chair**, Jeremy Shefner, Barrow Neurological Institute
- Jinsy Andrews, Cytokinetics, Director, Clinical Research and Development
- Frank Bennett, ISIS Pharmaceuticals, SVP Research
- Chris Coffey, University of Iowa
- Robin Conwit, NINDS, Medical Director, Office of Clinical Research
- David Ennist, Origent Data Sciences, Chief Science Officer
- Toby Ferguson, Biogen Idec, Medical Director, ALS iHub
- Madeline Kennedy, Person with ALS and Research Ambassador
- Steve Kolb, Family Member and Research Ambassador
- Michael McDermott, University of Rochester Medical Center
- David Schoenfeld, Harvard School of Medicine and Professor in the Department of Biostatistics, Harvard School of Public Health

Public Policy Working Group:

- **Chair**, Rick Bedlack, Duke University Medical Center
- Tricia De Santis, Biogen Idec, VP Regulatory Policy
- Steve Gibson, The ALS Association, Chief Mission Strategy and Public Policy Office
- Michael Gollin, Person with ALS
- Ghazala Hayat, St. Louis University
- Ed Kasarskis, University of Kentucky
- Jon Katz, Forbes-Norris
- Kim McCleary, FasterCures, Managing Director
- Paul Mehta, CDC, Agency for Toxic Substances and Disease Registry
- Jess Rabourn, ALS Emergency Treatment Fund, Managing Director
- Kristin Stephenson, MDA, VP Policy and Advocacy
- Pat Wildman, The ALS Association, VP Public Policy
- Stephen Winthrop, Person with ALS, Research Ambassador and ALS Association National Trustee

**FTD and ALS Working Group** (material was developed on FTD in ALS and then incorporated into other guidance sections as appropriate)
- *Chair*, Catherine Lomen-Hoerth, UCSF
- Sharon Abrahams, University of Edinburgh
- Lora Clawson, Johns Hopkins
- Laura Goldstein, Kings College
- Murray Grossman, University of Pennsylvania
- Jennifer Murphy, UCSF
- Steve Reznick, Person with ALS and Research Ambassador
- Susan Woolley, Forbes Norris

**ALS Patient & Caregiver Advisory Committee:**
- *Co-Chair*, Ted Harada, Person with ALS, ALS Association National Trustee
- *Co-Chair*, Stephen Winthrop, Person with ALS, ALS Association National Trustee, Research Ambassador
- Clay Ahrens, Person with ALS
- Pat Bradley, Person with ALS
- Steve Byer, ALS Worldwide, Co-Executive Director and family member
- Mike Cherepov, Person with ALS
- Clare Durrett, Team Gleason, Board Member
- Stephen Finan, Person with ALS
- Stephen Finger, HopeNow4ALS, Person with ALS
- Deb Gaudet, Person with ALS
- Michael Gollin, Person with ALS
- Bob Hebron, HopeNow4ALS, family member
- Shelly Hoover, Person with ALS
- Madeline Kennedy, Person with ALS
- Robert Kleiner, Project ALS
- Steve Kolb, Family Member, Research Ambassador
- Christopher Leidigh, Person with ALS
- Jehad Majed, Family Member, HopeNow4ALS
- Bob Murray, Family Member and Research Ambassador
- Jess Rabourn, ALS Emergency Treatment Fund
- Steve Reznick, Person with ALS, Research Ambassador
- Shay Rishoni, Prize4Life, CEO and Person with ALS
- Wes Rose, Person with ALS
- Craig Sanderson, Kevin Turner Foundation
- Karen Shideleff, Person with ALS
- Kathy Thomas, Family Member and Research Ambassador
• Joe Wise, Person with ALS
• Evan Yegelwel, Person with ALS
• Steven Ziegler, Person with ALS

Technical Support:
• Lisa Bain, Science & Medical Writer
• Allison Durham, FaegreBD Consulting
• Tim Franson, Chief Medical Officer, YourEncore
• Steve Gibson, The ALS Association
• Pat Wildman, The ALS Association
• Dave Zook, FaegreBD Consulting
ALS Guidance Outline

I. INTRODUCTION

II. BACKGROUND

III. BENEFIT RISK
   A. General Comments
   B. Evidence of patient and caregiver preferences
   C. Implications of patient risk/benefit preferences in ALS.
      1. Alternative endpoints.
         Proposed Guidance
      2. Placebo Control Arm.
         Proposed Guidance
      3. Threshold for statistical significance.
         Proposed Guidance
         Proposed Guidance
      5. Role of expanded access and accelerated approval.
         Proposed Guidance
      6. Patient protection and need for informed decisions in consultation with physician.

IV. THE CURRENT UNDERSTANDING OF THE NATURAL HISTORY OF ALS
   A. General Comments
   B. Testing and evaluation across the course of the disease
      1. ALS functional rating scales
      2. Upper motor neuron involvement
         a. Quantification of clinical signs
         b. Direct assessments of UMN dysfunction and degeneration
      3. Cognitive dysfunction
         a. Scope of cognitive dysfunction in ALS
         b. Influence of Cognitive Dysfunction on the Natural History of ALS
         c. Role of genetic mutations
         d. Detection and tracking of cognitive dysfunction in ALS
         e. Clinical characteristics of Frontotemporal Dementia and ALS
      4. Muscle Strength
         a. Rationale for use
         b. Current practice
      5. Respiratory Assessment
         a. Current practice
b. Comparison of SVC, FVC, and supine FVC  
c. Assessing inspiratory function  
d. Evaluation of current assessments  
6. Bulbar function tests  
   a. Current Practice  
   b. Future Assessments Needed  
7. Following Nutritional Status  
   a. Standard Measures  
   b. Total Daily Energy Expenditure  
8. Quality of Life and Survival Measures  
   a. Rationale for Quality of Life Assessments  
   b. Current practice  
9. Electrophysiological Measures  
   a. Rationale for electrophysiological measures  
   b. Current practice  
10. Laboratory measures  
11. Biomarkers  

C. Phenotypes, phenotype heterogeneity, and disease progression  
1. Clinical phenotypes  
   a. Introduction.  
   b. Clinical phenotypes based on level of involvement  
      “Typical” ALS  
      Primary lateral sclerosis (PLS)  
      Progressive muscular atrophy (PMA)  
   c. Clinical phenotypes based on body region of involvement  
      Bulbar and pseudobulbar palsy  
      Limp regional variants including flail leg, flail arm, 
      polynuertic pattern, and hemiplegia (Mill’s variant)  
      Upper extremity regional variant  
      Lower extremity regional variant  
      Mill’s variant (hemiplegic ALS)  
   d. Clinical phenotypes with involvement of non-motor regions  
      Frontotemporal dementia  
      Other system involvement  
   e. Clinical focality, stochasticity and spread  
   f. Progression, progression rates and survival  
   g. Progression and severity of frontotemporal dementia and ALS  
2. Staging  

D. Genetic Predictors of Disease  

E. Effect of standard interventions on the natural history of ALS  
1. Nutrition management  
   a. Rationale for nutrition management  
   b. Difficulties leading to weight loss  
   c. Use of GT in nutrition management  
      Survival benefits of enteral nutrition  
      Decision to obtain GT
GT procedures
   d. Nutrition management studies
2. Physical and occupational therapy
   a. Rationale for Physical and Occupational Therapy
   b. Exercise and ALS
   c. Forms of effective physical therapy
   d. Therapies to improve communication
3. Respiratory management
   a. Rationale for respiratory management
   b. NIV survival benefits
   c. Timing of NIV initiation
   d. NIV compliance/tolerance
   e. Advantages of NIV
   f. Alternatives and additions to NIV usage
   g. TIV as alternative to NIV
   h. Effects of TIV usage
4. Psychological support and End-of-Life Care
   a. Importance of end-of-life care
   b. Types of palliative therapies
   c. Relevance to clinical trial participation

V. DIAGNOSIS
   A. General comments
      1. Who makes a diagnosis of ALS
      2. Neurological history and examination
   B. Barriers to early diagnosis
      1. Impact of diagnostic milestones on delayed diagnosis
      2. Financial barriers to diagnosis
      3. Heterogeneity of ALS Disease Presentation and Progression
      4. Frontotemporal Dementia in ALS
         Clinical Diagnosis of FTD
   C. Diagnostic Criteria
      1. The El Escorial Criteria-Revised
      2. Awaji Diagnostic Algorithm
         Proposed Guidance
   D. Diagnostic Laboratory Investigations/Methods for Confirming Diagnosis
      1. Blood and urine tests
      2. CSF studies
      3. Electrophysiological Studies
      4. Imaging Studies
      5. Nerve/muscle biopsy
      6. Genetic Analyses
         Proposed Guidance

VI. BIOMARKERS
   A. General comments
B. Current State of Biomarkers in ALS
   1. Identifying and Using Biomarkers
   2. Current promising biomarkers (limitations/strategic)
      a. Biological-fluid-based Biomarkers
         Neurofilaments
         P75 Neurotrophin Receptor Extracellular Domain
         \( p75NTR_{ECD} \)
         SOD1 and C9RANT
      b. Eletrophysiological Biomarkers
         Motor unit number estimation (MUNE), index (MUNIX),
         and compound motor action potential (CMAP)
         Electrical impedance myography (EIM)
      c. Imaging Biomarkers

C. Use of biomarkers in trials
   1. Context and illustrations of biomarker incorporation in ALS trials
   2. Defining Biomarker Utility for a Trial; Context of Use as a Framework for
      Maximizing Biomarker Benefit in ALS Trials

D. Summary and Proposed Guidelines

VII. CLINICAL TRIALS AND OUTCOME MEASURES

A. General comments
B. Clinical trials
   1. Inclusion/exclusion criteria
      a. Familial ALS
      b. Genetic factors
      c. Disease heterogeneity
      d. Cognitive / behavioral status
      
      Proposed Guidance
   2. Feasibility issues
      a. Ease of use and missing data
   3. Trial Design
      a. Study Duration
      b. Choice of treatment arms
      c. Designs in rare/genetic forms.
      d. What constitutes an adequate control group?
         Advantages of historical controls
         Disadvantages of historical controls
         Discussion Summary
         Ethics of using historical controls
      e. Use of predictive algorithms in trial design
         Proposed Guidance
      f. Role of open-label extension studies
      g. Allocation ratio
      h. Function-specific therapies
      i. Trials of combination therapies

C. Outcome measures and endpoints
1. Outcomes sufficient to support approval
   a. Survival
   b. Function
      ALSFRS-R
      
      Proposed Guidance
   c. Strength
      i. Respiratory function
         Vital Capacity
         Sniff nasal inspiratory pressure (SNIP)
         Maximum inspiratory pressure (MIP) / maximum expiratory pressure (MEP)
         Maximum voluntary ventilation (MVV)
         
         Proposed Guidance
      ii. Quantitative Extremity Strength
         Tufts Quantitative Neuromuscular Examination (TQNE)
         Hand held dynamometry (HHD)
         Manual muscle testing (MMT)
         Hand grip strength
         Muscle strength fatigability
         ATLIS
         
         Proposed Guidance
   d. Cognitive and behavioral scales
      
      Proposed Guidance
   e. Specific tests of function
      i. 6 minute walk test
      ii. Timed up and go
      iii. Bulbar/Pseudobulbar Function
         Center for Neurologic Bulbar Function Scale (CNS-BFS).
         Speech Rate
         ALSFRS-R
         Timed Swallow Tests
         Modified Barium Swallow (MBS)
         Center for Neurological Study - Lability Scale (CNS-LS)
         
         Proposed Guidance

2. Outcomes that would support approval secondarily

   Proposed Guidance
   a. Patient / Caregiver reported outcomes
      i. SF-36
      ii. ALSAQ-40/ASLAQ-5
      iii. EQ-5D
      iv. McGill QOL
      v. Sickness Impact Profile (SIP)
      vi. Clinical Global Impression (CGI) scale
      vii. Caregiver Burden Inventory
   b. Structural measures
      MRI
PET
c. Neural function
   Electrical impedance myography (EIM)
   Compound Motor Unit Action Potential (CMAP)
   Motor Unit Number Estimation/Motor Unit Number Index
   (MUNE/MUNIX)
   Excitability testing
   Transcranial magnetic stimulation (TMS)

3. Outcomes that assay potential disease pathways
   a. Inflammatory markers
   b. Markers of oxidative stress
      activity of catalase
      glutathione
      uric acid
      8-oxo (or 8-hydroxy)deoxyguanosine
      15-F2t-isoprostane
      plasma protein carbonyl
   c. Markers of cell damage
   d. Other

4. Time-to-failure endpoints

5. Safety outcomes

D. Statistical issues and analysis challenges in ALS
   1. Missing Data
   2. Incorporating deaths in the analysis of longitudinal outcomes
      a. Approach 1
      b. Approach 2
      c. Approach 3

3. Use of less stringent significance levels for hypothesis testing

E. Preclinical requirements – should they be different than for less serious illnesses?

VIII. PUBLIC POLICY

A. General Comments

B. Expanded Access
   1. Patient eligibility for expanded access
   2. Drug eligibility for expanded access
   3. Benefits of expanded access protocols for drug development

C. National ALS Registry

D. FDA Stakeholder Communication

E. FDA Accountability in Using Patient-oriented tools
   1. Introduction
   2. Patient perspectives

F. Regulatory implications of targeted drug development, and development
   Incentives, including evidence for reimbursement
I. INTRODUCTION

The purpose of this guidance is to assist sponsors in the clinical development of human drugs and therapeutic biological products for the treatment of Amyotrophic Lateral Sclerosis (ALS).

This guidance is the result of a collaboration between the FDA, the ALS Association and other patient advocacy organizations, clinicians experienced in treating ALS patients, patients and caregivers, and researchers from academia, industry, the National Institutes of Health, and the Center for Disease Control and Prevention. It reflects both the FDA’s recognition of the substantial unmet medical need that exists for patients with ALS as well as the importance of engaging patients and caregivers in efforts to develop new treatments. The FDA invited the ALS community to develop the initial draft of this guidance in accordance with the FDA’s Good Guidance Practice provisions as described in CFR Title 21 (21CFR10.115).

Community awareness campaigns about ALS, its presentation, and the need for research have long been a part of efforts to improve communication about ALS. Social media and the Internet have now begun to play significant roles in disseminating information about symptoms and signs of ALS. Nowhere has this been more evident than the attention brought to ALS by the Ice Bucket Challenge, which began in 2014 and continues on an annual basis. The Ice Bucket Challenge helped promote ALS awareness within the lay community as well as raise financial resources to expand support for research and patient care. Organizations with longstanding commitments to supporting ALS clinics nationally and raising community awareness have included The ALS Association and the Muscular Dystrophy Association. Both the Muscular Dystrophy Association and The ALS Association maintain a network of specialized clinics throughout the US to deliver the most up-to-date care to patients within these communities. These organizations also help to communicate research breakthroughs and advances in clinical care to the patient community and general public. Additionally, in 2008, Congress enacted the ALS Registry Act to create the first National ALS Registry with the goal of identifying persons with ALS in the community, learning more about the disease presentation and progression, and identifying potential causes of ALS.

In February of 2015, The ALS Association launched this effort to develop the first-ever community-driven drug development guidance for ALS with funding from the ALS Ice Bucket Challenge, bringing together over 100 participants including people living with ALS, caregivers, researchers, clinicians, and industry experts from across the world to contribute their expertise and experience. In parallel with this effort, a committee was formed to update the ALS clinical trial guidelines that were developed more than 15 years ago. The guidance and guidelines have different audiences and different goals, yet are meant to be consistent. This guidance is intended to represent the Agency’s interpretation of, or policy on, evaluating new ALS therapies for approval in order to assist industry in navigating the regulatory process. The clinical trial guidelines incorporate stakeholder views across all phases of drug development, from pre-clinical to market approval, and serve as “best practices” for clinical trial design and are used by
researchers and industry to provide structure and direction for the design and conduct of clinical trials in ALS. The goal of clinical trial guidelines is to lead to more effective and efficient trials but they do not directly impact the FDA regulatory process.

This guidance reflects the FDA’s current thinking regarding the weight that should be given to the preferences of ALS patients and caregivers with regard to benefit/risk tradeoffs in light of the severity and rapid progression of the disease coupled with the lack of effective treatments. It also reflects the FDA’s recognition of the heterogeneity in the etiology, presentation, and progression of ALS and the challenges this heterogeneity presents with regard to the design and implementation of clinical trials.

The guidance further recognizes that the dynamic nature of ALS research and the continual emergence of new research discoveries mean that recommendations regarding treatment development will evolve. Thus, this guidance is meant to be viewed as a living document, capable of adapting to changes in our understanding of disease mechanisms and natural history, as well as to the development of new technologies.

FDA’s guidance documents, including this guidance, should be viewed only as recommendations unless specific regulatory or statutory requirements are cited. The word should in guidances is meant to denote a recommendation rather than a requirement.

II. BACKGROUND

ALS is a multisystem neurodegenerative disorder in which patients develop progressive paralysis involving all skeletal muscles as well as the bulbar and respiratory muscles involved in breathing, speaking and swallowing. Signs and symptoms reflect upper motor neuron (UMN) and lower motor neuron (LMN) dysfunction, which spreads within a region or to other regions over time. Commonly called Lou Gehrig’s disease after the legendary baseball player who was diagnosed with the illness in 1939 and died two years later, the disease is nearly always fatal, usually within 2 to 5 years of diagnosis (1). No cure is available and only one drug, riluzole, has been approved by the FDA for the treatment of ALS, yet even this drug has only modest benefits (2). A second drug, nuedexta, has been approved for the treatment of pseudobulbar affect (emotional lability) in patients with ALS (3).

U.S. prevalence statistics vary in different studies, ranging from approximately 12,000 with a definite diagnosis of ALS, according to the National ALS Registry (MMWR 2014), to about 25,000 when considered more broadly as part of the spectrum of motor neuron diseases, which include progressive muscle atrophy (PMA), primary lateral sclerosis (PLS), and progressive bulbar palsy (PBP) (4). This relatively low prevalence arises in part from the rapid progression of the disease, obscuring the fact that about 1 of every 800 Americans develops ALS, making it the most common adult-onset motor neuron disorder (MND) (5). ALS also co-occurs with frontotemporal degeneration (FTD) (6) in about 15% of ALS patients (7). The disease typically strikes adults over the age of 50, with the prevalence highest among those in their 70s. The prevalence is greater in men than in women and in whites than blacks.

The earliest symptoms are not specific to ALS and vary from person to person depending on the muscles affected. They include fasciculations, cramps, spasticity, muscle weakness,
slurred or nasal speech, and difficulty chewing and swallowing. Because the onset of these symptoms is typically gradual and subtle, and because no blood test or diagnostic biomarker exists, diagnosis is often delayed until the disease has progressed to the point that significant neurodegeneration has occurred. Delayed diagnosis thus has significant implications for treatment development, since demonstrating efficacy may require treating early in the disease process.

The causes of ALS are largely unknown (8). About 90-95% of ALS cases occur sporadically with no known associated risk factors and no family history. The remaining 5-10% of cases are inherited, most commonly from mutations in the chromosome 9 open reading frame 72 (C9ORF72) or the gene for the enzyme copper-zinc superoxide dismutase 1 (SOD1). Other mutations have also been linked to ALS, including genes associated with FTD, i.e., FUS and TARDPB (9). People with familial forms of ALS typically develop symptoms at a younger age than those with sporadic forms (4). A number of environmental exposures have also been proposed as risk factors for ALS (8). These include cigarette smoking, military service, head trauma, physical activity, and exposure to lead, pesticides, electromagnetic radiation, and other neurotoxins. Understanding the mechanisms by which mutations and environmental exposures may contribute to neurodegeneration is important because they may help identify molecular mechanisms of pathogenesis and thus, potential therapeutic targets (9-11). For example, one of the key molecular pathways that appears to underlie both familial and sporadic forms of ALS is glutamate-mediated excitotoxicity. Since riluzole is a glutamate antagonist, this may explain its neuroprotective benefits (12).

Riluzole has only modest benefits in terms of survival (2). Other treatments such as nuedexta are available for symptom management and to improve the quality of life of patients, but there remains an urgent need for an effective disease-modifying treatment. The lack of successful drug development results from a combination of factors: delayed or inaccurate diagnosis; an incomplete understanding of the biological mechanisms underlying disease; heterogeneity of disease presentation, progression; and etiology; as well as by the lack of reliable biomarkers, outcome measures, and novel trial designs.

This guidance reflects the need for the ALS patient and caregiver community to join with the FDA and sponsors to produce a clinical development strategy that will provide the best opportunity to demonstrate a treatment’s effectiveness and safety in this relentlessly progressive disease. We begin with a discussion of benefit/risk tradeoffs, recognizing that many patients with ALS are willing to accept increased risks in light of the ultimately fatal nature of the disease and the absence of effective treatments.
III. BENEFIT RISK

A. General Comments

There is currently one approved drug that improves survival in ALS patients, yet it only extends life expectancy by a matter of months. Without access to new effective treatments, roughly half of patients will die within 2 to 5 years of their first symptoms. Thus, it is no surprise that in numerous peer-reviewed publications and in statements by people with ALS and their caregivers, such as during the FDA’s public hearing in February 2013, people with ALS repeatedly stated their desire to accept greater risks in the search for potential therapies. The willingness of many patients to accept greater risk, the nature of this disease and the lack of current treatment options have direct implications for the drug development process with regard to: the use of alternative endpoints, the significance levels for hypothesis testing, the role of placebo observations in the control arm, the method of delivery, and the use of expanded access and accelerated approval mechanisms.

While the poor prognosis and lack of effective treatments to slow or reverse disease progression motivates patients to try experimental treatments and assume considerable risk, the increased tolerance for risk expressed by people with ALS makes them potentially vulnerable to investigational new treatments supported by only minimal evidence of tolerability, safety, and efficacy. Thus, while increased flexibility with respect to the evaluation of potential new treatments for ALS deserves serious consideration, rigorous standards must remain in place to ensure that accurate information is provided to patients and their doctors so that informed decisions can be made that do not unnecessarily compromise the health and safety of people with ALS.

It is the appropriate goal of all stakeholders in this endeavor to achieve the appropriate balance between the increased tolerance of people with ALS for drug development risk given the current prognosis and the continued protection of people with ALS from their potential exploitation.

B. Evidence of patient and caregiver preferences

In an online patient-driven survey regarding the level of risk ALS patients are willing to accept to participate in FDA clinical trials, patients indicated a willingness to accept the possibility of a serious adverse event, depending on the potential for benefit (13). Since patients often feel they have “nothing to lose,” potential treatments spur hope for patients with a disease that appears hopeless. This willingness to accept serious risk extends even to treatment where the probability of benefit is unknown and the possible risk of serious adverse events is high. This may be most evident in the eager demand from many patients with ALS for clinical trials of invasive stem cell treatments, where procedures include intrathecal or spinal cord injections (14). With other experimental treatments, many patients have expressed a willingness to accept risk even if efficacy has been shown only in preclinical models.

In the same survey, patients with ALS expressed a desire to participate in clinical trials, and would accept considerable risk if a drug showed promise. For a drug that showed some preliminary efficacy, 97% of patients surveyed would participate in a clinical trial (13). 28% percent of ALS patients would participate in a trial with potential serious or life threatening side
effects and an additional 38% of patients would participate if the drug could modify disease progression (13). 78% indicated that their patients would participate in a clinical trial regardless of efficacy data and that no risk would be considered too great by 58% of patients, although 20% felt craniotomy might be too great a risk. Most other serious adverse effects would dissuade fewer than 15% of patients (13).

ALS patients are more willing to take risks than some other populations with serious illnesses. Of oncology patients surveyed, 83% would participate in a clinical study compared to 96% of patients with ALS (15). Compared to oncology patients surveyed in 2006, ALS patients are less concerned regarding the possibility of adverse effects than oncology patients (22% versus 45%) (13, 16). ALS patients’ willingness to accept risk of adverse events when the mechanism of action is poorly understood has not been assessed. However, these survey findings suggest that investigational drug safety is less important than efficacy to many ALS patients.

Patients with ALS and their caregivers are highly interested in treatments that provide perceptible benefit in ALS, and they are most concerned with loss of independence from muscle weakness and shortened survival. Acute improvement in motor function is easier to perceive than a latent increase in survival with minimal symptomatic benefits. Riluzole has been shown statistically to improve survival, but not to improve strength or other parameters of function in ALS (12, 17). Despite the survival benefit, rates of riluzole use have declined to 53% in the USA (18) and 66% in the UK (19) since the initial FDA-approval of this medication. One reason for this decline may be the concerns that have been expressed through social media groups such as PatientsLikeMe (20). Patients and caregivers often cite the lack of symptomatic benefit as a reason to decline taking this medication, despite its very limited adverse event risk.

Acute improvement in motor function due to a treatment would also be easier to perceive than a slowing in the rate of functional decline compared to a more rapid rate of decline in the absence of treatment. Put more simply, individual patients cannot tell if they are getting worse more slowly as a result of treatment compared to how they would have progressed in the absence of treatment. Nevertheless, they clearly would prefer to slow their progression even if they experience no functional improvement. Thus, many people with ALS may prefer a benefit of improved or relatively preserved function over an isolated improvement in survival. This preference may lead many to accept a riskier set of possible adverse events, and to accept the risk associated with drugs approved based on non-traditional endpoints that may not be supported by as rigorous evidence of efficacy as is typically required, as long as adequate safety can be demonstrated. The balance between a potential increase in risk to people with ALS associated with the use of such endpoints for approval against the potential benefit of making new treatments available more quickly is discussed further below in Section III.C.1, Alternative Endpoints.

The hope that an experimental treatment will have benefit can, in some ways, be a benefit itself. Some patients, especially those with the familial form of the disease, are eager to participate in clinical research in an effort to provide knowledge that might benefit future patients, even recognizing the lack of or limited direct potential benefit they may derive from a clinical trial.

Many patients with ALS who want to participate cannot, since they are either ineligible or physically unable to access research centers. Patients want criteria that are less strict and more inclusive. However, there are statistical concerns that including patients with variable
presentations, slower disease progression, longer duration of illness, advanced respiratory insufficiency, or other measures of more advanced motor impairment might lessen the likelihood of a drug showing a statistically significant effect in a placebo-controlled trial. In addition, there is a spectrum of motor neuron diseases that may share pathophysiological deficits and share similar functional motor deficits to ALS (progressive muscular atrophy [PMA], progressive bulbar palsy [PBP], primary lateral sclerosis [PLS], and frontotemporal dementia [FTD] with ALS), yet these patients are currently not be eligible for ALS drug trials. Some of these patients eventually develop clinical ALS such as PMA (50%) and PBP (the majority) (21). (See also, Section IV, Natural History).

C. Implications of patient risk/benefit preferences in ALS.

1. Alternative endpoints.

   Given the Agency’s public desire to speed drug development by developing and selecting outcome measures that are more specific or sensitive to changes in the manifestations of the disease or more quickly demonstrate safety or efficacy than existing measures, dialog amongst FDA, people with ALS, ALS clinicians, and sponsors attempting to develop drugs for the potential treatment of ALS could be fruitful and should occur as expeditiously as possible.

   As noted above in Section B, many people with ALS are willing to accept the risk associated with drugs approved based on endpoints that may not be supported by as rigorous evidence of efficacy as more traditional endpoints might have indicated, as long as adequate safety can be demonstrated. In the absence of validated and qualified surrogate markers for ALS, the use of new pharmacodynamic or progression endpoints must be considered as acceptable alternatives (see Biomarkers section VI.A).

   Time to the first occurrence of death or tracheostomy has been viewed as an acceptable endpoint for phase 3 clinical trials of new drugs to treat ALS. However, even in an inexorably progressive and uniformly fatal disease such as ALS, such trials would be prohibitively large (especially for smaller companies which often have no marketed products to support such an effort) and would take several years to complete. For example, in the recent Phase 3 EMPOWER study of dexpramipexole in patients with ALS (22), mortality at 1 year was approximately 18%. A 20 – 25% reduction in that rate sure surely would be viewed as a clinically meaningful benefit; however, a Phase 3 clinical trial with 90% power to demonstrate such reductions with conventional statistical significance (i.e., p < 0.05) would require enrollment of approximately 1,400 to 2,100 patients and collection of approximately 500 to 850 deaths (i.e., to demonstrate reductions of 25% and 20%, respectively) with a mean follow-up duration of approximately 2.5 to 3.5 years. The time from the first patient enrolled to the last patient visit in such a trial would be approximately 3.5 to 4.5 years, which does not take into account the start-up time before the first patient could be enrolled, nor the time after the last patient visit until the data could be collected and analyzed and the results made available.

   There are too few patients with ALS in the United States to make enrolling such a mortality study feasible. The prevalence of ALS is believed to be between 4 and 6 cases within a population of 100,000, with about 25,000 Americans living with ALS at any one time (4). Many of them will not meet the study’s entry criteria; for others, an investigational center will not be practically accessible. Clearly, there is a need for clinical trial endpoints that could support
registration but that would require far fewer patients to be enrolled into trials that could be completed much more expeditiously.

While ALSFRS-R is the most commonly accepted clinical endpoint and does capture the effect of the disease on each of four functional areas, it also has its limitations. It is a self-reported, discrete measure with less than perfect test characteristics (23). Relying on the ALSFRS-R as the sole method of measuring disease progression substantially increases the risk that a study will fail to demonstrate a statistically significant treatment effect even if the treatment is efficacious.

The lack of surrogate measures for motor neuron loss in ALS that might relate to disease progression is a glaring deficiency that affects the benefit-risk assessment for development of treatments. Other alternative endpoints are also needed.(24-27).

Proposed Guidance

In view of these limitations, the FDA supports the development and use of alternative endpoints such as vital capacity, measures of muscle strength, functional measures, and pharmacodynamic or progression biomarkers to support approval of new therapies, recognizing that the value of substantiating their correlation with irreversibly morbid and mortal endpoints would require much larger and longer studies. The ALS community further recognizes (yet is also concerned) that if such studies are required before approval, then scores of patients with ALS must experience irreversible morbid and mortal events before a new drug can be approved or a new endpoint utilized.

2. Placebo Control Arm.

ALS patient and caregiver preferences regarding benefit/risk tradeoffs are of particular importance with regard to clinical trial participation. Placebo-controlled studies can be frustrating, and the possibility of randomization to placebo has been cited as a reason not to participate in clinical trials (13). While 97% of patients surveyed indicated they would enter a trial if efficacy was likely, only 57% would enter a drug trial if there was a chance they would receive placebo (13). Increasing the ratio of patients receiving active drug versus placebo may increase patient interest; however, it also increases the number of patients needed to show a drug effect in a placebo-controlled trial.

Individuals with ALS often advocate the use of historical controls or predictive algorithms since they can eliminate or reduce the size of a concurrently randomized placebo arm. This allows a larger percentage of the enrolled patients to receive active treatment, thus improving recruitment, and also reducing the number of patients required and, in turn, study costs (28). However, the use of historical controls can yield results fundamentally at odds with results of controlled studies (DiPALS vs. DPS registry). There is a real possibility that a historical control group may not be well matched to the active group under study. Sponsors utilizing single arm studies must clearly document and substantiate their choice of historical controls or the unbiased nature of their predictive algorithms before treatment begins. It is essential that baseline characteristics of historical controls and active participants are well matched and clearly described so that results can be accurately interpreted. When conducted appropriately, single arm studies without placebo control may provide critical information in the
early stages of treatment development, particularly if a positive effect on a clinical endpoint is being assessed (29).

**Proposed Guidance**

Single arm studies seeking a positive effect on disease progression have the potential to reduce the number of enrolled patients required to definitively determine the effectiveness of a treatment if appropriate historical controls or predictive algorithms are utilized and clearly justified. In supplemental analyses to support the results of double-blind, randomized clinical trials with a concurrently enrolled placebo group, we support using controls or predictive algorithms from studies with similar selection criteria that resulted in a patient population with similar baseline characteristics and used similar outcome measures, which may reduce the number of patients required in the concurrently controlled placebo group. Such approaches may limit bias and be reasonable to show drug efficacy if a drug has a robust effect (28).

**3. Threshold for statistical significance.**

The required significance level for any trial should minimize the impacts of Type I (false positive) and Type II (false negative) errors. In a disease such as ALS, with few clinically significant treatment options and a stark prognosis, the cost of approving an ineffective or dangerous treatment must be weighed against the cost of delaying or rejecting an effective treatment given the alternative for patients (30). The demonstrated appetite for risk of the patient population and the fact that roughly half of patients will die within 2 years of diagnosis (31) without the discovery of a new treatment justify an appropriate choice of statistical standard. The purpose of an efficacy trial is to determine whether or not a treatment is more effective than the baseline control. Therefore, a one-tailed test may be the more appropriate statistical analysis.

A \( p \)-value \( \leq 0.05 \) represents the current accepted standard for clinical statistical significance, meaning that statistically, the probability that the observed result is entirely due to chance is \( \leq 5\% \). This value often plays a role in the decision to approve a drug for use; however, the use of a \( p \)-value \( \leq 0.05 \) is essentially an arbitrary convention. Given their prognosis in the absence of treatment, some people with ALS may be willing to accept somewhat more uncertainty than 5% regarding whether the observations from a trial of a particular intervention are likely entirely due to chance, especially if there are no serious safety concerns associated with the intervention under study. Furthermore, the \( p \)-value is a function of the treatment effect size, the sample size, and the standard deviation. However, a drug with a large clinically meaningful effect in a subgroup of ALS patients may not demonstrate a conventional statistically significant outcome if the sample size is small. Further, challenges specific to ALS including the relatively small number of patients who have the disease and satisfy inclusion criteria, difficulty traveling to a study center (resulting in small sample sizes), disease heterogeneity and imperfect endpoints (resulting in a larger standard deviations) may further complicate statistical testing. Drugs lacking potential side effects may be judged differently from drugs with serious and known side effects; again, the ALS community is willing to accept the latter if efficacy is proven.

**Proposed Guidance**
The FDA supports continued testing, or conditional or full approval of a drug with a conventionally non-statistically significant outcome should the effect size be clinically interesting and meaningful. This is especially the case when the drug does not appear to be associated with major safety concerns, when secondary and surrogate endpoint measures appear to confirm the likely efficacy, and when post-marketing surveillance can continue to support evidence of safety and efficacy.


The method of drug delivery may be a consideration for some when considering the benefits and risks of a new therapy. Oral therapies represent the most common method of drug delivery, although this method is not necessarily trivial in subjects with ALS due to dysphagia and, in those with feeding tubes, incompatibilities with the tube material. Nonetheless, this will likely represent the most desirable method of delivery for most patients.

Subcutaneous therapies have been tested in ALS and are generally well tolerated (32-34). At least one intramuscular (IM) therapy has been investigated, but to date, no clinical trial data have been published. Intravenous therapies represent a slightly more invasive therapy compared to oral and subcutaneous routes due to the need for central access. A study of ceftriaxone tested subjects with daily or twice-daily infusions. Special attention was made to train caregivers to reduce risk of catheter based infection, resulting in an infection rate of 2.9-11.3 per 1000 catheter days (35).

Intrathecal (IT) delivery of drugs represents the next escalation of invasiveness and is used for treatment with antisense oligonucleotide (ASO) and stem cells. This route of delivery requires a lumbar puncture, which is a common outpatient procedure that most neurologists consider to be of minimal risk. In a recent trial of ASO for SOD1, 22 participants underwent a 11.5-hour IT infusion via an external pump. Post-lumbar puncture syndrome was experienced in 11/32 infusions (drug and placebo) and back pain in 8/32 (drug and placebo). In this trial the catheter tip was placed near the T8-T10 level, which has a theoretical risk of spinal cord injury (36). Several stem cell therapies have been delivered via the IT route with generally good tolerability (37).

Intraparenchymal delivery methods are also undergoing trials including stem cell injections into the cerebral cortex and spinal cord. A phase 1 trial of intraspinal stem cell transplantation study (14) utilized a risk escalation strategy whereby subjects with more severe stages of the ALS composed the initial study groups. The rationale for this study design was that these subjects were non-ambulatory, and the location for initial injections was the lumbar spine; therefore, should a severe adverse event occur, it was felt to pose a less severe risk to the subjects. At the 2015 annual meeting of the American Neurological Association, initial safety results of a phase 2 intraspinal stem cell transplantation study reported that one subject developed weakness related to the therapy or procedure. Some safety concerns regarding injection site hemorrhages and edema have also been published related to frontal cortex olfactory stem cell injections, but these AEs have not been published by the study team.

Many therapies may necessitate more invasive delivery methods. Several factors are postulated to influence patient willingness to accept a particular delivery method including disease severity. Stronger preclinical and early phase clinical trial data should increase the likelihood that patients will tolerate a riskier procedure. Further, the presence of a willing and
able caregiver will likely make subcutaneous and intravenous therapies more tolerable. Finally, the more “hope” surrounding a therapy, whether related to clinical and/or preclinical data or social media hype, the more likely this therapy may be accepted by subjects despite the delivery method risk. Given patients' eagerness to accept risk, care must be taken so that patients do not incur unnecessary delivery risks and information must be provided so that patients, caregivers and physicians understand potential complications.

5. Role of expanded access and accelerated approval.

Recognizing the devastating nature of ALS, patients’ benefit/risk preferences, the unmet medical need and the high proportion of patients who were not candidates for clinical trials, the FDA encourages sponsors of ALS products to utilize accelerated approval and expanded access programs to the extent that they do not impede the clinical trials program intended to support the formal approval of the drug, as expressed in current FDA guidance. Under the current standard of care, roughly half of patients will die within 2 years of diagnosis (31). Alternative endpoints discussed above may be used to support accelerated approval. For products approved under accelerated approval using expedited procedures, sponsors must conduct required post-approval studies and approval may be revoked if required studies failed to verify the predicted effect or other evidence demonstrates that the product is not safe or effective, recognizing that placebo-controlled trials may not be feasible or even ethical in jurisdictions where the drug has been made available under an accelerated approval designation.

Proposed Guidance

The FDA strongly supports the use of accelerated approval and expanded access programs, in conjunction with post-approval studies, for therapeutic products in development for ALS. There is a large population of individuals with ALS who are not eligible for typical clinical trials that are very eager to take part in expanded access programs. The generally high risk tolerance of these individuals would allow for lower-cost safety monitoring and utilization of alternative endpoints could allow for longer duration of study, potentially allowing discovery of benefit in patients with more advanced ALS than is typical under clinical trial investigation.

6. Patient protection and need for informed decisions in consultation with physician.

The increased tolerance for risk of many people with ALS makes them potentially vulnerable to investigational new treatments supported by only minimal evidence of tolerability, safety, and efficacy. In addition, the spectrum of the disease and its complexity during the evolution of clinical changes in ALS, as discussed in the Natural History section, leads to patients having differing degrees of risk tolerance and interpretations of the potential benefits they would receive from an investigational therapy. Therefore, patients must be provided information with respect to the safety and likely potential benefit of therapy, to the extent possible, based on evidence from pre-clinical research, Phase 1 and/or Phase 2 clinical trials. Neurologists should play an active role in helping each patient to make informed decisions regarding whether their participation in a trial or program is in the patient’s best interest.
IV. The Current Understanding of the Natural History of ALS

A. General comments

ALS is a clinical syndrome named for its neuropathological hallmark: degeneration of motor neurons in the spinal anterior horn and motor cortex and loss of axons in the lateral columns of the spinal cord. **Clinically**, ALS is defined by history establishing muscle dysfunction over time and space, and by physical examination showing signs of both UMN and LMN dysfunction in one or more body regions. **Neuropathologically**, ALS is defined by degeneration of motor neurons and their axonal projections in brain, brainstem, spinal cord, and peripheral nerve, and now increasingly by a sophisticated repertoire of pathological biomarkers. The neuropathological molecular signature common to almost all sporadic ALS (SALS) and most familial ALS (FALS) is TDP-43 immunoreactive neuronal cytoplasmic inclusions. The molecular pathological features of ALS variants PLS and PMA are less certain, but also appear to share the primary features of ALS. **Genetically**, at least five to ten percent of ALS cases can be attributed to a genetic cause, such as mutations in SOD1, FUS, and C9ORF72. The disease spectrum thus genetically comprises distinctive molecular and neuropathological signatures of ALS-related genes. The key clinical, pathological, and genetic differences form the basis for general classifications (Table 1).

Table 1: ALS Classifications

<table>
<thead>
<tr>
<th>Based on Phenotype</th>
<th>Based on Molecular Pathology</th>
<th>Based on Genetics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Typical ALS</td>
<td>TDP-43 proteinopathy</td>
<td>Sporadic ALS</td>
</tr>
<tr>
<td>Bulbar/pseudobulbar ALS</td>
<td>(ubiquitinated pathology)</td>
<td>Familial ALS, incl.:</td>
</tr>
<tr>
<td>Limb onset variants:</td>
<td>• Without repeat expanded C9ORF72</td>
<td>• ANG</td>
</tr>
<tr>
<td>• Typical limb onset</td>
<td>• With repeat expanded C9ORF72</td>
<td>• C9ORF72</td>
</tr>
<tr>
<td>• Flail arm or bibrachial ALS</td>
<td>SOD1 proteinopathy</td>
<td>• FIG4</td>
</tr>
<tr>
<td>• Flail leg</td>
<td>FUS proteinopathy (basophilic inclusion body disease (BIBD))</td>
<td>• FUS</td>
</tr>
<tr>
<td>• Polyneuritic variant</td>
<td></td>
<td>• OPTN</td>
</tr>
<tr>
<td>• Hemiplegic ALS</td>
<td>(Mill’s variant)</td>
<td>• SOD1</td>
</tr>
<tr>
<td>(Mill’s variant)</td>
<td>Primary lateral sclerosis</td>
<td>• TARDBP</td>
</tr>
<tr>
<td>Progressive muscular atrophy</td>
<td></td>
<td>• UBQLN</td>
</tr>
<tr>
<td>ALS with associated FTD or impairment of higher cortical function</td>
<td></td>
<td>• VAPB</td>
</tr>
</tbody>
</table>
B. Testing and evaluation across the course of the disease

1. ALS functional rating scales

As in many neurological disorders, functional scales have been developed for ALS. The ALSFRS and revised version containing respiratory function (ALSFRS-R) are the most widely used to measure function in ALS clinical trials. Other functional scales include the Norris Scale (1974), Appel Scale (1987), and ALS Severity Scale (1989). The ALSFRS-R is described in more detail in the Clinical Trials and Outcome Measures section (Section VII). Aggregate data from multiple clinical trials using the ALSFRS provides progression rates and demonstrates the wide variability in ALS regarding progression rates (see below regarding heterogeneity). ALSFRS may also be a useful measure of disease progression for natural history studies of genetic subsets of ALS.

In the classic ALS phenotype, the average decline in the ALSFRS-R has been estimated at about 1 unit/month (see Table 2 below) (38). However, the rate of decline differs widely among individual patients. Some studies categorize patients into rapid- or slow-progressors based on their initial progression rate (initial ALSFRS-R score/months from symptom onset) of greater or less than 0.5 points/month (39) or interval progression >8 points over three months (40). However, longitudinal studies indicate that the decline in the ALSFRS-R score may be greater in the first year to 18 months of symptoms and at late stages of illness, particularly after 5 years (41, 42). Older age, bulbar onset of symptoms, and the initial ALSFRS-R progression rate are associated with a faster rate of the decline in the ALSFRS-R (43). Although the decline in the ALSFRS-R tends to be more linear in the intervening years, short plateaus of stability of several months may occur in up to 25% of patients (44). There are few studies comparing rate of decline of the ALSFRS-R in other motor neuron disorder phenotypes. In a study of 37 patients with PMA, the ALSFRS-R score declined by 1.85 points/3 months (45). In a study of 12 PLS patients, the ALSFRS-R declined by 3 points over 2 years (46). Several studies have shown that the ALSFRS-R is a good predictor of survival (47-51).

Table 2: Progression of ALS Disease Outcome Measures

<table>
<thead>
<tr>
<th>Outcome Measure</th>
<th>Average Progression Rate</th>
<th>Range of Progression Rates</th>
<th>Characteristics of Progression Rate</th>
<th>References (Placebo Cohorts of ALS Clinical Trials)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALS Functional Rating Scale (ALSFRS-R)</td>
<td>-1.03 units/month</td>
<td>-0.70 to -2.20 units/month</td>
<td>- Fastest changes are early (first 18 months) and late (&gt; 5 years) in disease progression (43, 52)</td>
<td>CNTF, 1996 (54), BDNF, 1999 (55), Topiramate, 2003(56), Creatine, 2004 (57), Celecoxib, 2006 (58), Pentoxifylline, 2006 (59), TCH346, 2007 (60), Minocycline, 2007 (61), IGF, 2008 (34), Dexpramipexole, 2013 (22),</td>
</tr>
<tr>
<td></td>
<td>At onset:</td>
<td></td>
<td>1st year -0.60 units/month; 2nd year -0.34 units/month (41)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>&gt; -0.5 units/month</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>At onset, slow progressors:</td>
<td>&lt; -0.5 units/month (39)</td>
<td>Modifying factors: age, bulbar onset, initial progression rate/shorter delay</td>
<td></td>
</tr>
</tbody>
</table>


| Vital capacity (FVC/VC) | -2.22% per month | -1.10% to -2.80% per month | • FVC/VC of bulbar onset ALS declines faster than upper and lower limb disease onset(66)  

2. Upper motor neuron involvement

UMN involvement in ALS is clinically identified by increased deep tendon reflexes, the emergence of pathological reflexes (e.g. Babinski's sign, Hoffman's sign, etc.), or the development of spasticity. These signs reflect dysfunction and/or degeneration of cortical motor neurons, and are important features that help establish a diagnosis of ALS. Determining the relative contributions of UMN and LMN dysfunction to a patient's clinical syndrome is difficult except in the extremes of PLS (pure upper motor neuron) and PMA (pure lower motor neuron). Furthermore, it is often the case that the progressive loss of lower motor neurons eventually masks the presence and degree of upper motor neuron dysfunction. As a result, UMN involvement or progression has not routinely been incorporated into clinical trials. This may also be due to the fact that the available clinical examination signs (deep tendon reflexes and tone evaluation) are relatively insensitive and show poor inter-rater reliability (72-74). To address these limitations, methods have been developed to quantify examination findings (reflexes and spasticity) while other approaches have aimed at more directly investigating upper motor neuron function or numbers (imaging and electrophysiologic function).

a. Quantification of clinical signs: Heterogeneous scales are used by clinicians to rate reflexes (72), but most utilize a 5-point scale similar to that put forward by the NINDS (0=no response, 1=hypo-active response, 2=response in lower half of normal range, 3=response in upper half of normal range, 4=exaggerated response, including with clonus) (75). Because of variability and reliability issues (72), methods to quantify various facets of reflexes have been developed. These methods have focused on standardizing limb position, standardizing the force of the stimulus delivered, and the quantitative measurement of the latency and force of the resulting muscle contraction (76, 77). The most commonly-used rating scale for spasticity in ALS is the Ashworth spasticity scale (78) or modified versions of it. However, the reliability and natural history of these scales and techniques have not been established in ALS, limiting their utility as outcome measures.
b. Direct assessments of UMN dysfunction and degeneration: Imaging paradigms aimed at assessing UMN involvement include measuring cortical thickness (morphometric MRI), integrity of corticospinal tracts (diffusion tensor imaging), and metabolism (MR spectroscopy). While all show abnormalities in patients with ALS compared to controls and may facilitate diagnosis, none have been validated as longitudinal markers of progression. Similarly, electrophysiologic function of cortical neurons is abnormal when investigated by transcranial magnetic stimulation techniques, but there is disagreement about whether measured parameters deteriorate with disease progression (79, 80).

3. Cognitive dysfunction
   a. Scope of cognitive dysfunction in ALS: ALS is increasingly recognized as a multi-system disease in which cognitive dysfunction is common. The onset of cognitive dysfunction can be difficult to date, but can precede motor symptoms of ALS by years, begin concurrent with weakness, or develop during ALS progression (81). Some degree of cognitive impairment (typically fronto-executive dysfunction) is found in 40-50% of patients (82, 83), with 10-15% meeting research criteria for frank frontotemporal dementia (FTD) (82, 83).

   b. Influence of Cognitive Dysfunction on the Natural History of ALS: The presence of FTD or fronto-executive cognitive impairment negatively influences the natural history of ALS in several ways. First, both are associated with decreased survival in proportion to the severity of the impairment (84, 85). There is some evidence to suggest that the timing of cognitive symptom initiation may influence this effect, with one study showing that patients with the simultaneous onset of cognitive and motor symptoms had substantially worse survival compared to patients with cognitive symptoms and delayed motor involvement or motor involvement with later cognitive impairment (81). The presence of FTD is also associated with reduced compliance with standard of care treatments (non-invasive ventilation, gastrostomy placement for nutritional support) (84), and an increased burden on caregivers and caregiver stress (86).

   c. Role of genetic mutations: The prevalence of FTD in ALS with known genetic cause varies by the implicated gene. SOD1 mutations are almost never associated with FTD, while it more frequently occurs with TARDBP and FUS, and is quite common in those with C9ORF72 mutations. Because of this association, some centers routinely screen for C9ORF72 mutations in ALS patients with cognitive involvement.

   d. Detection and tracking of cognitive dysfunction in ALS: Because the most common impairment is in fronto-executive circuits, screening tools developed for the detection of Alzheimer’s disease-type cognitive dysfunction (such as the mini-mental state exam, or MMSE) may not be sensitive (7). Several ALS-specific screening tools have been developed and validated but the field has not settled on a preferred screening instrument yet. The two most commonly utilized in clinical care are the Edinburgh Cognitive and Behavioural ALS Screen (ECAS) (87, 88) and the ALS Cognitive Behavioral Screen (ALS-CBS) (89). A full neuropsychological evaluation battery remains the gold-standard for research use. Understanding how cognitive impairment progresses during ALS is an urgent research need currently being addressed (90) but complicated by high attrition rates and rapid declines in an ALS patient’s ability to participate in assessments.
e. Clinical characteristics of Frontotemporal Dementia and ALS: The behavioral and cognitive characteristics of FTD can be a feature of ALS, occurring in approximately 15% of ALS patients. Milder cognitive and behavioral disorders that do not rise to the level of severity found in FTD can be seen in an additional 35% of patients with ALS. Thus, up to half of ALS patients may have cognitive difficulty, making it an integral component of the disease. The link between ALS and FTD is confirmed by the presence of TDP-43 pathology that is common to both disorders, and families with C9ORF72 repeat expansions feature carriers with ALS, FTD and/or both disorders (see below). The presence of cognitive difficulty that is consistent with ALS and not due to another condition thus should not be a criterion for exclusion from a trial. However, the presence of FTD may need to be taken into consideration when assessing outcomes of a trial; for example, ALS patients with FTD have worse prognosis and are less compliant with interventions such as NIV.

The most prevalent type of dementia in ALS is behavioral variant FTD, occurring in about 70% of those with concurrent ALS and FTD, and subtypes of progressive aphasia have been described in the remaining patients (PNFA, SD). Behavioral features in ALS patients with FTD as well as the milder cognitive and behavioral disorders include personality change, apathy, disinhibition, loss of sympathy/empathy, obsessive-compulsive behaviors, and hyperorality; related deficits in social cognition and executive dysfunction include difficulties in higher order tasks such as verbal fluency, dual-tasking, planning, theory of mind, and working memory. Language features can include effortful and non-fluent speech characterized by grammatical simplification and errors, although a smaller number of ALS patients can exhibit empty speech with word-finding problems and poor comprehension of words and objects. Late-onset psychotic symptoms, such as somatic delusions, also associate with FTD.

4. Muscle strength

a. Rationale for use: Progressive muscle weakness is a hallmark sign of ALS and therefore an important outcome measure of disease progression. Quantitative strength measures are closely correlated with motor neuron loss (91, 92) and demonstrate a remarkably linear, predictable loss of strength within each patient (93-95). However, strength data among groups of patients is difficult to interpret because of the heterogeneity of ALS (site of onset, pattern of spread, 10-fold or more variations in disease duration) (96, 97).

Unlike other neurological diseases such as Parkinson’s or multiple sclerosis, ALS has no exacerbations or remissions. Instruments capable of accurately testing both very strong and very weak muscles in all extremities and converting raw data to a percent of predicted normal based on biometric factors over a sufficient time period provide a disease progression rate for each individual. Thus, measuring each individual’s strength decline during a sufficient lead-in phase may provide improved evidence of efficacy.

b. Current practice: Tests currently used to evaluate muscle strength include: Manual muscle testing (MMT), Tufts Quantitative Neuromuscular Examination (TQNE), Quantitative strength testing using hand held dynamometry (HHD), Hand Grip strength using a JAMAR hydraulic dynamometer, and Muscle strength fatigability. Full descriptions of these tests can be found in the Clinical Trials and Outcome Measures section (Section VII).
5. Respiratory Assessment

a. Current practice: The respiratory measurements used to evaluate the pulmonary status of patients with ALS vary between clinics and countries, and may include forced vital capacity (FVC), slow vital capacity (SVC), arterial blood gasses, end tidal CO2, overnight pulse oximetry, polysomnography, peak cough flow, maximum inspiratory pressure (MIP), maximum expiratory pressure (MEP), sniff inspiratory pressure (SNIP) and MVV (103-106).

Of the various tests available, in the United States, FVC is most commonly performed in part because insurance guidelines to institute noninvasive ventilation are predicated on an FVC below 50% (though other measures such MIP < 60 cm H20 or nocturnal desaturations also may qualify for insurance). FVC correlates with survival; it is typically monitored every 3 to 6 months (67, 105, 107). The rate of decline in FVC in ALS patients averages 2.4-3.5% per month and those with faster rates have earlier onset of dyspnea (108, 109).

b. Comparison of SVC, FVC, and supine FVC: When performing FVC, patients with bulbar weakness may have abrupt cut off because of glottic closure related to upper motor dysfunction. In recent years, SVC has been gaining popularity compared to FVC in clinical trials (22, 110-112) as it may be more reproducibly performed in patients with bulbar involvement (104). Measuring supine FVC or SVC, while logistically challenging in patients with extremity and truncal weakness, may be a more sensitive way to detect early changes in diaphragmatic strength (113-115). VC measurements are not only influenced by muscle strength, but are also impacted by the airways, chest wall, and lung parenchyma.

c. Assessing inspiratory function: Inspiratory muscle strength is mediated primarily through the diaphragm, with accessory muscles also contributing. MIP is the most widely used measure of inspiratory function, but is effort-dependent and may be difficult for bulbar patients, while SNIP is felt to be easier for a patient with bulbar weakness to perform. Mouth weakness may limit a patient’s ability to perform MIP and MEP, but is not the case with SNIP so it can readily be done throughout the course of the disease. However, both are less reproducible if done by an inexperienced evaluator (104, 113, 116). Abnormalities in SNIP, MIP, sniff transdiaphragmatic pressure, and oximetry desaturations of <90% for a cumulative minute tend to occur before a decline in FVC (116, 117). Expiratory muscle strength is evaluated through MEP and PCF, and poor values may be associated with a weak cough and reduced ability to clear secretions (113).

d. Evaluation of current assessments: Comparing the usefulness of various measures in predicting survival, a single normal MIP (>70 cm H20), MEP (>70 cm H20) or FVC (>80% of predicted) were highly predictive of being alive at one year, while arterial pressure of carbon dioxide was not (115). Given the desire to identify a measurement that is not patient dependent, non-invasive, and easily performed, phrenic nerve stimulation recording a compound muscle action potential over the diaphragm has appeal. Decreased amplitudes and prolonged motor distal latencies were more common in those with reduced VC and respiratory symptoms (118). Small amplitude responses correlated with worse survival in patients with both spinal and bulbar onset disease (118).

6. Bulbar function tests

a. Current Practice: There is no standardized method for measuring bulbar changes over the course of disease in ALS. The most commonly used clinical scale is the ALSFRS-R
(64), which contains one item each grading speech, salivation and swallowing. Less commonly used scales are the ALS severity scale, which includes one item for speech and one for swallowing (119) and a patient-developed CNS-bulbar function scale (120) with one item grading speech, swallowing and salivation. The CNS lability scale is a patient self-report instrument that has been used in clinical trials to demonstrate changes in pseudobulbar affect (121). For clinical decision making, a videofluoroscopy swallowing study is most commonly used to determine whether patients may safely continue oral intake.

b. Future Assessments Needed: Voice quality, rate of speech, and communication effectiveness were identified as early predictors of bulbar dysfunction in ALS (122, 123). Comprehensive measures of the motor systems contributing to these features of speech have been developed (124, 125) but require specialized instrumentation and time for administration that limit their widespread use. Electrical impedance myography of the tongue is emerging as a more rapid measurement that can be followed longitudinally, but also requires specialized equipment (126). Efforts are ongoing to develop a comprehensive quantitative assessment battery (125) and screening tests simple enough (e.g. timed reading and swallowing) to administer at each clinic visit (127). To date, there are few data regarding their longitudinal reliability.

7. Following nutritional status

a. Standard measures: Weight, body mass index, and percentage change in weight or BMI may be used to follow the nutritional status in patients with ALS (128). In addition to recording this information, patients should be asked about symptoms of dysphagia on average every 3 months (107). As patients become increasing immobile, utilizing a wheelchair scale in the clinic or in the patient’s home is essential so accurate weights can be tracked. 'Bedside' swallowing studies done in the clinic, fiberoptic endoscopic evaluation of swallowing, or a videofluoroscopy swallowing study may be done depending upon symptoms and circumstances but are not required in documenting dysphagia before advising a patient to consider a gastrostomy tube (GT) (107). Dysphagia, muscle atrophy, and depression or respiratory symptoms causing a poor appetite have long been thought to play a role in weight loss in patients with ALS (128-130), and mounting evidence suggests patients with ALS may be hypermetabolic (131-133), further contributing to the loss of weight.

b. Total Daily Energy Expenditure: Total daily energy expenditure (TDEE) is a combination of the thermal effects of feeding, physical activity, and resting metabolic rate (RMR), with the latter contributing 75% of the TDEE in a sedentary person. Traditionally, a patient’s total daily caloric needs are calculated using the Harris Benedict equation, which takes into account gender, age, and weight and then multiplying it by a number determined by the person’s level of activity (134). While ALS patients may be physically inactive, there may be increased caloric consumption related to the increased work of using weak muscles including those associated with breathing. ALS patients may also have increased energy expenditure related to the non-functional work of spasticity, cramps, fasciculations, and pseudobulbar affect. Measuring TDEE through the doubly-labeled water method and estimating the resting metabolic rate through the Harris Benedict equation found that for the majority of ALS patients the TDEE exceeded the RMR by much more than expected. A potentially more accurate way to measure an ALS patient’s total daily caloric needs was developed using the Harris-Benedict equation but
also incorporating scores of 6 items from the ALS functional rating scale related to speech, handwriting, dressing and hygiene, turning in bed, walking and dyspnea (135), and is as follows:

\[
TDEE = [\text{Harris-Benedict RMR}] + (55.96 \times \text{ALSFRS-6 Score}) - 168
\]

It remains to be seen if utilizing this novel, ALS specific determination of caloric needs will lead to better weight preservation compared to traditional methods.

8. Quality of Life and Survival Measures

a. Rationale for Quality of Life Assessments: An additional consideration in determining the value of an experimental treatment is its impact on the ALS patient’s quality of life (QOL). Scales to measure QOL or the impact of illness on the patient and caregiver are frequently incorporated as secondary outcome measures in clinical trials. While QOL measures can be discordant with physical outcome measures of disease progression (136), they capture values of particular importance to patients.

b. Current practice: The most commonly used QOL scales in ALS clinical trials are the 36-item Short-Form Health Survey (SF-36; www.sf-36.org), the Amyotrophic Lateral Sclerosis Assessment Questionnaire (ALSAQ-40) and its short version, the ALSAQ-5 (137, 138). Longer questionnaires, such as the Sickness Impact Profile, a 136-item patient-reported questionnaire to measure sickness-related dysfunction (139), are more comprehensive but are less commonly used in clinical trials because of the time involved and patient acceptance (140). More information on specific QOL measures can be found in the Clinical Trials and Outcomes Measures section (Section VII).

9. Electrophysiological measures

a. Rationale for electrophysiological measures: Functional rating scales and quantitative strength assessments are important tools for charting the progression of weakness in ALS but only indirectly reflect the death of motor neurons. Because motor neuron death is the pathological step underlying ALS progression, considerable effort has been invested in developing electrophysiological paradigms that provide longitudinal estimates of the remaining number of these cells. Functional rating scales and quantitative strength assessments are evaluated through tests such as Compound Motor Action Potential (CMAPs), Motor Unit Number Estimation (MUNE), Motor Unit Number Index (MUNIX), Electrical Impedance Myography (EIM), and Transcranial Magnetic Stimulation (TMS). Additional discussion on these tests can be found in the Clinical Trials and Outcome Measures section (Section VII).

b. Current practice: In the clinical evaluation of ALS, none of the electrophysiological parameters above are routinely obtained to track the progression of ALS. This is largely because methods are time consuming for patients and evaluators, require specialized training and equipment, or, as of yet, do not clearly give information about progression that is not readily obtained from functional and strength assessments.

10. Laboratory measures

Overall, laboratory measures are not currently used to help understand progression rates or other measures of the natural history of ALS.

11. Biomarkers
As discussed in the Biomarkers section (Section VI), there are many efforts underway to establish biomarkers of ALS. Most of these studies are cross sectional and will need further evaluation with longitudinal samples to better define their utility both in clinical trials and as measures of ALS natural history.

C. Phenotypes, phenotype heterogeneity, and disease progression

1. Clinical phenotypes

   a. Introduction. Heterogeneity of clinical phenotypes is characteristic—these are mostly determined by the anatomic location of neuropathology, which is imputed. There are vastly different body regions that are affected, degrees of involvement of UMN and LMN, and progression rates. In addition, there are varying degrees of involvement of other systems, especially cognition and behavior but also body metabolism and appetite. FALS and SALS are not clinically distinguishable.

   b. Clinical phenotypes based on level of involvement:

      **"Typical" ALS (Table 3):** Weakness in classical ALS has simultaneous UMN and LMN characteristics. The weakness typically begins insidiously in discrete body regions and advances steadily over time and space. It presents in any of the three main body regions (face, arm, and leg), although occasionally begins in the muscles affecting the trunk and/or respiration. The co-mixture of UMN and LMN clinical signs is variably distributed with a possible skew to LMN predominance (141).

      **Primary lateral sclerosis (PLS) (Table 3):** PLS is the designation for a clinical presentation with solely or predominantly UMN level involvement. It remains unknown whether PLS is a discrete syndrome or a variant of ALS (142-147). In over half of PLS patients, symptoms begin insidiously in the legs and ascend smoothly and relatively symmetrically to arms and bulbar muscles. Others have a patchy progression, often with prominent bulbar symptoms. There is disagreement concerning the degree of LMN involvement especially as identified by EMG findings (148, 149). Patients with clinically pure PLS and no EMG changes four years after symptom onset have decades-long survival (148, 150), whereas minor EMG or LMN findings predict a poorer prognosis, consistent with typical ALS patients presenting with predominant UMN signs (151). Thus, the diagnosis of PLS should be made only after four years of disease duration (148). PLS may stabilize after a few years of progression (152), although similar stabilization may occur in UMN-dominant ALS (UMN-D ALS) (151, 153).

      Frontotemporal dementia (FTD), cognitive impairment, and altered behavior occur in PLS, at levels comparable to ALS (154). Ultimately, PLS is a clinical diagnosis that relies upon exclusion of other known causes of progressive spasticity, such as sporadic presentations of hereditary spastic paraparesis (155).

      **Progressive muscular atrophy (PMA) (Table 3):** PMA is the designation for clinical syndromes with solely or predominantly LMN involvement. Onset can begin in any body region and compared to typical ALS, PMA patients are more likely to be men and have a higher age of onset. Approximately 30% of PMA patients develop UMN signs within 18 months (45, 156). A subset of patients, characterized by segmental involvement for more than four-years duration, have slow progression and prolonged survival, though transition to ALS can occur even in this group (157, 158). Patients with PMA demonstrate the same frontotemporal pattern of cognitive
involvement as is seen in typical ALS and thus the degree of UMN involvement does not correlate with cognitive involvement (159).

c. Clinical phenotypes based on body region of involvement:

**Bulbar and pseudobulbar palsy (Table 3):** While the designations PLS and PMA are based on the level of the underlying pathology, another set of designations is based on the body region first affected at the outbreak of disease. When ALS begins by affecting the muscles of speech, mastication and swallowing, it is designated bulbar-onset ALS. The designation bulbar has traditionally signified predominantly LMN involvement and the designation pseudobulbar has traditionally signified predominantly UMN involvement, but often bulbar is used as parlance for both. EMG positive (meaning LMN is involved) and EMG negative (meaning only UMN involvement) have similar progression. Interestingly, there is a female predominance in bulbar palsy, compared with other limb regional forms where there is male predominance. Bulbar onset is more highly associated with affect and cognition and often has the added feature of altered and exaggerated emotional expression; delineation has permitted a clearer understanding of the natural history (160). Bulbar symptoms are often directly correlated with depression. Neurophysiological studies have identified neural networks underlying corticobulbar control of swallowing that are especially affected during repetitive movements (161). Functional MRI studies of the course of bulbar and limb-onset ALS are providing insights into the interrelationship between brainstem derived and spinal cord derived neural networks (162). A treatment based on dextromethorphan has an attenuating effect on pseudobulbar affect.

**Limb regional variants including flail leg, flail arm, polyneuritic pattern, and hemiplegia (Mills’s variant) (Table e):** When ALS begins by affecting muscles of the limbs, as it does two-thirds of the time, it is referred to as limb-onset, or typical, ALS as discussed above. But within this group, a few variant phenotypes have stood out and been given special designations, with a view that these variants may have variant biology. Typically, these variants are predominantly LMN syndromes and tend to be very slowly progressive.

**Upper extremity regional variant:** This is a regional variant consisting of weakness exclusively confined initially to the upper extremities. Cases have also been described as hanging arm syndrome, neurogenic man-in-the-barrel, flail arm syndrome, brachial amyotrophic diplegia, and the Vulpian-Bernhart syndrome. These patients have bilateral upper extremity weakness and atrophy that affects predominantly the proximal arms and shoulder girdle (158, 163). The average age of onset does not differ from that of ALS, but compared with ALS, this syndrome is significantly more common in men. Average survival is approximately 5 years; however, the definitions used for these patients has been slightly different. Some patients presenting with this phenotype can go on to develop a typical ALS course. Katz et al used an 18-month time of weakness confined to the arms and no UMN signs; Wijeskera used 12 months and patients could have UMN signs. In the original series of Katz, after a mean follow-up of 5.5 years, weakness remained restricted to the upper extremities in 7 out of 19 patients (163).

**Lower extremity regional variant:** This LMN variant confined to the legs is known as the pseudopolyneuritic variant of ALS, the Marie-Patrikios form, flail leg syndrome, the peroneal form of ALS, and leg amyotrophic diplegia (158, 164, 165). It is rare (about 3-3.5% of all motor neuron disease cases), predominantly male, predominantly LMN, and relatively slowly progressive with mean survival ranging from 76 to 96 months.
**Mill’s variant (hemiplegic ALS):** This is a disputed rare variant phenotype characterized by a progressive hemiplegic pattern of motor deficit that ascends from the leg or descends from the arm. This could represent a variant of PLS. The existing scarce literature suggests that it is simply a descriptive clinical term. A positron emission tomography (PET) study in one such patient demonstrated a striking lateralization of microglial activation in the hemisphere contralateral to the hemiplegia (166).

d. Clinical phenotypes with involvement of non-motor regions:

*Frontotemporal dementia (Table 3):* The overlap of FTD and ALS has been well documented in FTD patients with co-morbid motor neuron degeneration and in ALS patients with frontotemporal dysfunction (82, 167-169). Up to 15% of FTD patients and 30% of ALS patients experience the overlap syndrome. The syndrome may be difficult to identify since ALS patients’ behavioral or cognitive abnormalities may be subtle and since patients are often seen in either a neuromuscular clinic or a memory disorders center. New designations, called behaviorally impaired and cognitively impaired ALS, were created to reflect uncertainty as to whether or not they may have different underlying biologies (170). Key tests that are useful to look for cognitive behavioral impairment and exclude depression are beginning to emerge (52). Survival is impacted for both disorders in the co-morbid condition, making identification of this syndrome critical. There is a survival difference of more than a year between patients with co-morbid disease versus ALS alone (84).

*Other system involvement:* In addition to dementia, other systems can be involved in what otherwise seems to be typical ALS. These include the extrapyramidal motor systems (171-177), supranuclear gaze systems (178-181), and the autonomic nervous system (182, 183). Defects in energy metabolism, including weight loss, hypermetabolism and hyperlipidemia, have increasingly been identified and implicate that other CNS regions such as the hypothalamus are involved, that ALS is part of a systemic disease, or both (reviewed in (132)). Such “atypical” involvement is sometimes referred to as “ALS-plus syndromes” but there is ample clinical, neuropathologic, and neuroimaging evidence to suggest that these should be considered to be part of the neuropathologic spectrum of ALS/MND (184).

**Table 3:** Phenotype Classification Based on Clinically Imputed Anatomy of Neuropathology

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>CNS Anatomical Region Affected as Implicated by Clinical Characteristics</th>
<th>Somatic (Peripheral Body) Region Affected*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>UMN</td>
<td>LMN</td>
</tr>
<tr>
<td>Based on CNS Anatomical Region Affected</td>
<td>ALS</td>
<td>++</td>
</tr>
<tr>
<td></td>
<td>PLS</td>
<td>++++</td>
</tr>
<tr>
<td></td>
<td>PMA</td>
<td>-</td>
</tr>
</tbody>
</table>
e. Clinical focality, stochasticity and spread: How the heterogeneous ALS phenotypes relate to each other is not understood, but neuroanatomic considerations are important (160, 185-189). ALS usually begins in discrete body regions that are randomly (stochastically) located. In these regions, the degree to which UMN and LMN degeneration contributes to the motor deficits (the distribution of disease burden) is variably distributed (141). In this light, PLS and PMA differ especially in distribution of the pathologic burden between UMN and LMN levels; limb variants differ especially in the neuroanatomic location of onset; and FTD and ALS differ especially in the cerebral distribution of pathology. Once triggered, disease spreads or propagates either to proximate neuroanatomical regions or along neuronal networks and thus disease progression is the result of summating motor deficits (141). One recent study identified that up to 14% of second regions involved in disease progression were not contiguous (189). Bifocal or multifocal onset has been proposed (190). Two recent studies using different approaches, one traditional groupings and one unbiased cluster analysis, identified a variety of demographic factors that are significant determinants of phenotype (187, 191).

f. Progression, progression rates and survival: ALS is progressive and fatal and 50% of patients die within 2 years and 90% within 5 years (192). Older age, bulbar onset, and dementia are associated with a poorer prognosis (84, 193, 194). Distinctions between LMN and UMN predominant ALS phenotypes are relevant since prognosis differs for the various syndromes: generally, typical ALS is more rapid than either the PMA or PLS variants (195, 196). The primary cause of death is almost always hypercarbic respiratory failure from involvement of respiratory muscles. A number of advances in symptomatic care, particularly respiratory care, have improved survival, although these have not altered disease progression (65), highlighting an important difference between survival and progression in discussions of natural history. Progression is best thought of in terms of the rate of change and the anatomy of change. The rate of change is largely linear in any one patient, at least during the middle portion of the disease, though there is also evidence for curvilinear progression (42). The rate of change is highly variable between different patients, ranging from malignant (<1 year) to indolent (>8 years). Anatomic variability of disease onset and spread is unrelated to rate of progression, but rate of spread from one region to another is prognostic (188).

Anatomy of disease has a significant impact on survival, since involvement of respiratory muscles is the main cause of death. As many as 2-5% of patients either present with respiratory failure, or have early respiratory involvement, and these patients have shortened survival
although functional impairments may not be severe. Progression rates are determined by the UMN and LMN degeneration and their impact on somatic motor function. Most studies of disease progression measure overall accrual of functional deficits and do not independently separate and measure progression at the UMN and LMN levels. Detailed studies of overall progression by Munsat et al in the 1970’s and 1980’s showed rates of decline in different body regions. Important questions are whether or not progression is similar and additive between the two levels. One reason that PLS and PMA generally have better prognoses than ALS is that disease is primarily only at one level. It is also not clear whether or not LMN involvement but not UMN involvement is a more critical contributor to respiratory failure.

**g. Progression and severity of frontotemporal dementia and ALS:** Progression of cognitive and behavioral symptoms of FTD has been difficult to assess longitudinally due to increasing disability and very high attrition rates at follow up assessment. Cognitive impairment at initial assessment, particularly executive dysfunction, has been associated with higher attrition, faster motor progression, and a more rapid cognitive decline. Furthermore, FTD executive dysfunction or cognitive impairment more generally has been reported as a negative prognostic factor in ALS.

**2. Staging**

The concept of staging in ALS is recent, first described by Roche and colleagues in 2012. Using a database of nearly 1500 patients, they developed a four stage system with hallmarks related to involvement of different regions (bulbar, diaphragmatic, upper extremity, and lower extremity) and need for interventions related to breathing and eating. Stages are defined as follows: Stage 1 – symptom onset; Stage 2A – time of diagnosis; Stage 2B – involvement of the second region; Stage 3 – involvement of the third region; Stage 4A – gastrostomy needed; Stage 4B – noninvasive ventilation needed. Disease course is defined as the time from symptom onset to death or tracheostomy. Stage 2A occurred approximately 35% of the way through disease course, Stage 2B 40%, Stage 3 60%, and Stage 4A or 4B 80%. Individuals are staged based upon their worst milestone (197). Additional studies are needed to determine if this generalizes to other populations.

**D. Genetic predictors of disease**

In the two decades since SOD1 was discovered as the first ALS-related gene, advances in sequencing technologies have substantially improved our understanding of genetic contributions to the disease. Of the nearly three dozen genes associated with ALS, many show sufficient penetrance in that they are recognized in familial forms of the disease and considered definitively causative. These include the most commonly found and best studied genes: C9ORF72, SOD1, TAR DNA binding protein (TARDBP), fused in sarcoma/translated in sarcoma (FUS), profilin 1 (PFN1), optineurin (OPTN), valosin containing protein (VCP), ubiquilin 2 (UBQLN2), and vesicle associated membrane protein (VAPB). Importantly, it has also been appreciated that one in ten ALS patients without a family history (so-called “sporadic” cases) also harbor mutations in these same genes. In total, it is now possible to find a plausibly causative mutation in ~15% of sporadic patients (198, 199). Further details regarding the function and description of each gene are available elsewhere, including on the ALSOD website (http://alsod.iop.kcl.ac.uk/).
While many causative genes have been identified, efforts to find risk factors or genetic factors influencing outcome have proven more challenging. However, it may be prudent to build in methods to genotype current research patients due to the possibility that future studies may render their genetic results more interpretable.

Our expanding understanding of the genetics of ALS has yielded several key insights with direct relevance to natural history and clinical trial design:

First, the genetic heterogeneity of ALS reveals that a diverse set of pathomechanisms underlie motor neuron degeneration. By implication, patient genotypes may be highly relevant for testing a specific mechanism of an intervention, whereas others patients may have different biological causes of disease based on their mutations. Sequence data might be used to enrich for patients with mutations in a specific gene or subset of genes, or to exclude patients with mutations in genes where the targeted pathway is believed to be less relevant for a specific trial. For symptom-based therapies, genotypes may be less important and patients may be selected for trials based on their clinical presentation.

Second, the clinical heterogeneity of ALS with identifiable mutations is very broad, and there are relatively few reliable correlations between genotype and the presentation or natural history of an ALS phenotype. However, several genotype associations are consistent enough to mention:

- Genes vary in the likelihood of concurrent dementia. The simultaneous presentation of FTD and ALS is very common in patients with C9ORF72 mutations, can be present in a subset of patients with TARDPB, FUS, OPTN, UBQLN2, and VCP mutations, and is very rare in those with SOD1 and PFN1 mutations (200-202).

- Most ALS genes are associated with the full range of ALS phenotypes (UMN or LMN predominant forms, bulbar versus limb onset, lower limb versus upper limb onset, early versus late onset). However, SOD1 mutations more commonly begin with symptoms in the lower extremities and are consistently lower motor neuron predominant or exclusive (203). Some analyses have suggested that mutations in TARDBP are more likely to begin with upper extremities (204). Finally, mutations in FUS, particularly those that are de novo, are associated with onset in early adulthood or adolescence (205).

- Several specific mutations have consistent phenotypes. The SOD1 p.A5V (p.A4V in the legacy nomenclature) mutation is the most common in the United States due to a Northeastern founder (206). Its age and site at onset are variable, but the rate of disease progression is uniformly malignant – progression is rapid and the majority of patients die within 18 months (207). In contrast, several other SOD1 mutations (p.G38R, p.D11Y) are characterized by early onset, usually in the lower limbs, and slow progression over decades (208, 209). A recessively inherited SOD1 mutation (p.D91A) is associated with lower extremity onset, ascending involvement, and bowel and bladder involvement (210).

- The predictability of the genotype and corresponding phenotype is a key determinant for whether or not historical controls may be considered in a clinical trial rather than concurrent, randomized controls (See subsection F, Historical Controls, below).
Third, the relative contribution of known genes differs widely between populations and should be taken into consideration. The prevalence of C9ORF72 repeat expansions mirrors the degree of northern European admixture in a population. In Caucasian populations, it accounts for 30-60% of familial ALS and is found in 5-10% of “sporadic” cases (211). C9ORF72 is virtually absent in Asia, where a higher proportion of FUS and TARDBP mutations are seen (211, 212). Founder events make SOD1 p.A5V more common in the United States and p.D91A almost exclusive to Scandinavia (206, 210). In Sardinia, TARDBP p.A382T, G295S) and C9ORF72 are quite common due to the island's history and relative geographic isolation (213).

E. Effect of standard interventions on the natural history of ALS

1. Nutrition management

a. Rationale for nutrition management: A number of studies have found that weight is related to survival in ALS. Both weight loss of greater than 10% (214) and rapid changes in BMI (215, 216) in the time frame from symptom onset to diagnosis are associated with a worse prognosis in ALS. More rapid changes in BMI pre-diagnosis are associated with a lower ALSFRS-R and lower vital capacity at the first clinic visit, but are not associated with bulbar onset disease (216). Patients who have a greater loss in their BMI in the first 2 years they are followed in clinic also have shorter survival (217). Being obese may offer some protection; mild obesity at time of study entry was associated with a better prognosis (218). A prospective nutritional and cancer study of over one half of a million subjects followed for 14-28 years found higher body fat was associated with a lower risk of dying from ALS (219).

b. Difficulties leading to weight loss: Early signs of swallowing problems include taking longer to eat a meal and throat clearing while eating. Patients may spontaneously start to avoid certain foods, realizing they are too difficult to eat. For those patients with early signs of dysphagia, changing the consistency of food including adding thickener to thin liquids or blending tougher foods, cutting food into smaller pieces, eating more slowly, and avoiding talking while eating may all be suggested. A speech therapist may determine that chin tucking or head turning may assist in swallowing for a given patient. As swallowing worsens or if weight cannot be maintained, patients may start adding nutritional supplements to their food intake (107, 128, 129). Constipation is a common side effect for many ALS patients, due to decreased movement, dehydration, and weakness of the diaphragm (220).

c. Use of GT in nutrition management: In the absence of evidence based medicine to determine when it is best to obtain a gastrostomy tube (GT), timing may be driven by a variety of factors, including weight loss of more than 10%, aspiration, distress over choking, long meal times, dehydration, and occasionally for practical reasons related to upper extremity weakness that make it impossible to eat without assistance (107, 221, 222). In addition, GT placement has been suggested when the FVC reaches 50% even if the patient has no trouble with swallowing because of concerns with increased risk of the procedure when breathing is compromised (107, 223). However, recent studies have suggested patients can safely get GT even with a low FVC (224-226), and using NIV during the procedure may prove helpful (225, 227).

Survival benefits of enteral nutrition: Despite the data demonstrating that weight loss is associated with worse prognosis, the evidence is mixed regarding survival benefits of enteral
nutrition via GT in ALS patients. Survival benefits were found by some (48, 228, 229) while others found no benefit (230-233). None of the trials were randomized controlled trials, and controls varied between studies and included both patients who refused GT and those who did not need it, which may have contributed to the variability in results (107, 234). For patients receiving a GT, bulbar onset (229), abnormal overnight pulse oximetry prior to procedure (235), older age and more than a 10% weight loss were all associated with shorter survival compared to those without these attributes (236). The impact of weight loss raises the possibility that failure to consistently demonstrate benefit may in part relate to patients being malnourished at the time the procedure is performed. It remains to be seen if recommending a GT, for example, when 5% of weight loss has occurred would lead to better outcomes. Alternatively, rapid weight loss could simply be a marker for more aggressive disease. Another challenge with enteral nutrition is that noncompliance with the nutritionist’s recommendations is not uncommon (237); patients who tolerated high caloric intake following GT had prolonged survival compared to those who did not (238).

**Decision to obtain GT:** In considering a GT, quality of life, providing a way to receive medications, and what subsequently happens to weight may also play a role in the patient’s decision. Enteral nutrition resulted in stabilization and or improvement in the BMI or weight following gastrostomy placement (223, 232, 239, 240). There are very limited studies on how quality of life changes following a GT and results are mixed (241-243), which led the AAN to conclude that there was insufficient evidence to determine the impact of GT on quality of life (223, 232, 239-243).

**GT procedures:** GT can be inserted surgically, radiographically, via an endoscope, and in a procedure that is a hybrid of the latter two (per-oral image guided gastrostomy or PIG); few patients with ALS have it done surgically (236). Some studies have found radiographically inserted gastrostomy tubes (RIG) may confer a survival benefit (244), be associated with fewer insertional failures and post procedure episodes of aspiration (245), and less pain (246) compared to percutaneous endoscopically placed gastrostomy tubes (PEGs). However, a recently published study compared patients who received RIGs, PEGs, and PIGs and found no difference in outcome (236) based upon the procedure used, and a study comparing PEG and RIG also found no survival difference (247).

**d. Nutrition management studies:** Studies investigating different diets in ALS have focused on food content as well as on calories. Limitations include both short duration and small subject numbers. Findings are described in Table 4. Results have varied, and no conclusions about specific diets to recommend can be made based upon available data.

<table>
<thead>
<tr>
<th>Table 4: Diet and ALS</th>
<th>Author</th>
<th>Diet</th>
<th>N=</th>
<th>Duration (mos)</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stanich et al, 2002 (129)</td>
<td>High protein</td>
<td>20</td>
<td>6</td>
<td>No change in disease progression</td>
<td></td>
</tr>
<tr>
<td>Silva et al, 2010 (248)</td>
<td>High protein</td>
<td>16</td>
<td>4</td>
<td>ALSFRS-R stabilized</td>
<td></td>
</tr>
<tr>
<td>Dorst et al, 2013 (249)</td>
<td>High fat, high calorie</td>
<td>22</td>
<td>3</td>
<td>Weight stabilized with high fat</td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Intervention</td>
<td>Control</td>
<td>ALSFRS</td>
<td>Outcome</td>
<td></td>
</tr>
<tr>
<td>------------------------------</td>
<td>----------------------------------</td>
<td>---------</td>
<td>--------</td>
<td>----------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Wills et al, 2014 (250)</td>
<td>High fat, high calorie</td>
<td>Control</td>
<td>8</td>
<td>Fewer adverse events</td>
<td></td>
</tr>
<tr>
<td></td>
<td>High carbohydrate, high calorie</td>
<td></td>
<td>9</td>
<td>Dropouts and deaths in the high carbohydrate, high calorie</td>
<td></td>
</tr>
</tbody>
</table>

2. Physical and occupational therapy

a. Rationale for physical and occupational therapy: Multidisciplinary clinics, which are the standard in ALS care, usually include occupational, physical and speech therapists as part of the team. Patients who attend such clinics are more likely to receive adaptive devices than those who do not get care in specialized centers (251). Early in the disease, emphasis may be placed on maintaining mobility by using braces or a cane, and permitting independence through adaptive equipment such as plump utensils and zipper pulls (252, 253).

b. Exercise and ALS: Limited information is available on the effect of exercise in ALS, in part because ever-changing strength and fatigue make developing an exercise program problematic (253). However, early in the disease, exercise may produce modest benefits and does not appear to be harmful (254, 255); when stretching is added to exercise, spasticity in particular may improve at least for the short term (255).

c. Forms of effective physical therapy: As the disease progresses and the patient becomes more dependent on a caregiver for activities of daily living, therapists may evaluate the home for modifications, train family members on transfers, and instruct on range of motion exercises to prevent contractures and reduce pain (252, 253, 256). Massage may be of benefit for reducing pain (257). Gait belts, pivot discs, and mechanical lifts may improve safety of transfers (252, 253). Elevated toilet seats or grab bars, shower seats and transfer boards were viewed as particularly helpful by the patients (258). When weakness progresses to the point that ambulation is either unsafe or no longer possible, a power wheelchair may be prescribed, with adaptations that may include items such as a seat that allows for variable heights to make transfers safer, elevated leg rests and reclining backs for repositioning, head support and head controls (252, 259). The majority of patients were satisfied with the comfort and ease of their power wheelchair and a minority felt that they had waited too long to get one (260).

d. Therapies to improve communication: With modest changes in voice, the speech therapist may focus on strategies to improve intelligibility (261). With worsening dysarthria, augmentative and alternative communication (AAC) devices come into play and range from low tech pen and paper, eye blinks and communication boards to elaborate computers that not only generate speech but also allow for sending emails and controlling the room environment (261). If hand weakness precludes typing, then eye gaze mechanisms coupled with the AAC may be useful (262). Dysarthric patients who receive AACs report stable to improved quality of life (263), although patients with cognitive dysfunction may not be able to learn how to use the more
elaborate systems (261). Although there is hope that brain computer interfaces may allow patients with marked weakness to communicate and interact with their environment, this is not yet a reality (264).

3. Respiratory management

a. Rationale for respiratory management: The majority of patients with ALS die from neuromuscular respiratory failure. Consequently, providing bi-level noninvasive ventilator (NIV) support was recognized as a treatment that offered survival benefits in the late 1990’s (265, 266). There is no consensus as to the optimal time point to initiate NIV support, and there is variability across countries and even within the same country as to whether the decision should be driven by symptoms, respiratory measurements, or both (105, 221). In the United States, Medicare regulations play a role in dictating when NIV is instituted, since they will pay for the necessary equipment if the FVC is below 50% or the maximum inspiratory pressure (MIP) is less than -60 cm H2O. Alternatively, it is covered if there is desaturation on nocturnal pulse oximetry less than or equal to 88% for 5 continuous minutes, or if the ABG reveals carbon dioxide to be greater than 45 mm Hg (267). In Canada, symptoms of orthopnea, dyspnea, and morning headaches are the most common reasons NIV is suggested (268) while in Italy PaCO2 greater than 45 mm Hg was the most common trigger to recommend NIV (105).

b. NIV survival benefits: A number of studies using different parameters for NIV initiation have demonstrated survival benefit in ALS patients. In the only randomized controlled trial, the rate of decline in FVC was slowed (-2.2%/month before NIV compared to -1.1%/month after) when used more than 4 hours at night and median survival was increased by 205 days. In this trial, NIV was initiated when patients had a MIP of less than -60 cm H2O with orthopnea, or had symptomatic hypercapnea (269). In an open label trial, survival improved by 15 months for those who tolerated NIV a minimum of 4 hours at night after starting when FVC was below 50%, and the rate of change in the FVC also slowed (265). While universally studies have found survival benefits in patients who are tolerant of NIV, not all have found that its use impacts the rate of decline in FVC (270). Despite the number of studies supporting survival benefits from using NIV, utilizing information from the ALS CARE database found that of those patients with an FVC of less than 50%, only 36% of the patients were using NIV (271). There are likely multiple reasons, including NIV not offered by the treating physician, or the patient either refusing it or being intolerant of the device.

c. Timing of NIV initiation: Starting NIV early has been proposed as a way to potentially improve tolerance and perhaps confer a larger benefit (267). When NIV was initiated in patients with an FVC of <75% who also had overnight desaturations on pulse oximetry, survival was improved at one year compared to those who refused or were intolerant (272), though using intolerant patients as a control may be problematic, as intolerance may be associated with other poor prognostic features such as bulbar dysfunction. Another study comparing initiation of NIV in patients with an FVC above or below 65% of predicted capacity also found patients who began the study with a better FVC lived approximately a year longer (273). Survival benefits from NIV may be lost in patients with prominent bulbar involvement although they may still have improvement in sleep quality (269). One month following NIV initiation, patients whose overnight pulse oximetry revealed they spent less than 5% of the time with SpO2 below 90% had improved survival compared to those who spent more than 5% of the time with such oxygen desaturations (274). Patients using NIV who also had a higher BMI had
improved survival, so nutritional status may also be important if benefits from NIV use are to be realized (275). Using NIV in patients with poor diaphragmatic function resulted in a decrease in resting energy expenditure along with eliminating accessory neck muscle activity during inspiration (276).

d. NIV compliance/tolerance: Compliance with NIV is reduced in patients with prominent bulbar involvement (275, 277-279) as well as in patients with frontotemporal dysfunction (84). Patients with orthopnea (278) and dyspnea (271) are more likely to be compliant, as are those with better upper extremity function (117). Use of other interventions, such as gastrostomy, riluzole, and augmentative language devices are also associated with improved NIV compliance (271), and earlier NIV initiation can also result in improved compliance (117, 277, 280). Improved tolerance to NIV is associated with less airway secretions at the time NIV is begun (279). No clear-cut relationship between age and NIV tolerance has been found (117, 271, 279). However, with technological advancements that include machines whose settings can be altered remotely, equipment that makes adjustments in pressure delivered based upon the tidal volumes being drawn by the patient, batteries that have an extended back up, and mask interfaces that are softer, come in a variety of sizes, and include nasal pillows for those who are claustrophobic, compliance may be better than reports from early studies that found only one third of patients tolerating NIV for 4 hours or more at night (265). A more recent study found tolerance was achieved by 75% of patients although this was done in Europe, where patients are admitted to the hospital for NIV initiation, while this is almost always done at home in the U.S. (279). Recently, besides pressure mode NIV, volume mode NIV has also been used in ALS patients. While there was no difference in survival when comparing patients who received the different modalities, more effective ventilation was achieved with the volume mode devices (281).

e. Advantages of NIV: Improvement in quality of life following NIV use has included benefits regarding dyspnea, sleep quality, fatigue, mood, energy, concentration and daytime somnolence (117, 269, 282, 283). However, overall quality of life may not improve following NIV initiation perhaps related to declines in other areas secondary to ALS progression (241).

f. Alternatives and additions to NIV usage: For patients with an impaired cough and trouble clearing secretions, a mechanical insufflator-exsufflator (MeIE), which simulates a cough, may be of benefit particularly in the setting of an acute infection, though patients with severe bulbar dysfunction may not benefit because of airway collapse during the exsufflation cycle (284, 285). Breath stacking, abdominal thrusts, and use of NIV may also improve peak cough flow and thus promote secretion clearance from airways (286). A device that delivers high-frequency chest wall oscillation (HFCWO) has been used in patients with cystic fibrosis and other pulmonary diseases for airway clearance, but only recently in neuromuscular diseases (287). In a small study, HFCWO helped with breathlessness but did not impact FVC, peak expiratory flow, dyspnea, or oxygen saturation (288). Another small study comparing patients who used HFCWO to those who did not also failed to show benefit regarding decline in FVC and did not improve survival (289). Suction machines are frequently used by patients with ALS to help handle secretions, though no trial has demonstrated their benefit (107). Additionally, diaphragmatic pacing has been shown to allow some patients with high cervical cord injuries to be weaned from a ventilator (290), and an open label trial in ALS using historical controls found it was safe and slowed decline in FVC (291). However, a placebo-controlled trial in Europe was
recently terminated prematurely because patients who were paced along with using NIV had a shorter survival compared to those who used NIV alone (224).

g. TIV as alternative to NIV: When NIV no longer offers sufficient neuromuscular respiratory support, tracheostomy invasive ventilation (TIV) may be considered. The percentage of patients who pursue TIV varies widely between countries, and is estimated to be 1.5% in Canada (268), under 10% in the United States (292), 10-32% in Italy (293, 294), and up to 33% in Japan (292). This may in part be related to how the treating physicians approach the discussion of TIV and whether or not they appear to support or oppose it (295). It is done emergently rather than as a planned procedure 52-67% of the time and in those circumstances may be done without the patient actively involved in the decision making process (281, 296, 297). Younger age, shorter disease duration, and being married are features associated with a higher likelihood of choosing TIV (292).

h. Effects of TIV usage: Challenges faced after TIV implementation include worsening cognition, limited ability to communicate, infections, pressure sores, and the need for indwelling catheters and sedating medications (293). For patients using TIV for more than 5 years, 33% had minimal ability to communicate and 18% became locked in, meaning the patient was able to function cognitively but was only able to communicate with their eyes or, for some, unable to communicate at all due to complete loss of movement (298). There have been mixed reports for how patients with TIV view their quality of life, although the burden of care is consistently reported to increase for the caregiver (297, 299-302). Twenty percent of patients undergoing TIV do not survive the initial hospital stay (294, 303). Survival after TIV varies widely, ranging from 8 to 49 months (294, 302-305), and those over the age of 60 do not have a clearly meaningful survival extension by TIV (302, 304). Following TIV, 27-31% of patients do not return home but rather are discharged to a skilled nursing facility (294, 303), and patients with NIV are less likely to reside in a skilled nursing facility than those with TIV (297). Patients with TIV living at home survive longer than those who reside in a skilled nursing facility – 43 months compared to 2 months in one study (303). With improvements in the machines and interfaces that deliver NIV, along with better secretion management, more patients are able to use NIV up to 24 hours per day and this may decrease the number of patients opting for TIV over time (297).

4. Psychological support and end-of-life care

a. Importance of end-of-life care: ALS places heavy physical, psychological, and financial burdens on patients and caregivers. Patients and caregivers need support and counseling from the time of diagnosis through the end-of-life (107, 306). Multidisciplinary care can improve both quality and length of life of patients with ALS (107, 221, 307). The care team has an important role in providing education to patients and caregivers to help guide medical decisions, including gastrostomy insertion, non-invasive ventilation, and life-sustaining interventions such as tracheostomy with mechanical ventilation (308). It is recommended that the care team discuss the patient’s preferences for aggressive or palliative care early in the course of disease, particularly since the quality of early care affects the quality of care at the end of life (293, 309). A palliative care approach, focused on relief of symptoms and maximizing the quality of life, can be initiated as soon as the diagnosis of ALS is made (221). Late referral to palliative care was found to have a negative effect on the terminal quality of life (310). Discussions of advance directives and hospice care, which refers to palliative care at the end of life, are best raised early on, rather than at the time of rapid decline (311). The hospice care team seeks to remove
obstacles to a peaceful and dignified death, and to provide psychological and spiritual support to the ALS patient, and emotional support to the family at bereavement.

b. Types of palliative therapies: Many of the physical symptoms near the end of life are amenable to medical management including pain, sialorrhea, and respiratory distress (220, 312). Early introduction of assisted communication devices may alleviate patients’ fear of an inability to communicate and can improve quality of life (313). Psychological symptoms include depression, anxiety, fear, and grief. Estimates of the prevalence of depression vary across studies, but may be lower in ALS patients than in the general population (314, 315); however, depression is associated with greater suffering and a poorer quality of life (315, 316). Expressing a “wish to die” may occur in ALS patients who are not clinically depressed (315). Depression is experienced by many caregivers, particularly in patients with mechanical ventilation with its increased caregiver burden and fatigue (317, 318). One small study showed a benefit of cognitive behavioral therapy for patients and caregivers (319), but best practices for providing psychological support to patients and caregivers remain to be determined (320). Most ALS patients professed a wish to die at home (321). The end of life often occurs suddenly; indicators include a rapid physical decline, cognitive impairment, and infection (306, 322). Palliative sedation may be appropriate for severe dyspnea (323). Mechanical ventilation, which is used in less than 10% of ALS patients in the US, can result in transfer to a care facility (295) as well as a decline in quality of life, as perceived by surviving caregivers (293). ALS places heavy emotional and financial demands on family and caregivers. Members of the multidisciplinary team play a role in providing information regarding social services and community resources, advance directives and the naming of a healthcare proxy.

c. Relevance to clinical trial participation: Progressive disability near the end of life in ALS affects patients’ participation in clinical trials, for example leading to early drop-out or exclusion. Methods for remote assessment of patients, by telemedicine or home devices to generate data would extend the ability of patients to be included in clinical trials.
V. DIAGNOSIS

A. General comments

1. Who makes a diagnosis of ALS?

Due to the absence of any diagnostic tests or biomarkers that can confirm ALS, the diagnosis relies heavily on the history and physical examination and the exclusion of other disorders by imaging, laboratory, or electrophysiological measures. While a general neurologist may diagnose ALS, a second opinion by a neuromuscular specialist is often requested given the gravity of the diagnosis and complexity of the management.

There are formal criteria for a diagnosis of ALS (the El Escorial Criteria discussed below). However, practically speaking, patients often come to the attention of physicians with symptoms and signs that can be confused with other disorders. In light of this, an early diagnosis of ALS is often dependent on clinical acumen and extensive electrodiagnostic studies.

2. Neurological history and examination

Key features in the history related to LMN degeneration may include symptoms of weakness, atrophy, and muscle cramps. The atrophy may be so profound that a patient may present with a chief complaint of weight loss, prompting a workup for malignancy. At the time of initial presentation, the patient may have evidence of a monoparesis, paraparesis, hemiparesis, or quadriplegia. Some patients may present with weakness limited to the bulbar muscles resulting in complaints of slurred speech and difficulty swallowing. Symptoms of UMN degeneration may include loss of dexterity, slowed movements, stiffness, and emotional lability. Important historical features that may suggest other etiologies include prominent sensory features, pain, visual disturbances, abnormal movements, bladder or bowel incontinence, or autonomic dysfunction. Historically, cognitive function was thought to be spared. However, subsequent studies and genetic discoveries clearly demonstrate that cognitive and behavioral dysfunction may occur.

On examination, signs of LMN dysfunction may include muscle atrophy and weakness, which is often asymmetrical and can be accompanied by fasciculations. Signs of UMN dysfunction may include spasticity, hyper-reflexia, loss of abdominal reflexes and pathologic reflexes (jaw jerk, extensor plantar response or Hoffman’s sign).

B. Barriers to early diagnosis

Total diagnostic time, defined as the time from symptom onset to confirmed diagnosis, ranges from 8 to 15 months in ALS (232, 324-328). This delay in reaching a diagnosis of ALS is problematic for several reasons: It represents a significant proportion of total disease duration, it prolongs a period of uncertainty that adds to patients’ stress, and it often leads to unnecessary, costly, and sometimes painful diagnostic tests and procedures. Diagnostic delays also represent missed opportunities to begin treatment with riluzole, address ALS-related symptoms, and allow for early entry in clinical research trials. Moreover, in patients diagnosed late, after marked neuronal loss has already occurred, even when disease-modifying treatments become available they will likely have limited effectiveness.
1. Impact of diagnostic milestones on delayed diagnosis

The causes of diagnostic delay in ALS are likely multifactorial. Non-modifiable predictors of prolonged diagnostic timelines include age >60 years at onset, sporadic vs. familial disease, and limb vs. bulbar onset (324, 328, 329). Other factors that contribute to delays in diagnosis are potentially subject to improvement, including three interim diagnostic milestones: (A) time from presenting symptom to first doctor visit, (B) time from first doctor visit to suspected ALS diagnosis, and (C) time from suspected to confirmed ALS diagnosis (325, 327, 328).

(A) Time from symptom onset to first doctor visit is negatively impacted by several factors. The onset of ALS is usually slow, insidious, and painless. Patients may wait to seek medical attention until symptoms become more noticeable or cause a functional limitation. The impact of increased public awareness about ALS regarding the time to medical evaluation has been largely unexplored.

(B) Time from first doctor visit to suspected diagnosis: this diagnostic milestone is an area of potential high impact. A major barrier to diagnosis is the fact that ALS is a rare disease with multiple possible presenting symptoms that may mimic other, more common diagnoses such as neuropathy, spinal disease, and carpal tunnel syndrome. General practitioners not familiar with ALS may want to rule out other “benign” diseases before exploring the more uncomfortable possibility of ALS. Thus, it is perhaps not surprising that 27-61% of the ALS patient population is initially misdiagnosed (324, 328-330). Fasciculations are probably the ALS symptom most commonly associated with the disease and their presence at onset is associated with a faster time to diagnosis (328, 329). Increased awareness about the implication of ALS symptoms among general practitioners could potentially result in faster diagnostic timelines. A “red flag” system to alert community physicians about the possibility of ALS when certain symptoms occur together is being investigated. Even with increased awareness, however, ALS remains a clinical diagnosis with no definite diagnostic biomarkers. Laboratory, radiological, and electrophysiological investigations performed to exclude disease-mimics are time-consuming, adding to the delay in diagnosis.

(C) Time from suspected diagnosis to confirmed diagnosis. Even when the diagnosis is suspected, confirmation is generally deferred to dedicated ALS centers, in part because of reluctance by non-ALS specialists to deliver such a severe diagnosis. As a result, patients see an average of three physicians before reaching the final diagnosis (328). We speculate that faster referral to dedicated ALS centers may shorten the ALS diagnostic timeline (325). However, ALS clinics are usually located only in academic settings or large hospitals, which may present additional geographical and financial barriers. Research on the impact of logistical factors on diagnostic delays (e.g., rural vs. urban setting and insurance coverage) is warranted to optimize referral patterns and prioritize resource allocation.

2. Financial barriers to diagnosis

A diagnosis of ALS typically requires multiple doctor visits and various forms of diagnostic testing. Distance to an ALS specialist and inadequate insurance coverage can
compound the financial difficulties related to ALS diagnosis and care. Additionally, potential lost work hours may cause patients to delay testing and doctor visits.

3. Heterogeneity of ALS Disease Presentation and Progression

The heterogeneity of ALS presents additional diagnostic challenges, adds to the complexity of trial design, and complicates randomized trial enrollment and analyses. Classical ALS is defined as a mixed UMN and LMN disorder, although initial manifestations vary among patients with regard to the anatomic region of onset (bulbar, cervical, thoracic, lumbosacral), degree of UMN and LMN involvement, rate of progression, and features such as cognitive dysfunction and pseudobulbar affect. The diagnosis of ALS may also be applied to incomplete presentations with only LMN or UMN signs, or solely bulbar features, an approach validated by post mortem and genetic data suggesting that patients with these seemingly distinct syndromes and classical ALS may represent the range of phenotypes that constitute the spectrum of ALS, as described in Section IV. Natural History (195, 331, 332). MND with exclusively LMN features is classified as PMA (45). Generalized, purely UMN disease is termed PLS, but LMN signs, to a limited degree, have been accepted by some authors in defining PLS, and the diagnostic criteria for PLS are less well established than those of ALS (149, 151, 333). PBP is an UMN and/or LMN disorder restricted to the bulbar region. Most patients presenting with these syndromes eventually develop the full clinical picture of ALS, but ~10% of adult MND patients retain the diagnosis of PMA, PLS, or PBP (332). The diagnostic challenge is exemplified by patients clinically diagnosed with PMA who at autopsy demonstrate UMN pathology, those with features of PLS in whom post mortem examination reveals LMN pathology, and patients with pure LMN or UMN syndromes who carry a major ALS risk gene variant (196).

Survival in ALS typically ranges from 2-5 years but may be as short as a few months; about 10% of patients survive longer than 5 years. Older age of onset tends to be associated with a poorer prognosis (193). PBP, which typically progresses to the full clinical picture of ALS, also tends to be associated with a poorer prognosis than ALS with limb onset (193). Clinically significant cognitive impairment is an additional clinical feature associated with shorter survival compared to cognitively normal patients (84, 194). Distinctions between LMN and UMN predominant ALS phenotypes are relevant prognosis differs for the various syndromes (195, 196). (82, 83, 85, 154, 170, 195, 334-337)

4. Frontotemporal Dementia in ALS

Up to 50% of patients with ALS develop features of frontotemporal cognitive impairment, meeting criteria for frontotemporal dementia in up to 15% (82, 83, 170, 195, 334, 335). Deficits include impaired executive functions, disinhibition, impulsivity, apathy, affective symptoms, compulsivity, and irritability (85, 195, 336, 337).

Cognitive, behavioral and language changes can occur before, during or following the development of motor neuron disease (154, 195). Quantitative difficulty in phonemic fluency has been reported early in disease. Behavioral changes may occur prior to the detection of cognitive change and may be difficult to discriminate from an exaggeration of long-standing personality traits or motor features of the disease.

Early detection, even in the presymptomatic phase of the disease, is one of several important targets for treatment trials since effective treatment could potentially prevent the
disease from occurring. Presymptomatic studies by definition depend on non-clinical ascertainment with biomarkers, and genetic testing in individuals who are members of families with ALS and/or FTD can be performed to ascertain future risk for clinical features of ALS and/or FTD. The presence of the C9orf72 repeat expansion has been associated with increased cognitive and behavioral dysfunction, although this does not account for the majority of patients with such cognitive and behavioral changes.

Clinical Diagnosis of FTD. The gold standard for diagnosis is neuropsychological assessment. This requires qualitative assessment of cognitive functions using measures which control for motor disability, and behavioral assessment based on patient interview and informant reports. Diagnoses across the frontotemporal spectrum are defined by the current consensus criteria (170), although revised criteria are expected in 2016.

ALS cognitive impairment (ALSci) reflects deficits that do not meet criteria for dementia. ALSci is defined by impairment on two distinct measures of executive functioning, characterized by scores at or below the 5th percentile compared to age and education-matched norms. Consensus criteria authors and associated experts uniformly agree that language dysfunction is also prevalent and distinguishable from executive dysfunction in ALS patients, despite the absence of this specification in the existing diagnostic criteria. The revised criteria will include diagnosis on the basis of language impairment. It is therefore recommended that ALSci should be diagnosed when language impairment is demonstrated on two or more distinct language measures (such as syntactic processing or receptive language).

ALS behavioral impairment (ALSbi) is a sub-clinical syndrome defined by the presence of two or more behavioral abnormalities that are not better explained by premorbid personality or a psychological condition. These symptoms can include apathy, disinhibition, loss of empathy, and other characteristics seen in behavioral variant FTD. Evidence for these behavioral symptoms should be obtained from at least two sources, such as patient observation and caregiver report on structured interview. Changes in social cognition impact behavior and are seen in ALS patients, although the criteria for diagnosing social cognition deficits are not well articulated for this population.

ALS-FTD (bvFTD, PNFA, SD) is diagnosed using the Neary criteria (338), although the revised criteria may include the Rascovksy criteria (339) for behavioral variant FTD (bvFTD). BvFTD is characterized by a marked change in social comportment and behavior. Progressive Nonfluent Aphasia (PNFA) is characterized by motor speech deficits, and Semantic Dementia (SD) is characterized primarily by loss of semantic knowledge.

Specific diagnostic tools are recommended in the consensus criteria (170), and language-specific measures are also referenced in recent publications (340, 341). Additionally, the NINDS CDE website provides a review of ALS-specific measures.

Quantitative assessment by a neuropsychologist using tools designed for ALS to accommodate for motor disability is encouraged in all individuals with ALS to ascertain early features of cognitive difficulty. Separate interview of an informant is also essential to determine behavior and personality change.
Use of additional measures cannot substitute for neuropsychological diagnosis, but neuroimaging may increase diagnostic specificity. Use of MRI/CT could delineate frontal lobe atrophy (bvFTD), left perisylvian atrophy (PNFA) or anterior temporal lobe atrophy (SD). Use of PET/SPECT could provide evidence of frontal and/or anterior temporal lobe hypometabolism/hypoperfusion in some cases. Use of neuroimaging in trials offers the ability to compare ALS data to FTD datasets. However, most ALS patients cannot tolerate this procedure except in early-stage disease due to respiratory complications and/or not being able to lie supine. Pathological markers could include histopathology on biopsy or post-mortem, or detection of known genetic mutations. Pathological heterogeneity in ALS suggests that use of these markers may not improve diagnosis but could further characterize subgroups.

C. Diagnostic Criteria

1. The El Escorial Criteria-Revised

The formal diagnosis of ALS has been outlined by the World Federation of Neurology Research Group on Motor Neuron Diseases and referred to as the El Escorial Criteria for the diagnosis of ALS. The El Escorial criteria were developed to generate a common understanding of diagnostic procedures for ALS and was originally outlined in 1994 (342) and then revised in 2000 (343). The revision allowed for the use of EMG/nerve conduction studies to detect denervation that could not be observed clinically to aid in the diagnosis of ALS. These criteria are used much more to determine the inclusion of ALS patients in clinical studies than they are used in the general community for determining a diagnosis. While these criteria continue to be the current standard for a formal ALS diagnosis, there are some shortcomings that were recently recognized and detailed by Agosta and colleagues and addressed in the 2015 revisions discussed below (344).

After a 15-year span in which the El Escorial criteria were employed as defined by Brooks et al in 2000, the El Escorial Criteria were once again revised in 2015 by Ludolph and colleagues on behalf of the WFN Research Group on ALS/MND (345). These revisions in the criteria were aimed at recognizing restricted phenotypes of ALS including progressive bulbar palsy, flail-arm and flail leg syndromes, progressive muscular atrophy, and primary lateral sclerosis. Importantly, primary lateral sclerosis can be included as ALS if and when clinical or electrophysiological evidence of involvement of the lower motor neuron in at least one limb or body region is present.

These criteria also recognize that a diagnosis of ALS can be made if the former criteria for ‘possible ALS’ are fulfilled. Given the recognition that cognitive dysfunction is present in a substantial percentage of ALS patients, the presence of cognitive dysfunction should not be considered exclusionary to a diagnosis of ALS, as has been the case in the past. There is also the recognition that deficits in sensory, oculomotor systems and sphincter disturbances can be features of ALS even though they may be uncommon. As the identification of causative ALS genes is now being increasingly recognized, the finding of a pathogenic mutation in a known gene can substitute for either lower or upper motor neuron signs, so that diagnosis of ALS can be made on the basis of UMN or LMN signs in one body region, associated with a positive genetic test.
With respect to specific signs at the time of diagnosis, the diagnosis of ALS requires, at minimum, one of the following:

- Progressive UMN and LMN deficits in at least one limb or region of the human body; i.e. meeting the revised El Escorial criteria for “possible ALS”.
- LMN deficits as defined by clinical examination (one region) and/or by EMG in two body regions (defined as bulbar, cervical, thoracic, lumbosacral). The EMG findings consist of neurogenic potentials and fibrillation potentials and/or sharp waves.

The diagnosis of ALS is made possible by:

- History, physical and appropriate neurological examinations to ascertain clinical findings, which may suggest ALS.
- Electrophysiological examinations to ascertain findings that confirm LMN degeneration in clinically involved regions, identify LMN degeneration in clinically uninvolved regions, and exclude other disorders.
- Neuroimaging examinations to ascertain findings that may exclude other disease processes.
- Clinical laboratory examinations, determined by clinical judgment, to ascertain possible ALS-related syndromes.
- Neuropathologic examinations, where appropriate, to ascertain findings that may confirm or exclude sporadic ALS, coexistent sporadic ALS, ALS-related syndromes or ALS variants.
- Repetition of clinical and electrophysiological examinations at least six months apart to ascertain evidence of progression.

2. Awaji Diagnostic Algorithm

The International Federation of Clinical Neurophysiology (IFCN) produced the Awaji diagnostic algorithm, which utilized electrophysiological measures in combination with clinical measures to increase the sensitivity of an ALS diagnosis. This algorithm is particularly relevant for evaluating bulbar-onset ALS (346). Currently, these criteria are not widely used for enrolling ALS patients into clinical trials although these may provide insights into potential alternative or supplemental strategies for diagnosing ALS.

Proposed Guidance (taken from Ludolph and colleagues) (345)

1. A diagnosis of ALS can be made if the former criteria for ‘possible ALS’ are fulfilled.
2. More widespread LMN disease (i.e. two or more body regions) in the absence of UMN signs or any other explanation for the LMN clinical signs is sufficient for the diagnosis of ALS.

The decision whether to include all or a subset of ALS phenotypes in a particular clinical trial will depend on the individual trial design and therapeutic target. However, these criteria for an ALS diagnosis offer an opportunity to be more inclusive of phenotypes which were previously thought too restrictive to the diagnosis of ALS.
D. Diagnostic Laboratory Investigations/Methods for Confirming Diagnosis

1. Blood and urine tests

There are currently no laboratory tests that confirm a diagnosis of ALS, although this is an active area of investigation. Results from most laboratory tests are generally within normal ranges for ALS, except for an elevation in creatine kinase. They are usually performed as clinically indicated primarily to rule out diseases that mimic ALS (Table 4). The studies in table 4 should be used as a guideline. It is not necessary that all of these studies be performed in order to exclude all of these disorders if not clinically indicated.

<table>
<thead>
<tr>
<th>Lower Motor Neuron Predominant</th>
<th>Multifocal Motor Neuropathy</th>
<th>GM1 antibodies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spinal Muscular Atrophy</td>
<td>Genetic testing/muscle biopsy</td>
<td></td>
</tr>
<tr>
<td>Hypothyroidism</td>
<td>Thyroid-stimulating hormone, free T4</td>
<td></td>
</tr>
<tr>
<td>Hyperparathyroidism</td>
<td>Intact parathyroid hormone; ionized calcium</td>
<td></td>
</tr>
<tr>
<td>Paraneoplastic</td>
<td>Antibody panel (serum, CSF)</td>
<td></td>
</tr>
<tr>
<td>Inclusion Body Myositis</td>
<td>Muscle Biopsy</td>
<td></td>
</tr>
<tr>
<td>Lyme Disease</td>
<td>Serology (serum, CSF)</td>
<td></td>
</tr>
<tr>
<td>West Nile Virus</td>
<td>Serology (serum, CSF)</td>
<td></td>
</tr>
<tr>
<td>Polio</td>
<td>Serology (serum, CSF), stool antigen</td>
<td></td>
</tr>
<tr>
<td>Myasthenia Gravis</td>
<td>Acetylcholine receptor and MuSK antibodies</td>
<td></td>
</tr>
<tr>
<td>Syphilis</td>
<td>RPR or CSF VDRL</td>
<td></td>
</tr>
<tr>
<td>Hexosaminidase A Deficiency</td>
<td>Hexosaminidase A level</td>
<td></td>
</tr>
<tr>
<td>Heavy Metal Intoxication</td>
<td>24hr Urine Collection for Heavy Metals</td>
<td></td>
</tr>
<tr>
<td>Lymphoma</td>
<td>Serum protein electrophoresis</td>
<td></td>
</tr>
<tr>
<td>Upper Motor Neuron Predominant</td>
<td>Copper Deficiency</td>
<td>Serum Copper and Zinc</td>
</tr>
<tr>
<td>B12 Deficiency</td>
<td>Vitamin B12, methylmalonic acid, homocysteine</td>
<td></td>
</tr>
<tr>
<td>Human Immunodeficiency Virus</td>
<td>HIV</td>
<td></td>
</tr>
<tr>
<td>Tropical Spastic Paraparesis</td>
<td>Human T-cell lymphotrophic virus 1</td>
<td></td>
</tr>
</tbody>
</table>
2. CSF studies

As with blood and urine laboratory tests, there are currently no cerebrospinal fluid (CSF) diagnostic markers that confirm an ALS diagnosis, although this is an area of active research and. CSF analysis is not usually performed except in cases where the patient presents at a young age of onset and/or in order to rule out infectious, inflammatory, neoplastic or paraneoplastic disorders. The CSF is usually normal in ALS, although 33% of patients with pathologically confirmed ALS have elevated CSF protein (>45 mg/dL) (1) and values in excess of 75mg/dL are occasional encountered (347, 348). Rarely patients with ALS have a mild pleocytosis or oligoclonal bands. None of these non-specific findings are diagnostic for ALS.

3. Electrophysiological Studies

Neurophysiological testing provides important data to support a diagnosis of ALS and exclude ALS mimics. In particular, electromyography aids in demonstrating LMN involvement in ALS and is central to published diagnostic guidelines, the El Escorial criteria, recently revised as discussed below (342, 343, 346, 349). The Awaji electrodiagnostic criteria for diagnosis of ALS addresses recognized limitations in the electrodiagnostic guidelines of the El Escorial criteria, which require signs of ongoing and chronic partial reinnervation in order to be considered sufficient to support the diagnosis of ALS (350). In the appropriate clinical context, the Awaji guidelines consider fasciculation potentials to be equivalent to fibrillation and positive sharp wave potentials, and chronic motor denervative changes on EMG to be considered signs of LMN involvement in the examined limb or spinal segment, to be supportive of the diagnosis of ALS (350). Multiple studies support their utility in the clinical diagnosis of ALS, but the Awaji criteria have yet to become a standard component of the eligibility criteria of ALS clinical trials (351-353).

The revised El Escorial criteria advise electrophysiological studies to confirm signs of LMN dysfunction, identify electrophysiological evidence of LMN dysfunction in clinically uninvolved regions, and exclude other pathophysiological processes. Evidence of LMN dysfunction by needle electromyography (EMG) in the form of fibrillation and positive sharp wave potentials, and long-duration, high amplitude and unstable, complex voluntary motor unit potentials (MUP) with reduced recruitment supports the diagnosis of ALS but can be found in other disorders of the LMN (349, 350, 354). Relevance of these findings in a given patient in relation to a diagnosis of ALS must be considered in context, but are readily interpreted by experienced examiners. EMG confirming the presence of LMN involvement aids in distinguishing ALS from PLS, in which EMG is expected to show no evidence of active denervation (150, 151). Nerve conduction studies are an essential companion to EMG to establish the diagnosis of ALS and investigating other possible diagnoses. General guidelines for interpretation of nerve conduction abnormalities in ALS have been published (349, 350, 354, 355).
Additional neurophysiological techniques applied in ALS to evaluate LMN involvement include electrical impedance myography, motor unit estimates and the neurophysiological index. But these are used mainly as measures of disease progression rather than diagnosis (356-358).

Transcranial magnetic stimulation for measurement of central motor conduction time (CMCT) and motor evoked potentials (MEP) has been studied as a potential indicator of UMN pathology in ALS, but may be more sensitive to the primary pathology in PLS than ALS and is not an established diagnostic tool (359-361). Cortical hyperexcitability, assessed by transcranial magnetic stimulation to evaluate the cortical silent period, may represent an indicator of UMN pathology in ALS but requires further validation (362, 363).

4. Imaging Studies

The role of neuroimaging is predominantly focused on excluding other etiologies of motor dysfunction such as central demyelinating disorders, vascular disorders, radiculopathies, mass lesions, and other rare diseases. The field of neuroimaging is advancing at a fast pace (364, 365). Innovative MRI acquisition sequences and new PET ligands could, in the near future, help in assessing the UMN component of ALS in a way similar to the role of EMG in assessing the LMN dysfunction (366, 367).

5. Nerve/muscle biopsy

Nerve and/or muscle biopsies are not necessarily indicated or necessary for an ALS diagnosis (342). However, they may be utilized if there are confounding co-existing medical conditions or other clinical, electrophysiological or laboratory studies that suggest other etiologies to account for the presentation.

6. Genetic Analyses

Recent advances in genomics have identified many of the genes responsible for ALS and we now know the genetic etiology of approximately 15% of ALS cases (368). The genetics of ALS is described in more detail in Section IV Natural History. The insights gained have dramatically altered how ALS is perceived as a disease entity, and enhanced our understanding of the cellular mechanisms underlying motor neuron death. Genetics has an even greater potential to improve clinical care of patients devastated with this fatal neurodegenerative disease. Moreover, genetics will have an increasing impact on three distinct clinical areas. The first, and the most obvious, is that genetic screening can help to establish the diagnosis of ALS at an earlier stage. This allows disease-modifying treatments to be instituted sooner in the disease course, thereby magnifying their therapeutic benefit. Identification of certain mutations also predicts clinical course and prognosis. For example, patients carrying mutations in VCP or the pathogenic repeat expansions in C9ORF72 are more likely to develop FTD. (369). Patients carrying the p.A54V mutation of SOD1 uniformly manifest extreme decline, whereas the p.D90A mutation in the same gene are associated with a much slower clinical course (193).

This second area where genetic information will be central to clinical research is in clarifying clinical heterogeneity. ALS is not a single clinical entity, but rather a collection of clinical syndromes. The resulting variability likely obscures positive signals in clinical trials of therapeutic agents and in biomarker discovery efforts. Genetics offers a method to overcome this, or at least to compensate. Patients can be selected based on their genetic mutations, or the genetic information can be included in the statistical analysis to minimize confounding variables.
This focus on patients sharing specific genetic etiologies leads us the third, and perhaps the most exciting area, namely gene therapy. The notion of targeting the underlying genetic defect to ameliorate the disease is not new, but our advancing knowledge of ALS genetics and development of novel therapeutic modalities are making it a reality. Antisense oligonucleotide therapy has already been attempted in ALS patients carrying \textit{SOD1} mutations (36) and similar trials targeting \textit{C9ORF72} repeat expansion are planned. Success of these clinical trials will give impetus to the identification of the remaining genetic causes of ALS. As our genetic knowledge becomes more complete, we may even be able to predict if, and at what age, an individual may develop ALS and institute therapy to prevent or delay symptom onset.

Today, there are a number of methodologies that can be used to screen patients for mutations, ranging from Sanger sequencing to repeat primed PCR to next generation sequencing. Each platform carries advantages and associated disadvantages, and a variety of factors influence their selection including clinical scenario, family history, and the nature of the suspected mutation. Although the price of next generation sequencing is dropping, cost remains a barrier to accessing genetic testing. High-throughput, array-based technology may offer a cheaper alternative until whole genome sequencing becomes more affordable (370).

\textbf{Proposed Guidance}

The FDA recommends a flexible approach with regard to genetic testing in clinical trials as the list of causative mutations grows and newer genomic technologies come to the fore.

Genetic testing for the \textit{C9ORF72} gene is particularly important in cases where there is a family history of ALS or dementia. Genetic testing for \textit{C9ORF72} is recommended in sporadic ALS patients without a known family history of these disorders, particularly in the context of enrollment into clinical trials, where stratification of ALS patients may be of importance. Genetic testing for other genes should be considered in familial cases, or if there is a specific clinical clue (e.g. Paget’s disease of the bone or Inclusion Body Myopathy, suggesting mutations in the valosin-containing protein, or \textit{VCP} gene) to suggest a specific ALS subtype. Important genes to be tested in these familial cases may include \textit{SOD1}, \textit{TARDBP}, \textit{FUS}, and \textit{VCP}.

Genetic counseling should be offered to all ALS patients who are considering genetic screening.
VI. BIOMARKERS

A. General Comments

The term ‘biomarker’ has been defined as a “characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathologic processes, or biological or pharmacologic response to a therapeutic intervention” (371). The concept of a ‘biomarker’, however, is not a unitary one. Instead, there are several different types of biomarkers (diagnostic, prognostic, predictive, disease progression and pharmacodynamic), the desired characteristics of which may vary depending on the intended use or application. Moreover, within each of these biomarker categories, there are markers that are expected to be generic to motor neuron degeneration irrespective of the cause (e.g. CSF pNfH, CSF NfL and urinary p75NTRECD) and others that are specific either to subtypes of ALS (e.g. CSF SOD1 and CSF C9RANT products) or experimental therapeutics with particular mechanisms of action (e.g. TSPO-PET).

ALS trials have traditionally focused on evaluating clinical efficacy, with biomarker analyses omitted or of secondary importance. Enrollment criteria have relied solely on clinical parameters. However, promising clinical results in small pilot studies have evaporated in confirmatory larger trials. Given the clinical heterogeneity of ALS and its variable rate of progression, these confirmatory trials are often lengthy and require large numbers of patients. Upon completion, there still remain questions about whether a trial failed to demonstrate efficacy because of a flawed hypothesis related to mechanism of intervention, inadequate dosing, lack of target engagement, or whether a drug might work solely in a subset of patients. These questions highlight a number of important potential roles for biomarkers in ALS trials.

Biomarkers may serve a number of valuable roles in facilitating ALS clinical trials. A diagnostic biomarker, for example, could provide earlier and more definitive confirmation of a diagnosis of ALS when there is clinical suspicion, thus facilitating earlier trial enrollment. A prognostic biomarker could forecast the clinical course for a given individual, allowing for the selection of a more homogeneous trial population and increased trial power. A predictive biomarker identifies whether a given patient may be more likely to respond to a specific therapy and could help select an enriched patient population most likely to benefit from a particular treatment. A predictive biomarker might also serve to reduce variability, thus increasing the statistical power of a trial and reducing sample size. In a rare disease, a tool that permits a reduction in sample size can be uniquely influential, and these biomarkers could make a meaningful impact on the field. A biomarker that changes with progression of neurodegeneration in ALS might be referred to as a disease progression biomarker. Disease progression biomarkers in ALS should correlate, at least to some extent, with clinically meaningful outcome measures, such as the ALSFRS-R, though the nature of the relationship may be not be linear and the correlation may not be perfect. This is because 1) the biomarker may measure a disease aspect related to, but different from, the functional measurement and 2) because both measures will have some degree of random noise associated with measurement. In the case of clinical outcome measures, bias and subjectivity may also play a role. A disease progression biomarker might have specific applications in ALS clinical trials. First, such a marker might be used to enroll patients at an appropriate stage of disease. In this regard, it might supplant (or at least supplement) the clinical inclusion criteria in current use. Second, a biomarker of ALS progression might be applied as a type of pharmacodynamic biomarker, discussed below, demonstrating that the proposed therapy has exerted a biological effect.
Pharmacodynamic biomarkers may have the most influential effect on ALS trials by measuring the biological effect of a new therapy in an objective and quantitative fashion. Ideally, they can reduce subjectivity and bias and provide critical information about the effect of a therapy on the intended biological target. Pharmacodynamic markers, then, exist along a spectrum, including those that provide information about 1) target engagement, 2) effects on biological pathways, or 3) motor neuron and muscle condition. In practice, a trial might incorporate several pharmacodynamic biomarkers in order to capture all of these features. Ideally, one would learn first if the target was engaged, and if so, whether the biological pathway was altered sufficiently, and if so, whether disease biology was altered, and if so, whether clinical outcomes were affected. With this knowledge, investigators can then decide to press on with that therapy, change to another therapy in the same class, or abandon the class altogether, depending upon the profile of the biomarker response. In this way, the paradigm shifts from evaluating isolated responses to individual therapies to learning about broader therapeutic approaches and class effects.

Whatever its specific use, an optimal pharmacodynamic biomarker should be highly repeatable across laboratories and sites and should change demonstrably as a result of the biological impact of the therapeutic intervention. Because pharmacodynamic markers reporting on target engagement are often specific to the biology of a given therapy, they may need to be developed alongside the therapy, itself. Pharmacodynamic markers of biological pathway alteration (for example, markers of oxidative stress, excitotoxicity, or neuroinflammation) are more likely to correlate with clinical outcomes than markers of target engagement and may be more agnostic of the specific therapy. And pharmacodynamic biomarkers that are reflective of motor neuron and muscle condition need not be specific to a given therapy at all. Furthermore, they may be most likely to correlate with clinical benefit, though even this is not certain. However, if a pharmacodynamic biomarker is very strongly associated with clinical drug efficacy, it could eventually reach the level of a surrogate endpoint. A true surrogate must be quantifiable, substitute for a clinically meaningful endpoint and predict the effect of the therapy. While the process for qualifying a biomarker as a surrogate endpoint can be lengthy, even promising biomarkers that are not formally qualified by the FDA may be accepted by the FDA for use as a trial endpoint.

Table 5: Biomarker Types – from FDA Guidance on Qualifying a Biomarker

<table>
<thead>
<tr>
<th>Type of Biomarker</th>
<th>Description</th>
<th>Potential Utility in ALS Trials</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diagnostic:</td>
<td>A diagnostic biomarker is a disease characteristic that categorizes a person by the presence or absence of a specific disease.</td>
<td>- Early diagnosis for earlier inclusion in clinical trials</td>
</tr>
<tr>
<td>Prognostic:</td>
<td>A prognostic biomarker is a baseline characteristic that categorizes patients by risk of a disease or progression of a disease.</td>
<td>- Increased statistical power by selection of a more homogeneous trial population.</td>
</tr>
<tr>
<td>Predictive:</td>
<td>A predictive biomarker is a baseline characteristic that categorizes patients by their likelihood of response to a particular treatment.</td>
<td>- Improved likelihood of success and statistical power by selection of a trial population likely to respond to the proposed therapy.</td>
</tr>
<tr>
<td>Pharmacodynamic:</td>
<td>A pharmacodynamic biomarker is one for which a change in the biomarker shows that a biological response has occurred in a patient</td>
<td>- Improved understanding of the biological effects of a proposed therapy</td>
</tr>
</tbody>
</table>
who has received a therapeutic intervention.

| Disease Progression: | Absent from other lexicon on biomarkers, this term could be used to describe a biomarker that changes to reflect the current state of disease severity or advancement. | - Increased statistical power by selection of a more homogeneous trial population (same disease severity)
- Application as a pharmacodynamic biomarker demonstrating effect on ALS biology. |

B. Current State of Biomarkers in ALS (372)

1. Identifying and Using Biomarkers

Over the past 15 years, research to discover ALS biomarkers was typically performed by individual investigators with small sample sizes. Biofluid biomarker studies investigated unbiased proteomics, genome wide association screens (373-375), and more targeted approaches focused on proteins with known function suspected to be involved in ALS pathology (376-378). Imaging and electrodiagnostic studies explored existing technologies looking for specific changes in ALS, such as changes on MRI (379). Most of these studies explored putative diagnostic biomarkers, comparing healthy volunteers to people with ALS. Dozens of candidate ALS biomarkers were identified using these approaches (380, 381), but none have been fully validated.

Recently, biomarker discovery and validation studies compared ALS to more appropriate disease mimics (382-384) and leveraged large consortium efforts to boost sample sizes, explore reproducibility, and ensure standardization of pre-analytical and analytical techniques (385). Now, standard operating procedures (SOPs) for the collection, processing and storage of samples, and acquisition and analysis of imaging are being established. Biofluid SOPs have been put to work within the ALS community (386). Similarly, MRI SOPs are being developed and PET SOPs may follow. SOPs for some electrophysiological markers are also being developed.

Recent efforts have focused on collecting and longitudinally analyzing imaging, electrophysiologic, and biofluid biomarkers to characterize potential pharmacodynamic biomarkers, understand potential disease progression biomarkers and identify prognostic biomarkers for ALS (387-389). For example, CSF levels of SOD1 have been shown to remain stable in ALS patients and may be useful as a pharmacodynamic biomarker for SOD1 knockdown therapies (390). The application of motor unit number estimation (MUNE) to assess motor neuron loss with disease progression represents a second example (391). Current efforts are emphasizing combinatorial approaches to biomarkers (e.g., imaging and biofluids), larger sample sizes, and a focus on reproducibility and validation.

2. Current promising biomarkers (limitations/strategic)

It is critical to state at the outset that while no ALS biomarkers have been validated, there are a host of candidate ALS biomarkers at some stage of discovery/development, and these can be divided conceptually into categories by modality. The most common categories are electrophysiologic, biofluid, and imaging-based biomarkers. Ongoing drug development programs are beginning to incorporate exploratory biomarkers when possible, both to encourage
new biomarker discovery and to provide more robust information about the promise of new
drugs being tested. A more thorough review of the most promising biomarkers being
incorporated into clinical trials today can be found in Section VII Clinical Trials and Outcomes
Measures. By way of brief review, some of the most promising ALS biomarkers in development
include:

a. Biological-fluid-based Biomarkers

**Neurofilaments.** Neurofilaments are neuron-specific structural components of motor axons that
would be expected to be released as motor neurons degenerate. Blood and cerebrospinal fluid
(CSF) levels of neurofilament light chain (NfL) and phosphorylated neurofilament heavy chain
(pNfH) have emerged as leading candidates for investigation as pharmacodynamic biomarkers.
The majority of published data support the conclusion that NfL and pNfH are elevated in the
CSF of patients with ALS relative to healthy controls (392-398), (382, 399), and early indications
suggest that levels remain stably elevated through the course of disease (399). Moreover, the
degree of elevation of these neurofilaments may have prognostic value insofar as higher levels
predict more rapid disease progression and shorter survival (396-398), (382, 399). These
findings have been replicated in numerous laboratories throughout the world.

**p75 Neurotrophin Receptor Extracellular Domain (p75NTRECD).** Urinary quantification of
p75NTRECD is a leading candidate with promise as a generic biomarker of disease progression.
p75NTR, the low affinity NGF receptor, is highly expressed in motor neurons during
development and declines after birth. However, it is re-expressed following nerve injury and its
extracellular domain (ECD) is excreted into the urine. At present, urinary p75NTRECD appears to
increase as disease progresses in ALS patients (400). Moreover, higher levels predict more rapid
disease progression and shorter survival (401). These findings must be validated.

**SOD1 and C9RANT.** Limited available data suggest that CSF levels of SOD1 are stably elevated
in patients with ALS, with the potential that they might be reduced by introduction of anti-sense
oligonucleotides to SOD1 (390). Similarly, repeat-associated non-ATG (RAN) translation of
sense and anti-sense RNA containing the expanded GGGGCC repeat in the C9ORF72 gene
yields dipeptide repeat proteins (C9RANT products) (402-404) that can be quantified in the CSF
of patients who harbor this intronic hexanucleotide repeat sequence (49). Quantification of both
CSF SOD1 and CSF C9RANT products might have potential as pharmacodynamic biomarkers
of treatment effect in the sub-populations of patients who harbor an SOD1 mutation or a
C9ORF72 repeat expansion, respectively.

b. Electrophysiological Biomarkers

**Motor unit number estimation (MUNE), index (MUNIX), and compound motor action
potential (CMAP).** The utility of MUNE, which attempts to estimate the number of motor
neurons innervating a muscle or muscle group (391), as a biomarker of LMN disease progression
has been limited by the need for considerable training and real-time decision-making. MUNIX is
simpler to perform than MUNE but its reliability is less well studied and its derivation is less
intuitive. Repeatability, a weakness of MUNE and MUNIX, is usually lowest early in the disease
and improves only as the disease progresses toward end-stage when there are fewer, larger single
motor unit potentials (SMUPs). Finally, the CMAP itself, either alone or in combination with
other electrical physiological parameters, (356, 405) also has the potential of serving as a useful
biomarker of disease progression in ALS.
Electrical impedance myography (EIM). This technique, which quantifies the conductive and capacitive properties of muscle, yields reliable data that correlate with muscle strength (406) and are sensitive to disease progression (357, 407, 408), is easy to perform, requires minimal training, and can be applied to most superficial muscles including the tongue. Several ongoing multicenter studies (both therapeutic and observational) will help determine whether EIM measures can serve as bona fide biomarkers of disease progression.

c. Imaging Biomarkers. A host of brain imaging-techniques, including diffusion tensor imaging, magnetic resonance spectroscopy (MRS), functional MRI and, voxel and surface-based morphometry, as well as PET have been investigated for their potential to serve as biomarkers of upper motor neuron dysfunction. Most of these studies have been cross-sectional and relatively small.

Newer molecular imaging approaches, for example, targeting the translocator protein (TSPO), may hold more promise. TSPO, formerly known as the peripheral benzodiazepine receptor (PBR), is highly expressed in activated microglia and astrocytes and serves as marker of neuroinflammation and gliosis (409, 410). As such, TSPO PET imaging holds potential as a molecular imaging modality that could serve as a pharmacodynamic biomarker for ALS therapies that specifically target neuroinflammation. Ongoing and future studies will provide insight into its potential use as a prognostic or predictive biomarker or as a biomarker of disease progression.

Additional biomarkers in each of these realms are under investigation and should be considered for incorporation into trials as soon as enough data has been generated to support their validity and reliability. Furthermore, concerted efforts should support the development of even more biomarkers.

C. Use of biomarkers in trials

1. Context and illustrations of biomarker incorporation in ALS trials

Based on numerous high-profile late-stage ALS trials reporting negative results (22, 35), the ALS clinical research community has started to embrace the use of pharmacodynamic and disease progression biomarkers in early phase trials to improve the transition from Phase 2 to Phase 3 trials (96). An early study of memantine in ALS used elevated CSF levels of cytoskeletal proteins including tau and pNfH to monitor drug effects, with patients exhibiting the greatest decline in CSF biomarkers also showing the slowest rate of disease progression as measured by ALSFRS-R (411). Some trials have also begun to use predictive biomarkers to enroll subsets of participants most likely to respond to a given therapy.

For example, a phase 2a trial of fingolimod in people with ALS included a marker of target engagement, namely lymphocyte count in peripheral blood, and extended this with an exploration of gene expression in peripheral blood, a potential marker of immune function (NCT01786174). A Phase 2 study of NP001, a novel immune-modulatory drug, included an exploration of markers of inflammation, for example wide-range C-reactive protein and IL-18, which show promise as prognostic markers of response for people with ALS receiving NP001 (412, 413). These biomarkers are being incorporated at screening to enrich the study population in a subsequent trial of NP001. In fact, an upcoming trial of tocilizumab, an anti-inflammatory
biologic, will require the presence of a specific inflammatory gene expression signature as a predictive biomarker, used as an inclusion criterion for the trial (NCT02469896).

An ongoing trial of neural progenitor cells transplanted into the spinal cord of people with ALS is incorporating EIM as an outcome measure, using it as a biomarker of disease progression (NCT01772810) (414). A trial of retigabine, an anti-epileptic agent that increases opening of the inward rectifier potassium channel and reduces nerve hyperexcitability, is incorporating both nerve threshold tracking and transcranial magnetic stimulation in an effort to record the disease-relevant biological activity of the drug (NCT02450552). Finally, an early-phase ALS trial of inosine, which raises uric acid in plasma and glutathione (an antioxidant) in the CNS, is incorporating MRS to assess CNS glutathione levels as an ALS biomarker (NCT02288091).

These, and other trials, are attempting to bridge the drug development gap between early phase safety, tolerability, and dose-ranging studies, and later-phase efficacy studies by using biomarkers to indicate target engagement and even provide feedback on disease modification. To be sure, these efforts are only early-stage; most of the biomarkers lack validation and some have only preliminary evidence to support their use. Still, their incorporation represents an important signal from ALS researcher community that markers of target engagement and biological activity should be incorporated when possible into early phase trials to help guide choices about subsequent drug development.

2. Defining Biomarker Utility for a Trial; Context of Use as a Framework for Maximizing Biomarker Benefit in ALS Trials

The Context of Use framework is meant to guide investigators aiming to qualify a biomarker and needn’t be viewed as a binding recommendation, but rather as a framework for garnering the most useful information from biomarkers included in clinical trials. Not all elements described below are relevant for every biomarker. In addition, the COU statement does not need to have all the elements in the same order. The elements listed should be incorporated on an as-needed basis for the respective COU statement, typically based on the phase of development. For early phase trials in which the biomarker will not be used to support new drug approval, no qualification or COU is required by the FDA. However, at the same time, creating a COU and using it as a framework for the biomarker will ensure that its benefit is maximized for the development plan.
A COU statement contains a concise biomarker **purpose-of-use statement** and a comprehensive description of conditions for the biomarker to be used in the trial setting, termed the **conditions for use**.

- **The purpose-of-use statement** should be concise and include the name and identity of the biomarker(s) and purpose for use in drug development.
  - **Identity of the Biomarker.** The term “biomarker” may refer to a **single biomarker** with a single, specific COU, or to a **composite biomarker** that is made up of several individual biomarkers combined in a stated algorithm to reach a single interpretation. In the latter case, the COU applies to the composite biomarker as a unified entity. Individual components of the composite biomarker do not have separate COUs unless they are intended for use as stand-alone biomarkers. Examples of single biomarkers in ALS include a specific imaging modality (specific PET ligands, MRI sequences), a specific substance measured in a biofluid (e.g. CSF, urine, sweat, serum, plasma, PBMCs) or tissue (e.g. skin, muscle, fibroblast cell cultures), or a specific genetic/genomic marker. Examples of composite biomarkers for ALS might include a panel of CSF or plasma markers, a multiplex proteomic signature of muscle biopsy tissue from ALS patients, other –omic profiles/cluster patterns, or combinations of biomarkers across modalities, for example, imaging, biofluid and electrophysiologic markers.
  - **Aspect of the biomarker that is measured.** Examples include specific aspects of radiologic findings, such as volume, diameter, area, perimeter (e.g. volume of specific CNS regions), measures that quantify PET ligand binding, an electrophysiologic parameter, or concentration or enzymatic activity of a specific analyte in a biofluid. Certain biomarkers may require explicit temporal statements such as the window of measurement time, including specific time of day, steady-state, AUC, post-treatment minus pre-treatment, etc. Measurements may be graded or dichotomous based on a threshold, although graded assessments
are typically more informative (e.g. change relative to a baseline reference, historical control, normal range, or X-fold change).

- The **conditions for use** should contain a comprehensive description of conditions for the biomarker to be used in the qualified setting.
  - **Subject characteristics to which the biomarker should be applied.** A biomarker may apply only to a specific subset of patients. For example, patients with different:
    - **Disease phenotypes** – a biomarker of muscle atrophy might be more relevant in patients with predominantly lower motor neuron findings, while a CSF analyte purported to reflect CNS pathologic changes or an imaging biomarker looking at motor cortex changes might apply to patients with predominant upper motor neuron pathology. A biomarker that tracks cognitive changes might only be useful in patients with ALS-FTD.
    - **Specific windows of disease progression** – For people with ALS, certain biomarker classes might only be relevant during a limited time in disease progression. For example, electrophysiologic biomarkers, like CMAP, have a floor and fail to reflect further progression once they can no longer be elicited.
    - **Specific etiologies of disease or treatment modalities** – A biomarker to quantify knockdown of a specific target by antisense/miRNA treatment is likely not relevant in the absence of that treatment. Similarly, an analyte specific to ALS patients with a specific genetic mutation would only be useful in patients possessing that mutation. Another example would be measuring changes in serum reverse transcriptase levels in patients treated with antiretroviral medications. Such a biomarker would only be useful in the subset of patients with elevated serum reverse transcriptase levels at baseline and thus in whom endogenous retroviral activity is hypothesized to be pathogenic.
  - **Specific purpose of biomarker use in drug development, and circumstances for applying the biomarker.** It is important to describe specifically how the biomarker will improve the drug development process and through what mechanism it will do so. This might be a description of a type of problem that arises in drug development and for which the biomarker enables decision-making. At what stage of development is it being used? Will it be used to select the best drug candidate among several based upon a specific toxicity? Can it be used in phase 1 trials in healthy subjects to assess target engagement or guide dosing? Will it be used for dose selection in patients, i.e., to maximize target engagement (e.g., receptor occupancy, enzyme inhibition), which might translate into greater efficacy? Will it be used to demonstrate activity on the disease pathophysiology (i.e., proof of concept)? Is it biologically proximate to target engagement or more downstream in a pathway of interest? What is the biological interpretation of the biomarker and how will it be applied for the evaluation of a given treatment?
  - **Interpretation and decision-based action centered on the biomarker.** We strongly recommend inclusion of a decision-tree diagram that explicitly illustrates the application of the biomarker(s) in the COU and includes the actions that would be taken based on the biomarker results. For example, biomarker levels above X
indicate that an adequate physiological/pharmacologic response has occurred or that no significant toxicity has arisen. Or patients with biomarker positive for the presence of protein Z have at least an N-fold greater risk of an endpoint event rate or adverse reaction. For composite biomarkers, this includes the algorithm and rationale for any weighting used to combine the components in order to arrive at a single interpretation and how that single interpretation is then applied to decision-making.

D. Summary and Proposed Guidances

The ALS biomarker discovery and validation field is burgeoning and may be on the precipice of great progress. With more biochemical, cell, protein and RNA analyses, improved imaging techniques, and a broader array of diagnostic techniques, possibilities abound. Biomarker discovery and validation should remain a priority for the field. With appropriate attention and funding, the number and quality of ALS-specific biomarkers will continue to improve. Many of these potential biomarkers could play vital roles in future ALS clinical trials.

Proposed Guidance:

Investigators should incorporate relevant biomarkers into trials as soon as enough data have been generated to support their utility to the clinical development program. Furthermore, investigators must make concerted efforts to support the development and validation of even more ALS biomarkers.

ALS trials have traditionally focused on evaluating clinical efficacy but the number of negative late-stage ALS trials has highlighted the importance of biomarkers at every stage of clinical development. The Context of Use (COU) definitions outlined in the FDA Guidance, Qualification Process for Drug Development Tools (415) can provide a framework to help consider ways to incorporate biomarkers in ALS clinical trials.

Proposed Guidance:

Investigators may use the Context of Use as a conceptual framework to build biomarkers into ALS clinical trials. Investigators should seek to understand the limitations of each biomarker clearly and acknowledge its potential pitfalls to avoid unrealistic expectations.

Biomarkers come in a variety of forms, have numerous potential applications (Table 5), and may serve a number of valuable roles in facilitating ALS clinical trials. Of particular note, pharmacodynamic biomarkers may play a vital role in drug development by demonstrating target engagement and desired impact on the relevant biological pathways.

Proposed Guidance:

As a part of a clinical development program, investigators should develop and use pharmacodynamic biomarkers to demonstrate target engagement, help refine dose selection, and clarify biological impact of a potential therapy. Importantly, full qualification of these biomarkers need not be undertaken in order to use these markers to help guide therapeutic development, especially in early phase trials.
If investigators foresee the need and have the capability to seek FDA qualification of a biomarker as a surrogate endpoint for a clinical development program, or for acceptance as an *in vitro* diagnostic, they should engage the FDA early in the drug development process.

Finally, the initiation of new ALS trials need not be formally dependent upon the inclusion of a biomarker in the trial design. There may be circumstances in which biomarkers are not helpful, are prohibitively expensive, or otherwise complicate trial implementation. But in such cases, it may still be possible to collect limited materials or data for future processing and analysis.

**Proposed Guidance:**

Investigators should coordinate with established ALS biobanking or data repository initiatives to facilitate standard collection, storage, and sharing of samples and clinical information to help support future ALS biomarker development.
VII. CLINICAL TRIALS AND OUTCOME MEASURES

A. General comments

Although ALS is a rapidly progressive neurodegenerative disease and almost uniformly fatal, only one drug – riluzole -- has been approved by the FDA as an ALS-disease modifying therapy. In the past decade, at least 18 drugs have failed in large phase 2 or 3 ALS-related clinical trials (416) and many more have been tested in earlier phase studies with disappointing results. The failure of clinical trials to identify treatments with substantial benefits for people with ALS can be attributed to many factors, including poor understanding of underlying disease biology and targets for therapies, heterogeneity with regard to both disease mechanisms and clinical presentation, inadequate disease models resulting in limited understanding of the pharmacokinetics and pharmacodynamics of investigational drugs, lack of reliable and validated biomarkers, lack of clinically meaningful outcome measures, and inefficient trial designs.

The challenge of conducting clinical trials in ALS is particularly acute because of the relatively low prevalence of the disease and the fact that diagnosis is often delayed, resulting in potential trial populations that are already in advanced stages of the disease. Moreover, long-duration trials with a placebo cohort are felt by some individuals with ALS to be inappropriate for a fatal and poorly-treated disease, despite the fact that these practices are widely-viewed as the gold standard for demonstration of efficacy in drug development. While participants with ALS do enroll in randomized placebo-controlled trials, there is a need for innovative and highly efficient trial designs as well as the inclusion of all members of the ALS community, especially patients and caregivers, in the drug development enterprise.

In this section, we discuss the specific trial design characteristics of clinical trials for ALS, highlighting the need for flexibility and innovation in all aspects, including selection of control groups, outcome measures, and statistical approaches.

B. Clinical trials

1. Inclusion/exclusion criteria

Many people with ALS have a strong desire to participate in clinical trials but are excluded for a variety of reasons. However, the current lack of effective treatments and the urgent need for patients to have access to potentially beneficial therapies argues for a greater degree of inclusiveness in clinical trials, as well as the need to state a clear rationale for inclusion/exclusion criteria. As noted above in the Diagnosis section (Section V), the current understanding of what constitutes ALS is broad. Patients from the entire spectrum of disease should be studied in trials when possible; however, it is recognized that hypothesized mechanism of action or requirements based on expected rate of progression may limit the patients who might be appropriate for a specific clinical trial.

Inclusion and exclusion criteria are based on the goals of the study, drug target, phase of development, and safety for participants. Because there are no set criteria that apply to all studies, criteria should be well justified for each study. Criteria can be different in early development trials where safety, pharmacokinetics, dosing and biomarker effects may be the
primary questions versus late confirmatory efficacy trials. In early phase trials where safety and dose finding are key outcome measures and follow up may be short, participants with fast disease course or people with more advanced disease can be included. For efficacy studies, rate of progression of disease course may influence decisions on inclusion criteria. For example, in a study measuring functional outcome measures over a long period of time, including people most likely to complete the trial may be an important consideration. Careful consideration is required for each trial on each inclusion and exclusion criterion, and the reasons for inclusion and exclusion criteria should be clearly explained in the study protocol. It is not possible to have a standard set of inclusion and exclusion criteria for all clinical trials in ALS.

Some therapies may be more effective in early versus late stage illness; thus, sponsors may exclude later stage patients to increase the likelihood of demonstrating efficacy. For example, clinical trials of edaravone and methylcobalamin have raised the possibility that specific disease populations (e.g., those with early but disseminated disease) may respond to treatment differently than others (62, 417).

Sometimes participants with slow disease progression are excluded from clinical trials because of possibility that treatment effect may not be easily measured. However, particularly in early phase development trials, participants with slowly progressive disease may provide important information on safety and pharmacodynamics or other biomarker effects.

Many ALS patients take a variety of nutritional supplements or non-prescription drugs. Occasionally the use of such products is a reason for exclusion in a clinical trial. This is primarily due to safety concerns, as these products may dramatically interfere with the metabolism of both approved and experimental agents. While excluding patients from clinical trials due to use of supplements and non-prescription drugs is not recommended globally, patients, physicians, and sponsors should all be aware of potential risks, and nutritional supplements or non-prescription drugs may be exclusionary if a specific interaction with the potential therapy is known or expected. In most situations, an appropriate washout period should be defined according to what is known of the kinetics of that product.

Many ALS trials set levels of vital capacity (VC) as inclusion criteria, using either forced vital capacity (FVC) or slow vital capacity (SVC) measures. These criteria are set based on several factors, including phase of drug development, duration of study, mode of treatment delivery, and mechanism of therapeutic intervention. For example, in a study that involves a surgical intervention, participants with lower VCs may be excluded for safety reasons. Similarly, in a study designed to delay or prevent the need for non-invasive ventilation, lower VCs or use of non-invasive ventilation may be exclusionary. Where possible, inclusion of people with lower VCs is encouraged, if safe and rational.

Duration of disease at entry into clinical trials varies by trial, with some clinical trials recruiting early-stage participants and others recruiting people who have had the illness for several years. The decision is made based on expected effects of the therapeutic intervention and often phase of therapy development. In general, when possible, broader inclusion of people who have had the illness longer is desired. However, these decisions need to be made trial by trial to ensure the optimal chance of success of the clinical trial. An alternative to rigorous entry criteria is to stratify participants by factors otherwise used to select subjects. However, while this
approach allows for more variable subject entry, it has the potential to reduce power and may lead to unacceptable dropout rates.

The inclusion and exclusion criteria will impact both sample size and future labelling indication should a therapy be found effective and safe. A trial with broader inclusion may require a larger sample size because of increased disease heterogeneity. Studies with more restrictive inclusion criteria may require fewer participants because of enrollment of a more homogenous population, but recruitment may be more difficult. Additionally, if a therapy is initially found effective in a limited population, it is possible that it may only be approved and/or covered by insurance for people with those characteristics. Additional studies may be needed to evaluate efficacy and safety in other patients with ALS

a. Familial ALS. The decision on whether or not to include participants with familial ALS depends on the therapeutic intervention. If the pathway targeted by the drug is thought to be involved in a particular genetic type of ALS, then it is recommended that participants with that type of ALS be included.

A more likely scenario would be studies that focus on a familial type of ALS (e.g., SOD1 or C9orf72 gene silencing studies), where the question becomes whether to include or exclude people with sporadic disease. The answer to this question is not yet clear, and will evolve as more information is learned about possible role of mutated proteins (e.g., SOD1) in sporadic ALS.

There are some forms of familial ALS that progress very rapidly (SOD1 A54V, or p.A4V in the legacy nomenclature) or very slowly (SOD1 I 113T). Because inclusion of these participants could have impact on ability to determine efficacy, there may be trials where these participants might be excluded.

b. Genetic factors. Knowing genetic status in trials is increasingly important. It is recommended that study protocols and informed consent forms include information on sharing the results of genetic testing with the participants and the availability of genetic counselors. In addition, it is also important to include information on how and where data will be shared.

For most studies, it is recommended to collect DNA for genetic testing. While this may not be feasible for all studies, understanding the impact of genetic mutations on disease course and response to treatment is likely to grow in importance in the future. At the conclusion of the trial and with appropriate safeguards, it is recommended that these data are shared with the community to add to knowledge about the disease. As we learn more about ALS, the impact of genetic changes on disease course and response to treatment could be important.

c. Disease heterogeneity drives sample size and impacts the ability to see drug effect. For this reason, for certain studies, it may be important to limit inclusion in the trial to a more homogeneous group most likely to respond to treatment. However, this may limit the generalizability of the study results and affect labeling of the drug, should it come to market. These factors thus are considered when choosing inclusion/exclusion criteria for studies.

d. Cognitive / behavioral status. In general, it is recommended that participants with cognitive dysfunction are included in ALS clinical trials unless there is a well described and justified reason for exclusion. It is important to include cognitive screens in all studies.
Participants with FTD may have faster disease progression, which can impact decisions on gastrostomy tube and respiratory support (81, 84, 418). Ideally, the rate of progression should be determined prior to initiating therapy so the groups can be balanced between slow progressing patients, average progressing patients, and rapidly progressing patients. In addition, caregiver involvement is important to establish before enrolling patients to ensure a minimum level of involvement, as caregiving quality can also influence survival.

It is important to assess cognitive/behavioral status during the trial/therapy as there may be a selective effect of a therapy e.g. only on the motor system while cognitive impairment progresses or vice versa. Any of the published brief cognitive and behavioral screens used with ALS patients is recommended along with full neuropsychological testing for any patients testing in a range suggesting the presence of FTD.

A therapy may have a secondary impact on cognition/behavior, such as exaggerating cognitive impairment directly or indirectly by affecting mood. Testing all patients prior to starting the study drug and continuing to monitor during the trial will help measure any effect of a drug on cognition/behavior. An example of a drug previously tested in ALS that can impact cognition is topiramate. Some trials may also specifically target cognitive pathology. For these reasons, cognitive status may be used as an inclusion or exclusion criterion, but should be justified.

Protocols also need to address ability of participants to provide consent since cognitive and behavior change may have impact on decision making and informed consent to participate. In the majority of patients, capacity is not an issue, but adaptation may be necessary to achieve informed consent. Where relevant, assent should be gained from the patient and engagement of caregiver or family is necessary.

Therapies for FTD, such as those targeting behavioral changes, may be relevant to ALS. Any therapy proven to be effective in treating FTD or potentially promising should ideally also be tested in ALS patients and vice versa. An example of this is riluzole, which has been approved to treat ALS patients and is currently being tested in FTD patients.

Identifying cognitive/behavior status is essential prior to trial/therapy because of the potential for subgrouping (e.g., by site of onset or underlying genetic factors), since a therapy may be effective for one subgroup and not another. It may be necessary to exclude potential participants with co-existing psychiatric disorders who are taking sedating medications, as well as individuals with other conditions that affect cognition. However, late onset psychiatric symptoms such as somatic delusions or paranoia starting in the 40s or 50s should not be excluded as these may be part of the ALS dementia process. Trial participants with pre-existing dementias such as Alzheimer’s disease, vascular dementia, parkinsonism, metabolic disease, or other dementia or motor illness should be excluded due to the cognitive confound and the reduced life span associated with these co-morbidities. Participants with significant, unstable medical conditions known to affect cognition such as liver disease, kidney disease, or active cancer should likewise be excluded.

Proposed Guidance
Given the desire of the ALS patient community to actively participate in clinical research, inclusion criteria should be as broad as scientifically reasonable; broad criteria also increase the generalizability of results obtained. Use of supplements should not preclude participation in trials except in cases where interactions with the experimental agent is known or suspected. However, as proposed therapies become more specific with respect to targets (genetic or otherwise), it is recognized that inclusion criteria will need to be designed to reflect the specificity of the proposed treatment. Cognitive status should not preclude trial participation to the extent possible.

2. Feasibility issues

a. Ease of use and missing data. Missing data can impact the interpretability of study results, yet for a variety of reasons, about 25% of people enrolled in ALS trials stop participation. Reasons for early termination of participation include adverse events, difficulty getting to the trial site, caregiver burden, inability to swallow study medication, or belief that the treatment is not helping. It is important to learn more about the reasons for dropouts and modify trial design to minimize them. Strategies include pre-screening for participants vulnerable to problematic adverse effects; minimizing subject burden by providing travel support, home visits, and home-care teams; keeping visits as short as possible; and timing visits when best for the participant. The use of outcome measures that can be done remotely in the home, by home care teams, or over the phone, as well as the use of wearable technology can also reduce participant burden and provide additional valuable data; however, these measures require validation. Missing data due to death is another important factor, although the effect of death on ALSFRS has been investigated by simulation and shown to be relatively minor. Strategies to minimize the effect of death on functional outcomes include the use of combination outcomes such as the CAFS score, which incorporates both death and a functional outcome into a single measure, and was used in the Phase 3 trial of dexpramipexole (419).

Finally, providing feedback to participants on the trial – as much as allowed – may also improve retention in a trial. Mechanisms for providing feedback include newsletters, enrollment updates, and finding ways to thank them for being part of a team that is trying to find treatments for ALS.

It is also important to educate participants and the research community on the impact of missing data on the ability of a trial to determine if a treatment is effective and safe, and the importance of remaining in follow-up for assessment of outcome even if the intervention is withdrawn. Discussing the impact of missing data during the informed consent process and at subsequent study visits, allowing individualized, flexible treatment regimens, and keeping the follow-up duration as short as is scientifically appropriate can also help maximize retention. Targeting sites with a good track record of recruitment and retention and monitoring missing data continuously and carefully are also essential to ensure a high quality study that can be interpreted easily. All these approaches should also be balanced with ensuring the right of the participant to withdraw at any time for safety or other reasons.

3. Trial design

a. Study duration. The duration of trial participant follow-up depends on the phase of development and study endpoints. In early phase drug development, questions may be answered
efficiently with short duration studies. For example, studies focused on early safety, dosing, pharmacodynamics and pharmacokinetics can often be completed in a few weeks. Later-stage trials designed to assess efficacy are often 6 to 18 months in duration depending on the primary outcome measure. The power of a survival study is increased with duration of follow-up. Often studies that assess survival are between 12 and 18 months in duration. Twelve months may be too short given the low number of events and would require a large sample size. Another approach for studies that assess survival is to follow each participant from the time they enter the study until the study is over (e.g., when last participant completes 12 or 18 months of follow-up) (22, 35). For instance, a study might have three years of accrual and a year of additional follow up. In this situation the average duration of follow-up would be four years, which could become problematic if the treatment discontinuation rate is very high. Selection of study duration for studies where survival is a primary or key secondary measure need to balance the factors that drive sample size (number of events) with the reality of participant dropout due to longer study duration.

Studies using the ALSFRS-R or measures of vital capacity or strength as the primary endpoint can have a shorter duration of between 6 and 12 months. This shorter follow-up may be preferable because participants are more likely to remain on the therapy for this period and less likely to die, which reduces the challenge of how to treat deaths statistically in an analysis of ALSFRS-R and other functional outcome measures.

Time-to-failure endpoints have the potential to further shorten trials to the extent that the failure endpoint is reached early by a sizable number of participants. Participation may be further encouraged if such designs are coupled with an open label treatment arm after the endpoint is reached. Clinical trials of agents that target specific disease disabilities such as emotional lability (420), cramps, or spasticity can also be shorter, typically 3 to 6 months duration. For such targeted therapies with expected short duration of action, crossover designs where participants receive both active treatment and a placebo in a randomized fashion should also be considered, as such designs increase the amount of data gleaned from each participant and allow all participants to receive active medication.

b. Choice of treatment arms. Multi-arm studies should be considered for both early and late phase ALS clinical trials, both in adaptive and non-adaptive designs (421). For instance, one could randomize participants between multiple new chemical agents, or multiple dosages of the same agent in a preliminary trial without a placebo group. Then the agent (or dosage) with the best response would be tested against a placebo in a confirmatory efficacy trial. Alternately, a seamless adaptive design using a single control group in a multi-arm study could enable selection of the optimal treatment or dosage; with the control group continuing through the confirmatory efficacy trial. If this latter design is used, the final statistical test must be chosen so that the type I error is not inflated (422). Whichever strategy is used, it is recommended that sponsors explore the dose range for safety and pharmacodynamic effects early in the development program for a therapeutic agent. This ensures that critical information on the therapeutic agent is known prior to entering the confirmatory efficacy trial.

c. Designs in rare/genetic forms. Randomized controlled trials (RCTs) are the gold standard for detecting the effect of an intervention, as randomization can reduce bias. However, RCTs are more difficult to design and implement in rare diseases such as ALS. Even so, recent large phase 3 randomized clinical trials clearly demonstrate that RCTs can be successfully
conducted in ALS, despite its rarity (22). However, it may not feasible to conduct similar large RCTs in the genetically defined subpopulations within ALS. These populations are often small in number and geographically dispersed, making such trials difficult to implement (423). To the extent that a specific form of ALS is phenotypically homogeneous, and the natural history of this form is precisely known, then it may be possible to design trials in which the comparator group is historical rather than a concurrent control. Historical controls are discussed in Section IV: Natural History. Pharmacodynamic markers may also prove extremely useful in such a situation. Alternatively, trials may be designed to detect large treatment effects only, thereby requiring substantially fewer participants. Upcoming therapies, particularly those designed for the genetic subtypes of ALS (e.g., antisense oligonucleotides, AAV-mediated gene therapy, etc.), might be expected to produce such large effects. To this end, efforts at understanding the earliest onset of disease may create a window for earlier intervention and also may increase the opportunity for therapeutic benefit (424). This approach will require additional screening to identify patients at risk for genetic ALS. Once identified, these individuals can be targeted for suitable clinical trials or observational studies designed to better understand pre-symptomatic and early disease manifestations (424, 425). Therefore, RCTs are likely feasible within a genetically-defined population of ALS patients; however, should RCTs not be feasible, alternate trial designs should be implemented.

d. What constitutes an adequate control group? Estimating the safety and efficacy of new therapies depends on a comparison between outcomes experienced when receiving the new therapy vs. outcomes experienced when receiving a comparator treatment, generally standard of care. Accuracy in estimating the effects of treatment depends on comparison to an appropriate comparator. Comparators may be concurrent or historical. If concurrent, they may be randomly or non-randomly selected, and if historical, they may be data collected from the same participant or others. In this section, we consider comparison to historical controls when evaluating the safety and efficacy of new therapies.

Since the 1950s, double-blind, randomized trials have been the standard for evaluating safety and efficacy of new therapies in medicine. Randomization and blinding have become the standard because they minimize bias in estimating the effects of treatment. Randomization facilitates unbiased estimation of treatment safety and efficacy by minimizing confounding between the assigned treatment and both measured and unmeasured baseline determinants of safety and efficacy outcomes. Blinding of participants and investigators to the assigned treatment, generally through use of a placebo control, facilitates unbiased estimation by preventing confounding with post-randomization determinants of trial outcomes. Safety and efficacy outcomes can be confounded with treatment after assignment when participant assessment, participant management, or expectations or behavior of participants differ by treatment group. This is true whether the comparator group is a concurrent, open-label, randomized control or a historical control.

Advantages of historical controls. The advantages and disadvantages of using historical controls has been discussed for more than 40 years (426-429). The primary advantage of using historical controls is increased efficiency as either participants serve as their own controls or a comparator group is drawn from existing data. When participants serve as their own controls, estimates are not subject to within-person variance, reducing standard errors and thereby narrowing confidence intervals and increasing power. When the comparator group is drawn from existing data, the time and cost of enrolling additional participants is avoided and statistical efficiency can be increased.
by including data from many more control patients than would be feasible if they were concurrently enrolled. In addition, all participants receive the experimental therapy in trials utilizing historical controls, potentially increasing rates of enrollment if patients are optimistic about the intervention and increasing medical benefit to trial participants if the treatment is truly effective.

Given limited resources, reliance on concurrent, randomized controls can delay progress by reducing the number of studies that are feasible and reducing power for estimating efficacy. In 1972, Chalmers et al. (426) reviewed 19 studies of estrogen as a treatment for carcinoma of the prostate to argue for the need for concurrent, randomized controls. Of two controlled trials, one indicated efficacy and the other not. Of 17 uncontrolled studies, 16 were interpreted as supportive of estrogen therapy and one was not. Chalmers et al. suggested that the uncontrolled studies were likely overly optimistic given that the controlled studies were equivocal about the benefit of estrogen. While diethylstilbestrol rather than estrogen is now the pharmacologic used, hormone therapy has been conclusively demonstrated to extend median survival (430). In this instance, the protections against bias afforded by use of concurrent, randomized comparators reduced power to detect the real benefit of hormone therapy in early studies. Nevertheless, it should be noted that it was only the evidence from future, larger and thus better powered randomized trials that provided the convincing evidence that is used to support hormone therapy as a standard of care for carcinoma of the prostate today.

Sacks et al. (431) reviewed evidence of treatment efficacy from 43 trials using historical controls and 38 randomized controlled trials and found dramatic optimism among the HCTs, leading to very high sensitivity for detecting treatment benefit and very low specificity (i.e., treatments were judged effective whether they were or not). Conversely, RCTs were dramatically pessimistic, leading to very high specificity but low sensitivity (i.e., treatments were judged ineffective even if they were effective). Avoidance of elevated type I errors in HCTs is not easily achieved because those errors are likely due to unrecognized bias in the selection of historical comparators. Avoidance of elevated type II errors in RCTs can be remedied by increasing sample sizes and using less stringent criteria for declaring efficacy, namely by accepting higher probabilities of type I errors.

**Disadvantages of historical controls.** Disadvantages of historical controls center on the comparability of comparator data that is not concurrent and obtained by randomization. The accuracy of estimates obtained using a participant’s own history to predict their future course in the absence of treatment depends upon the accuracy of the prediction model for extrapolating historical data forward in time. Comparison to data from external historical controls is subject to confounding of chance differences between membership in the historical control sample or the experimental sample and the outcome of interest. This confounding can create positive or negative bias in estimates of treatment effect. In the absence of any concurrent, randomized controls, neither the magnitude nor the direction of bias from using inappropriate comparator data can be estimated. Accurate estimates from use of non-randomized controls depends on unverifiable assumptions that no unmeasured confounders or measured confounders for which inadequate adjustments are made could exert more than negligible influence on the outcome of interest relative to the effect of the experimental intervention.
While past failures in the use of historical controls should not wholly indict future appropriate use of such data, they do raise a cautionary note. For an example of previous misinterpretation of results based on historical controls, see Kyle 2005 (432).

Nevertheless, we should acknowledge that all evaluations depend to some degree on the use of historical controls. Even a demonstration of efficacy in a double-blind, placebo-controlled, randomized trial is judged on the basis of similar experience in other patients, whether considering if results from the placebo arm match expectations for typical outcomes under standard of care or considering whether the reported outcome comports with the proposed mechanism of action and is of a magnitude that passes tests of face validity.

**Discussion Summary:** The gold standard for clinical investigation remains the randomized, placebo controlled design. However, the absolute requirement for such a design requires reconsideration in diseases such as ALS that are poorly treated, ultimately fatal, and for which the natural history is fairly well known. In what follows the word “placebo” refers to a concurrent randomized control rather than the actual use of a physical placebo to blind that control.

The practical and ethical use of a placebo control group has been challenged in both early and late phase trials. Some contend that a control group is needed only for studies investigating primarily the efficacy of a potential treatment. Thus, if the primary objective is to investigate a treatment’s safety, using either clinical measures or biomarkers, then the use of placebo might be perceived as unjustified, especially if a trial involves certain elevated risks (e.g., invasive procedures, vulnerable populations, etc.). However, a therapy does have the risk of making participants worse, and this may not be possible to detect without an adequate comparison group. The use of placebo controls in early phase clinical trial research may also lead to overinterpretation of favorable secondary or post-hoc findings on efficacy signals. The avoidance of placebo in early stage trials would, therefore, “protect” from over-interpreting pilot data.

The urgency to find disease-modifying therapies for ALS and the material costs of instituting a placebo arm have led investigators to investigate the validity of using historical controls. The use of historical controls can lead to a substantial decrease in the required sample size for a trial, although care is needed in determining the required sample size since there is uncertainty associated with statistics obtained from the historical control group (433). However, the use of historical controls introduces potential biases affecting the study outcome, such as placebo effects and investigator bias if the only prospectively studied group is the active-intervention group. Furthermore, trials using historical controls have a higher risk of baseline imbalances among the treatment groups-- for example, trials with no concurrent controls may attract participants that are different from participants in controlled trials. Attempts to correct such selection biases by matching or statistical modeling (adjustment) can induce biases due to differences in the measurement of confounders. Historical controls are best selected from the placebo groups of recent, similarly designed trials. Importantly, it is well known that “natural history” is not the same as “placebo.” A caution is that the treatment of neurological disorders, in particular ALS, has evolved over time. Therefore, historical controls may lose relevance. Specifically, Cudkowicz et al. (434) showed that survival in ALS is improving over time, in part due to earlier intervention with supportive care.
Safety observations, as well as efficacy observations, can be affected by bias in trials using historical (or no) controls. This is true for clinical measures under voluntary or involuntary behavioral influence, as well as for biomarkers. Objective measures such as biomarkers do not necessarily provide unbiased results. For example, a biomarker might be a marker of a biological process, which could itself be under the influence of a placebo-effect. Also, the process by which a biomarker is measured or obtained might change over time or be subject to bias.

Dose finding and first-in-human-studies often have one or two placebo patients for each 8-10 treated patients in order to have placebo controls. It is unclear what benefit these placebo control patients have or how their data will be analyzed. The main point of these studies is to determine a dose that is tolerable. If there are not enough placebo patients in each dose level to use in this comparison, the data on the placebo patients is not used.

It can be argued that historical controls may have a role in trials that screen potential treatments. Indeed, highly effective treatments may be adequately evaluated using historical controls. In such cases the criteria from Byar et al. (435) should be applied to justify the use of a historical control. A potential danger with this strategy is that the finding of a (false) large effect, particularly for a treatment that is readily available, may make it difficult, if not impossible, to enroll participants in confirmatory trials of that treatment. Some researchers proposed supplementing placebo controls with historical controls to improve clinical trial efficiency. The role of the placebo control may be to provide informal validation of the assumptions underlying the historical control group (436). Formal Bayesian methods for combining historical and concurrent control groups have also been proposed (437).

**Ethics of using historical controls.** Development of clinical guidance from pilot trials using historical controls can be viewed as unethical (426, 438) as the optimism created by biased estimates from historical controls can be difficult to overcome. Performance of future RCTs is impeded because it is viewed as unethical to randomize participants away from a presumptively beneficial intervention, but only the RCTs can provide conclusive evidence of the relative safety and efficacy of an intervention.

**e. Use of predictive algorithms in trial design.** At stages prior to a pivotal trial, predictive algorithms and tools based on the algorithms are increasingly used to design, inform, and analyze results of early-phase clinical trials. A number of algorithms that could form the basis of tools that aid in decision-making have been described in the literature (439-444). Possible applications of the algorithms include the development of tools designed to optimize stratification at participant enrollment or for definition of groups to be analyzed following a trial. In addition, the effects of an intervention can be analyzed using a virtual control comprised of predicted outcomes generated at the end of a trial in ways similar to traditional historical controls by comparing observed patient outcomes to predicted virtual outcomes.

The use in pivotal trials of predictive algorithms and tools derived using these algorithms should be thoroughly discussed with the FDA at the time of the initial protocol review along with other aspects of the statistical plan of the trial. Time can be saved during review of a New Drug Application (NDA) if an algorithm or specific analytic tool has been previously deemed by the FDA as “fit for use” for a pre-specified application in ALS drug development. Comprehensive review of an algorithm or tool will be part of the NDA review if a predictive algorithm has not been found “fit for use” prior to its application during a clinical trial. Predictions of specified
outcomes for individual patients enrolled in a trial should be made at enrollment, sealed, and opened as specified in the study protocol for interim analyses or during the final analysis phase of the trial.

Algorithmic tools for device development should be qualified for their intended use as described by the Center for Devices and Radiological Health (CDRH) (445). There is no current CDER guideline for analytic tools of this nature used in drug and biologics development. In the interim, algorithm developers should closely follow the CDRH guidelines with appropriate modifications for drugs and biologics in their discussions with CDER.

**Proposed Guidance**

Despite the temptation to use historical controls, the interpretation of safety and efficacy data is significantly limited in trials without randomized, concurrent controls. Both the use of historical controls and predictive algorithms should be explored in middle phase trials. The development of target-specific biomarkers may also facilitate the use of these designs. However, given misleading patterns observed in past studies employing historical controls, and the lack of validation of current algorithms, late phase trials should always incorporate a randomized control group unless completely impractical. This recommendation is made while recognizing the distress that such controls may induce in trial participants, and with the desire to ameliorate that distress as much as possible. Clear expectations of open label access to experimental medication at the conclusion of active treatment (see below) may help with participant acceptance of controls.

**f. Role of open-label extension studies.** Although RCTs are ideal in evaluating efficacy of an investigational therapy, open label extension studies may have a role in ALS clinical trials, given that the disease is often fatal and currently no curative therapy exists. Often, an open-label extension in which active treatment at a dose felt to be safe is offered to participants subsequent to a phase 3 RCT for the purpose of making available to participants an investigational therapy that may ultimately be effective but that is not yet approved or licensed. In some instances, this may enhance recruitment to the original RCT.

Given that ALS is considered an orphan drug indication, an open-label extension may also afford an opportunity to provide longer-term safety data for an investigational therapy subsequent to a shorter and/or innovative trial design that may be reviewed and considered for approval. There may also be instances where a subpopulation of patients participating in the study may have experienced a potential benefit (‘responders’) and longer-term observation of these patients is needed to provide further descriptive efficacy data.

Finally, another role of an open-label extension may be to demonstrate continued efficacy of the investigational therapy over a longer period of time or to show that participants randomized to receive the active treatment during the open-label phase achieved outcomes similar to those of participants who received the drug from the beginning of the parent RCT. This is a more complicated issue, as open-label extensions are not able to control for biases that may influence evaluation and interpretation of efficacy; yet it highlights the strength of RCTs to control for many biases in assessing true efficacy of an investigational therapy. Some of the challenges in deriving efficacy data from open-label extension studies include the lack of
blinding, selection bias (e.g., introduced if only the participants who complete trial choose to enter the open-label extension), and the lack of a controlled comparator group. Despite these challenges, a well thought out open-label extension study may be designed in a manner with careful considerations to biases, to provide supplement or descriptive analyses of the effect of an investigational therapy in ALS.

Open-label extensions can provide a means for obtaining additional safety or efficacy data when working to build a more expeditious drug development pathway. However, several factors can influence the availability of open-label extensions. First, a careful assessment of benefit and risk needs to be considered. In addition, sponsors must consider the impact of the open label extension in the scheme of the overall drug development program, including requirements for longer term dosing in patients, cost to patients, and manufacturing and maintaining drug supply.

**g. Allocation ratio (1:1, 2:1, etc.).** Clinical trials in ALS and in many other conditions have often incorporated the strategy of unequal allocation of participants to the different treatment groups in clinical trials. A common argument made for this practice is the purported enhancement of recruitment and retention, which may be particularly important in the context of rare diseases. Another reason to employ unequal allocation may be the desire to obtain additional data on safety of the active treatment, especially if there is concern regarding a serious adverse event that may occur relatively infrequently. Cost considerations can also be used to justify unequal allocation: if one treatment is much costlier than the other, it may be cost-effective to increase the overall sample size and allocate more participants to the less costly treatment. Although unequal allocation results in a higher sample size requirement than equal allocation for a given power, the required increase in sample size may be relatively modest (e.g., approximately 12% for 2:1 vs. 1:1 allocation in a two-arm trial).

Those who advocate equal allocation note the scarcity of direct evidence demonstrating the effectiveness of unequal allocation in promoting recruitment and retention in clinical trials (446-448). There are also concerns that unequal allocation may aggravate the problem of “therapeutic misconception” (446) and that it may enhance expectations among trial participants, resulting in increased placebo effects (449). Although it is sometimes argued that unequal allocation may alleviate ethical concerns regarding the assignment of trial participants to receive placebo, such concerns should not be present if the investigators are truly in a state of clinical equipoise at the start of the trial. Indeed, it may be argued that because equal allocation of participants minimizes the number of participants required for any given trial, ethical considerations favor such a design as a risk reduction strategy.

**h. Function-specific therapies.** The design and approach of a clinical trial intended for a therapy directed at certain functional disabilities in ALS takes into consideration the mechanism of action of the therapy that is being investigated and the anticipated clinical effect of the therapy. The outcome measure used is specific to the clinical disability that the therapy is intended to effect. An appropriate scale, patient reported outcome, or quantitative measure that has the ability to detect a clinical effect may be used as an outcome measure. Ideally, the efficacy of the therapy should be evaluated using a placebo-controlled trial. Disability-specific therapies, in principle, can be evaluated in shorter duration trials as compared to therapies that target neurodegeneration. Depending on the mechanism, pivotal efficacy trials can be approximately 3-6 months in follow-up duration.
The recent approval of a therapeutic to improve functional disability in patients with ALS, i.e., Nuedexta for the treatment of pseudo-bulbar affect (PBA), was based on a study involving 3 months of follow-up (420). The pivotal phase III trial, also referred to as the STAR trial, was a 12-week, randomized, double-blind clinical trial in patients with ALS and MS with clinically significant PBA. A total of 326 patients were randomized 1:1:1 to placebo and two different dosing regimens, with 283 patients (86.8%) completing the study. Among completing patients, there was a significant reduction in the episodes of PBA (as measured by the CNS Lability Scale) in those who were treated with dextromethorphan combined with quinidine. The short duration of follow-up provided satisfactory safety and tolerability data to support the approval of Nuedexta.

Similarly, the approval of the cholinesterase inhibitor, donepezil, for the treatment of cognitive function in Alzheimer’s disease was based on studies of 3 and 6 months in follow-up duration (450). The approvals of dopamine agonists for the treatment of motor symptoms in Parkinson’s disease, and of tetrabenazine for the treatment of chorea in Huntington’s disease (451), were also based on studies of 3 and 6 months in follow-up duration.

i. Trials of combination therapies. Given the multiple pathogenic mechanisms thought to underlie ALS, there is a strong likelihood that multiple drugs may be needed to effectively treat the disease (452). Two four-arm clinical trial designs, which look alike but are analyzed quite differently, may be considered for evaluating the effectiveness of two drugs in combination. Both designs have one control group, a group receiving treatment A, a group receiving treatment B, and a group receiving both treatments. Such a study is called a factorial trial if its purpose is to determine the efficacy of A and the efficacy of B separately, while it is called a combination trial, if its purpose is to determine whether A and B are better than the control treatment and whether the combination of A and B are superior to both A and B alone.

The 2 × 2 factorial design can be highly efficient, allowing the investigation of the effects of two treatments at approximately the same cost of investigating one treatment. The main effect of each treatment is estimated using data from all four treatment groups. For example, the main effect of Treatment A is estimated by comparing treatment groups AB and A with treatment groups B and placebo (and similarly for estimation of the main effect of Treatment B). An important assumption is that the treatment effects of A and B are additive on an appropriate scale, i.e., that the effect of Treatment A is the same whether or not Treatment B is given, and vice-versa. This additivity assumption requires thorough understanding of the treatment and disease mechanisms; it can also be affected by the scale of measurement and treatment non-compliance. The assumption needs to be carefully justified prior to trial onset because although this assumption can be tested, the power of this test is usually quite low to detect plausible non-additivity.

If interest centers on the effect of combination therapy (AB), then a design comparing this treatment with placebo will not address the question of whether or not both components of the combination are necessary; demonstration of the superiority of combination treatment to each of the individual components (as well as placebo) is typically necessary. This will typically require a large number of trial participants to provide adequate power for the comparisons of interest.
A final point is that the potential for increased toxicity and lower compliance with a more complicated treatment regimen are often concerns in trials of combination treatment. These issues need to be carefully considered in the design of both learning-phase and confirmatory-phase trials.

C. Outcome measures and endpoints

1. Outcomes sufficient to support approval

a. Survival. Survival has long been accepted as a clinically relevant outcome measure sufficient to support approval in ALS trials. As a stand-alone measure, it has both clear evidence to support its use, as well as reasons to consider other options. It is without question an outcome of critical importance to patients. Although defining survival operationally has been somewhat problematic, the current definition used by most recent trials (death, tracheostomy or permanent assisted ventilation [PAV]) has been stable in the U.S. as rates of tracheostomy and PAV have not systematically changed.

The main issue limiting use of survival is the fact that most ALS trials are not of sufficient duration for many patients to reach this endpoint, severely reducing power. For example, the one-year survival in the celecoxib trial was over 75%, detecting an influence of a therapeutic agent during that time is almost impossible. The only two options to resolve this problem are to increase study duration or sample size, both of which contribute to cost, and reduce trial efficiency.

Survival has been linked as a joint outcome to functional measures in order to reduce the effects of dropouts due to death on data integrity. Under these circumstances, inclusion of survival can serve an important goal. However, for ALS trials to meaningfully evolve, it is clear that future trials will need outcomes more sensitive to change than survival. Such a choice requires the assumption that the treatment would not negatively impact survival if it improved the alternative outcome.

b. Function

ALSFRS-R. The ALSFRS-R is the most widely used rating scale in ALS trials. It is primarily a self-report scale, although it is intended that the evaluator query the participant to ensure accurate reporting. Comprising 12 items each scored from 0 (worst function) to 4 (best function), it surveys four domains: bulbar function, respiratory function, fine motor function and gross motor function. In its original version and its revised form, it has been in use since 1991, so that there are extensive data regarding its reproducibility and behavior over time. Although rate of decline has varied somewhat from trial to trial, average decline is approximately 1 point per month, with rate of decline being linear over the course of a year.

Advantages of the ALSFRS-R include the ability to administer it by phone, its high reproducibility, and its high correlation with survival. Limitations include the fact that the 4 domains do not change equally over time (fine motor function is the most important domain with respect to change), a lack of assessment of cognitive function, and the fact that the scoring of individual items does not meet criteria for interval scaling. As is true with any functional
measure, missing data due to death occurs during all trials, and no agreed upon method for data imputation has been developed.

A simulation study that compared the different methods of analyzing ALSFRS-R suggested that a random effects model treating data after death as missing at random did not create excessive bias (419).

**Proposed Guidance:**

ALSFRS-R may be an excellent option as a measure for potential disease modifying treatments, or functional treatments intended to impact the range of behaviors surveyed by the scale. However, for agents with more specific goals (i.e., swallowing and speech, respiratory function, emotional lability, others), the ALSFRS-R is not sensitive enough along any of its dimensions to serve as an effective measure, and more target specific measures would be important to support approval (64, 453).

c. **Strength**

i. **Respiratory function**

**Vital capacity.** Vital capacity is the most commonly used measure in assessing respiratory function in ALS. Vital capacity is a significant predictor of disease progression and survival in patients with ALS. The prognostic value of vital capacity for survival in ALS has been shown in several randomized clinical trials (56, 454) and in ALS clinic populations (67, 455). Vital capacity is established as a recommended test for ALS clinical trials and an important standard of ALS management (456). It is the most commonly used and most studied measure of respiratory function in ALS. Vital capacity is important for making clinical decisions for patients with ALS including the introduction of non-invasive ventilation, determining post-operative risk and recommendations for initiating enteral nutrition via feeding tube (107).

Vital capacity can be measured either by using a forced exhalation (FVC) or slow exhalation (SVC) maneuver. Vital capacity is often measured as percent predicted vital capacity and declines on average 2 to 3 percentage points per month in patients with ALS and generally declines in a linear manner during the course of the disease. Vital capacity can be measured upright or supine. Although there can be variability in the presence or absence of respiratory symptoms with given vital capacity measurement, the overall decline is associated with disease progression and survival in ALS. Additional information on vital capacity decline can be found in the Natural History section (Section IV).

When performing FVC, patients with bulbar weakness may have abrupt cut off because of glottic closure related to upper motor dysfunction. SVC has been preferred in recent clinical trials (22, 110-112) as it may provide a more accurate and reproducible measure of total volume exhaled in patients with glottic or bulbar weakness(104). However, evaluators need to be trained to perform this test, equipment must be calibrated and patients are required to visit the clinic to obtain this measure. Given that it is an objective measure of respiratory function associated with disease progression and survival and has a fairly predictable decline in patients with ALS, vital capacity is an ideal choice for an outcome in ALS clinical trials.
Measuring supine FVC or SVC, while logistically challenging in patients with extremity and truncal weakness, may be a more sensitive way to detect early changes in diaphragmatic strength (113-115). VC measurements are not only influenced by muscle strength, but are also impacted by the airways, chest wall, and lung parenchyma.

**Sniff nasal inspiratory pressure (SNIP).** SNIP is a simple and noninvasive means of measuring inspiratory muscle strength. This is a very simple procedure measuring peak nasal pressure in one occluded nostril during a maximal sniff performed from relaxed end-expiration through the contralateral patent nostril (457). It does not involve the use of a mouthpiece and therefore, use of SNIP avoids some of the conceptual issues surrounding the use of vital capacity.

Inspiratory pressure has been used in both Europe and North America as a criterion for initiation of non-invasive ventilation; it is also correlated with transdiaphragmatic strength and can be an early predictor of risk of intubation or mortality (458). However, the natural history of decline in ALS is not entirely understood. Additionally, multiple trials are required to ensure that the maximum pressure measurement is obtained (can be up to 10 to 20 trials). For these reasons, SNIP has not been frequently used as a primary outcome measure in clinical trials.

**Maximum inspiratory pressure (MIP) /maximum expiratory pressure (MEP).** MIP is a volitional test of inspiratory muscle strength. It is measured as the highest mouth pressure sustained for one second. during a maximum inspiratory effort against a closed system. It is usually measured at residual volume (RV) because inspiratory muscle strength is inversely related to lung volume (in a curvilinear fashion).

MEP is measured during a similar maneuver at total lung capacity (TLC) because expiratory muscle strength is directly related to lung volume (again in a curvilinear fashion).

The information available from these maneuvers are dependent on a hermetic seal around the mouthpiece, can be effort dependent and it may be difficult to distinguish between insufficient effort, muscle weakness, or a neurologic disorder. This can potentially limit the ability to use these respiratory measures in ALS when patients develop facial muscle weakness or bulbar dysfunction.

**Maximum voluntary ventilation (MVV).** MVV is a maneuver that requires maximum inhalation and exhalation over a fixed period of time (usually 12 seconds) and is measured in liters/minute. It can be a useful test to measure respiratory fatigue. However, it requires a good seal around the mouthpiece and maximum sustained effort. This can be a tiring and difficult test to perform for ALS patients. Additionally, there is a paucity of studies that have employed this measure in a clinical trial setting with ALS and therefore, limited understanding at present of the natural decline of this respiratory measure over the course of the disease.

**Proposed Guidance:**

Measures of respiratory function are clinically relevant in ALS, and should support approval of an experimental agent. Decline in respiratory function is a direct result of the known pathophysiology of the disease. The measurements themselves are direct tests of function. The most common cause of death in ALS is respiratory failure. For all these
reasons, a reliably measured aspect of pulmonary function that has clearly described changes with disease progression should support approval in a well-designed phase 3 study.

ii. Quantitative Extremity Strength

**Tufts Quantitative Neuromuscular Examination (TQNE)** Quantitative isometric muscle testing has been an important component of ALS trials since the development of the Tufts Quantitative Neuromuscular Evaluation (TQNE) in the 1980s by Munsat. The TQNE uses a strain gauge attached to fixed bars around and examining table, and required the patient to assume variety of positions to measure maximal voluntary isometric strength (MVIC) of 20 muscle groups, of both upper and lower extremities. TQNE produces interval data that accurately tests both strong and weak muscles(93). Summary arm and leg scores can be calculated by converting raw data to percent of predicted normal using a regression equation for each muscle group comprised of biometric factors. However, TQNE equipment is expensive, requires difficult position changes for patients, and requires a dedicated room(100). Data obtained using TQNE were reproducible, and demonstrated clear declines over time in ALS patients; however, the test was extremely fatiguing and missing data were common due to inability of patients to assume required positions, which has contributed to the infrequent use of TQNE in recent large clinical trials.

**Hand held dynamometry (HHD).** For these reasons, a battery of isometric strength measurements using a hand held dynamometer (HHD) was developed, and has been in use for approximately the last decade HHD tests MVIC of specific muscles in the arms and legs and can assess many more muscles can be evaluated than with the TQNE. HHD is portable, convenient for patients, as they may remain in a sitting position, can be performed quickly due to the fact that patient positioning is constant, and produces interval level data. HHD data can be evaluated for individual muscles, single limbs, or can be incorporated into a single averaged muscle strength value for each patient. However, HHD relies on the strength of the evaluator’s wrist to overpower the subject’s strength, resulting in increasing variability when testing strong muscles(101). In fact, in a recent Phase III trial of Ceftriaxone (35) using HHD, evaluators were unable to overpower the knee muscles in the majority of the 513 participants during their first study visit. In addition, most muscles are tested in an anti-gravity position, causing a floor effect in muscles that have less than fair (3/5) MMT strength. Recent ALS trials have demonstrated that isometric strength measured with HHD is reproducible, is highly correlated with other strength measure such as vital capacity, and declines over time at a constant rate when multiple muscles are combined.

As loss of strength is a hallmark of disease progression in ALS, its measurement should be considered a critical component of clinical trials in which disease modification or functional preservation is being assessed. It should be stressed that appropriate use of strength testing with HHD requires standardized training and demonstration of competence by evaluators performing this test.

**Manual muscle testing (MMT)** is universally used as a clinical measure during a neurological exam. MMT requires no equipment and is extremely convenient. MMT grades each muscle on an ordinal scale of 0-5. Unfortunately, the uneven steps between grades can be vast. In fact, the grades of 4 and 5 using the MRC scale can cover 97% of a muscle’s expected strength
Thus, when ranked scores are added together as if they are interval-level data, MMT summary scores can be inaccurate and misleading especially in small groups of data (99). Therefore, MMT is not frequently used in clinical research.

**Hand grip strength** using a JAMAR hydraulic dynamometer has been in standard clinical use for several decades (102). It provides a sensitive, reliable measure of grip strength. The grip meter is inexpensive, portable and easy to use. Recently, grip meters using load cells with wireless and the same dimensions as the original dynamometer are replacing the hydraulic version.

**Muscle strength fatigability** can be used with any load cell that has appropriate software. Often, grip strength fatigue is tested with a digital grip that is set to a sub-maximal level and the patient is asked to maintain that force level for a period of time. Fatigue can also be measured as the amount of decline in strength with several repetitions.

**ATLIS.** Other methods for quantitative strength testing include a recently developed tool called ATLIS; it is comprised of a specialized chair with a strain gauge placed in different locations. Thus, ATLIS is similar to TQNE in that a patient exerts strength against a fixed gauge rather than a hand held device, but is less fatiguing for a patient as there are no position changes except for moving into the test chair. Fewer muscles are tested than when HHD is used, and more proximal muscles are tested. ATLIS has been studied in normal subjects in in a small number of ALS patients, and has shown excellent reproducibility.

iii. Qualitative Extremity Strength

Medical Research Council (MRC) strength grading has been used for many years by neurologists in clinical settings. This strength grading system is bounded at the highest level (5) by the clinician’s view of normal strength, and at the bottom (0) by the absence of any muscle contraction. In between, grades 2-4 are graded based on the presence of movement with gravity eliminated as a counterforce (2), movement against gravity (3), and “good” but not normal strength (4). Appropriate use of this scale assigns virtually the entire range of muscle force as grade 4, with very small gradations in strength determining 3-0 grades (98). Although by averaging many muscle groups together, an overall total strength is reliable, it is less so than quantitative assessments (459).

**Proposed Guidance:**

Like pulmonary function measures, muscle strength is a direct assay of clinical relevance to patients with ALS. The major source of disability in ALS, progressive loss of motor nerve fibers, has a direct effect on strength. Function declines in direct proportion to weakness. For these reasons, we believe that a well measured, uniformly performed, and reproducible measure of muscle strength should be an approvable endpoint in an appropriately designed trial. Strength measurements should be quantitative.

d. Cognitive and behavioral scales. There is growing evidence that cognitive and behavioral function is part of the disease spectrum in ALS. Depending on the clinical trial and therapeutic agent being investigated, there are a number of cognitive and behavioral scales that may be used. An appropriate scale may be chosen with good justification and rationale. At a
minimum, ALS clinical trials should be compliant with the Core ALS instruments required by the NINDS (460).

**Table 6: Core ALS Instruments**

<table>
<thead>
<tr>
<th>Cognitive</th>
<th>Edinburgh Cognitive and Behavioral ALS Screen-ECAS (87, 461)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ALS Cognitive Behavioral Screen</td>
</tr>
<tr>
<td></td>
<td>UCSF Screen</td>
</tr>
<tr>
<td>Mood</td>
<td>ALS Depression Inventory (ADI-12)</td>
</tr>
<tr>
<td></td>
<td>Beck Depression Inventory II</td>
</tr>
<tr>
<td></td>
<td>Hospital and Anxiety Depression Scale-ALS</td>
</tr>
<tr>
<td></td>
<td>UCSF Screen</td>
</tr>
<tr>
<td>Behavior</td>
<td>ECAS</td>
</tr>
<tr>
<td></td>
<td>ALS Cognitive Behavioral Screen-ALS CBS</td>
</tr>
<tr>
<td></td>
<td>UCSF Screen</td>
</tr>
<tr>
<td></td>
<td>Cambridge Behavioral Inventory- Revised</td>
</tr>
<tr>
<td></td>
<td>Frontal Behavioral Inventory- ALS Version (FBI-ALS)</td>
</tr>
<tr>
<td>Pseudobulbar affect</td>
<td>Center for Neurologic Study Lability Scale for pseudobulbar affect (PBA)</td>
</tr>
<tr>
<td></td>
<td>Emotional Lability Questionnaire: Persons with MND (ELQ-MND)</td>
</tr>
<tr>
<td></td>
<td>UCSF Screen</td>
</tr>
</tbody>
</table>

**Proposed guidance:**

There are several cognitive and behavioral scales that are available for use in clinical trials. An appropriate scale with clinical justification and rationale may be chosen to target specific aspects of cognition and behavior when there is a need to screen for these symptoms in a clinical trial or if there is an investigational drug that is anticipated to treat or affect cognition and/or behavior.

**e. Specific tests of function.** ALS is a clinically heterogenous disease and outcomes measuring specific areas of function that are clinically important in activities of daily living are valuable in a clinical trial setting.

**i. 6 minute walk test (6MWT).** This test was originally developed for use in patients with cardiopulmonary disease, but has since been used in a variety of neurological conditions. It measures the distance that a patient can quickly walk on a flat, hard surface in a period of 6 minutes (the 6MWD). This test is self-paced and assesses the submaximal level of functional capacity. Since most activities of daily living are
performed at submaximal levels of exertion, the 6MWD may better reflect the functional exercise level for daily physical activities. It is a quantitative measure of mobility and leg function performance. Although this measure has been extensively evaluated in other neurological disorders such as Spinal Muscular Atrophy and Duchenne Muscular Dystrophy, it has not yet been evaluated in large prospective studies in ALS. A small pilot study in ALS indicated that the 6MWT can be performed in a trial setting and demonstrate changes (462).

ii. Timed up and go (TUG). The original purpose of this test was to test basic mobility skills of frail elderly patients. The original Get Up and Go Test used an ordinal scoring system based on the observer’s assessment of the patient’s risk of falling. The TUG measures a person’s ability to get up from a seated position and walk a fixed distance in seconds. In a small prospective study in ALS patients (n=31), the TUG correlated with ALSFRS-R and manual muscle testing (MMT). Over the six-month observation period, the TUG time increased linearly and was able to predict falls. The subject wears his regular footwear and uses his customary walking aid (none, cane, walker). No physical assistance is given. The subject walks through the test once before being timed in order to become familiar with the test. Either a stopwatch or a wristwatch with a second hand and be used to time the trial. This assessment measures functional independence and fall risk (463).

iii. Bulbar/Pseudobulbar Function. Although there are no established best practice parameters for evaluation of speaking and swallowing function in ALS, there are several measures that have been used in clinical trials that assess different aspects of bulbar function in ALS and may be used to evaluate investigational therapy aimed at treating specific aspects of bulbar function.

Center for Neurologic Bulbar Function Scale (CNS-BFS). The CNS-BFS, a self-report scale, is designed to evaluate and monitor bulbar function in ALS patients. It includes three domains: speech, swallowing and salivation. It has fairly good correlation with the bulbar domain of the ALSFRS-R and speech rate. This scale is being used as a primary endpoint in therapy intended to improve bulbar function in ALS (127).

Speech Rate. Speech rate is an objective timed test of speech where a patient reads a standardized item, resulting in a calculation of speech rate.

ALSFRS-R. The ALSFRS-R scale contains a bulbar subdomain which has 3 questions related to speaking and swallowing.

Timed Swallow Tests. The Water Swallowing Test (WST) estimates swallowing speed. While sitting, subjects are asked to drink 30 milliliters (mL) of liquid. The Timed Swallowing Test assesses the subject's ability to swallow solids (e.g., a tablespoon of 5 cheerios).

Modified Barium Swallow (MBS). The MBS is an objective test of swallowing and can provide information regarding swallowing risks.

Center for Neurological Study - Lability Scale (CNS-LS). The CNS-LS is a 7-item self-report scale that assesses pseudobulbar affect (PBA) by measuring the perceived frequency of PBA episodes (laughing or crying).
**Proposed Guidance:**

Tests of specific function should be well characterized with respect to performance by ALS patients, and paired carefully with proposed action of the experimental agent.

2. **Outcomes that would support approval secondarily**

Monitoring the disease progression in ALS is challenging and generally assessed by indirect imaging and neurophysiological evaluations. Either patient self-report or objective, reproducible, and sensitive tests that could measure degeneration could be useful for evaluating the therapeutic potential of novel ALS therapies. Although no current imaging or electrophysiological measure is validated as a surrogate outcome of ALS disease progression, careful pairing of disease mechanism and outcome measure may offer value through support of either target engagement or proof of mechanism. This type of data may augment evidence of clinical improvement and may be important in interpretation of significant but more modest effects. For example, PET imaging enables *in situ* quantification of specific proteins or sugars (e.g. Amyloid-beta, Tau, TSPO, glucose) or receptors (e.g. NMDA). Changes in these types of measures may suggest modification of the degenerative process. However, given that none of these measures have been validated, the impact on supporting an approval, independent of survival and ALSFRS-R, is unclear.

**Proposed Guidance:**

In cases of serious and deadly diseases with unmet medical need, such as ALS, surrogate endpoints or intermediate clinical endpoints may be used for accelerated approval in order to save valuable time in the drug approval process. ALS drug developers are encouraged to develop, validate and use surrogate endpoints or intermediate clinical endpoints to measure therapeutic effect considered to be reasonably likely to predict the clinical benefit of a drug.

a. **Patient /Caregiver reported outcomes.**

To minimize missing data, it is often advantageous to incorporate outcome measures that may be obtained remotely. The most often used of these measures is the ALSFRS-R, which can be administered by telephone. In multiple prior clinical trials, it has been shown to be reliably obtained and provides results consistent with those obtained in clinic (453). Other remote measures that have been used include assessment of pulmonary function using the PICO 6 spirometer and activity measures obtained using accelerometers placed on limbs. These last measures have been used previously in stroke trials more than ALS trials, but may have applicability in the future (Deak, 2009).

Patient reported outcomes (PROs) regarding health related quality of life, overall quality of life, or clinical aspects of patient condition can be helpful in providing the patient or caregiver perspective of therapeutic efficacy. Generic and disease-specific scales are available, but even the use of a well-known scale should be validated and sensitive to change in ALS. The selection of a particular patient or caregiver reported outcome should be justified and can be supplemental to the primary efficacy data from a clinical trial.
i. SF-36: a generic, multi-purpose, survey of 36 questions which provides a functional health profile, psychometrically-based physical and mental health summary, and a preference-based health utility index.

ii. ALSAQ-40/ASLAQ-5: a disease-specific patient self-reported health status questionnaire. The ALSAQ-40 is a 40-item measures QOL in five areas over the previous two weeks: physical mobility, activities of daily living and independence, eating and drinking, communication, and emotional reactions. The ALSAQ-5 asks one item for each area. The ALSAQ-40 has been validated in multiple languages and correlates very well with the SF-36 (www.sf-36.org), which was developed for measuring general health outcomes (464-470).

iii. EQ-5D: a scale developed by the EuroQOL group, the EQ-5D is a standardized instrument for use as a measure of health outcome and is used to generate Quality Adjusted Life Years (QALYs). The EuroQOL-5 (www.euroquol.org) is a similar patient-report scale commonly used in European clinical trials.

iv. McGill QOL: a questionnaire relevant to all phases of the disease trajectory for people with a life-threatening illness. This questionnaire differs from most others in three ways: the existential domain is measured; the physical domain is important but not predominant; and positive contributions to quality of life are measured.

v. Sickness Impact Profile (SIP): a generic measure used to evaluate the impact of disease on both physical and emotional functioning. Patients are asked to respond to the items as they are on that day.

vi. Clinical Global Impression (CGI) scale: These rating scales are commonly used measures of symptom severity, treatment response and the efficacy of treatments.

vii. Caregiver Burden Inventory: A 24-item multi-dimensional questionnaire measuring caregiver burden with 5 subscales that include time dependence; developmental, behavior, physical burden, social burden, emotional burden.

b. Structural measures

MRI: several magnetic resonance imaging (MRI) sequences have been used to show differences between ALS patients and controls. These include as diffusion tensor imaging (DTI), surface based morphometry (SBM), voxel-based morphometry (VBM), magnetic resonance spectroscopy (MRS), and resting state functional connectivity. Some of these measures showed change over time such as SBM, VBM, and DTI. The challenge for these MRI signals is that they do not represent specific disease biology. In addition, there is a need to standardize data acquisition and analysis across multiple platforms.

PET: several PET tracers have been used to show differences between ALS and controls such as fluorodeoxyglucose (FDG), peripheral benzodiazepine receptor 28 (PBR28), and flumazenil. These tracers bind to specific proteins, sugars, or receptors, providing insight into disease biological status. The advantage of showing positive changes in PET signals in response to treatments is that they can provide additional valuable information such as proof-of-mechanism, target engagement, and change in ALS biology. For example, an experimental anti-
inflammatory therapy that shows reduction in PBR28 signal would have an added mechanistic value that compliments a positive clinical outcome.

c. Neural function

**Electrical impedance myography (EIM):** EIM is an electrophysiological technique in which current is applied to a muscle of interest, and resultant voltage and impedance are measured. Several parameters are obtained, some of which are abnormal in many neuromuscular diseases (471), including ALS. The major derived EIM parameter, 50 kHz phase, correlates with survival and presumably the pathologic state of denervated muscle in patients with ALS. EIM has been shown to be a sensitive measure of disease severity in ALS. Longitudinal natural history studies have also shown correlation of some EIM parameters with survival (357) and standard measures of disease severity, including ALSFRS-R upper and lower extremities subscores and HHD (406). Though measured values in EIM are not specific for ALS, change of EIM may suggest modification of muscle changes that occur in ALS and as such may complement a positive clinical outcome. EIM has the advantage of being straightforward to obtain, quick, non-painful, and reproducible, and therefore propels further investigations into its utility as an outcome measure for clinical trials.

**Compound Motor Unit Action Potential (CMAP):** CMAP is a standard electrophysiological measure generated by maximally stimulating a nerve such that all muscle fibers innervated by the respective nerve are depolarized. Reduction of CMAP amplitude reflects loss of motor axons and, therefore, is directly relevant to ALS. Furthermore, median nerve CMAP values decline substantially in ALS patients (358); typically 50% loss is required before the CMAP is detectable as small. Despite the simplicity and attractiveness of CMAP, it has been difficult to obtain reliable and repeatable CMAP values because it may vary substantially with stimulus intensity, electrode position, limb position, and temperature. Likely for this reason, it has not been widely used as an outcome measure in ALS clinical studies. However, with careful standardization CMAP may prove useful as a measure of disease progression.

Although the CMAP reflects the total number of remaining axons, it does not provide a count or estimate of the number of motor axons present. Muscles differ widely in how many muscle fibers a motor axon innervates, and therefore equivalent CMAP amplitudes could reflect vastly different numbers of motor axons. If the response generated by triggering a single motor axon can be determined, one can divide the CMAP size to yield an estimate of how many axons are present independent of whether sprouting has occurred.

**Motor Unit Number Estimation/Motor Unit Number Index (MUNE/MUNIX):** A series of techniques have been developed since the early 1970s to determine the number of motor neurons present, including the “incremental stimulation,” “statistical,” and “multipoint” methods. All methods show essentially linear reductions in the number of motor units with faster declines than strength or functional measures. However, some methods are very time-consuming and require many shocks, limiting the number of nerves that can be studied and being quite burdensome to study subjects. Efforts to make MUNE results quicker to obtain and more standardized across multiple sites have been challenging, but one multipoint method may have figured out how to overcome some of these issues (358).
MUNE and MUNIX both estimate the number of functioning motor units within a muscle. While multiple MUNE techniques exist, multipoint incremental motor unit number estimation (miMUNE), which combines incremental stimulation of the nerve and stimulation at multiple points along the nerve to estimate single motor unit action potential (SMUP), is attractive because of simplicity and reproducibility. miMUNE reliably estimates the number of motor units in a muscle or group of muscles, and it has been shown to be a sensitive index of motor neuron loss (358). Subjects with ALS showed an average rate of decline in miMUNE of approximately 9% per month (357, 358); the coefficient of variation of the rate of change from baseline was lower compared with the ALSFRS-R. In addition, since both measures decline more quickly than ALFRS-R (357), they could potentially be used in clinical studies with reduced sample sizes.

Like MUNE, MUNIX estimates individual functioning motor unit responses within a muscle. However, while MUNE directly estimates motor unit number, MUNIX uses a statistical approach to estimate functioning motor units within a muscle. CMAP and surface electromyography potentials (surface interference patterns) are obtained at various levels of voluntary effort, and MUNIX is estimated using power and area of CMAP and surface interference patterns. MUNIX may be used to interrogate any muscle in which a reproducible CMAP may be obtained. Like other techniques that estimate motor unit number, MUNIX appears to decline more quickly than ALFRS-R and may reflect disease progression(472). Single center studies also suggest that MUNIX may provide equivalent information to MUNE while being substantially faster and less variable (473, 474).

**Excitability testing.** Axonal hyper-excitability has been observed in ALS patients. The profound increases in threshold and latency reduction to depolarizing currents suggest that the alteration of membrane excitability may be a relevant component of ALS disease pathophysiology. Remarkably, the degree of this abnormal hyper-excitability appears to correlate with patient survival (475). These observations suggest that hyper-excitability is a phenomenon that could be present in various ALS populations. If this phenotype can be reliably measured in man it could serve as a proof-of-biology biomarker to track the effect of therapeutic interventions. Threshold tracking is a clinical neurophysiology technique that tests nerve axon excitability by repetitive stimulation aimed at eliciting a pre-defined CMAP (usually 30 to 50 % of its maximum amplitude). Threshold tracking can be applied to motor or sensory axons to determine changes in excitability caused by a single impulse (e.g. refractoriness and superfexcitability), by changing stimulus duration (strength- duration time constant), or by subthreshold polarizing currents (latent addition and threshold electrotonus).

**Transcranial magnetic stimulation (TMS).** Characterizing patterns and severity of UMN involvement is challenging in ALS due to the lack of easily standardized assessments and the fact that severe LMN involvement often blocks the ability to detect UMN involvement clinically. TMS is a non-invasive technique that can assess motor cortical and corticospinal function, which are abnormal even in early ALS. This fact has driven efforts to use TMS to improve diagnostic accuracy (363). There are several different techniques that can be employed including single-, paired- or multiple-pulse techniques. Potential measures that can be derived from TMS include motor threshold (MT), motor evoked potential (MEP) amplitude, central motor conduction time (CMCT), cortical silent period (CSP), intracortical inhibition and facilitation. Longitudinal TMS studies in ALS patients reported a significant reduction in MEP...
amplitude, MT and CMCT, and suggested that reduction in MEP amplitude may be an objective biomarker of disease progression in ALS (79). Others have failed to document any significant longitudinal changes in TMS parameters(80). Multiple prospective studies are being planned to better address the potential utility of TMS as a biomarker of progression. It should also be noted that equipment and the specialized expertise required for TMS are not currently available at many academic centers.

3. Outcomes that assay potential disease pathways

ALS is characterized by several interconnected pathogenic events that lead to motor neuron cell death, including inflammation, oxidative stress, protein aggregation, altered RNA metabolism, glutamate hyperexcitability, metabolic dysfunction, and altered DNA expression and damage (476). While numerous markers are being investigated as possible outcome measures, none have been validated as a primary outcome measure for ALS clinical trials. Given the biologic complexity of ALS, a composite biomarker consisting of a combination of several biomarkers may be necessary to reach an interpretive readout. There is a critical need for the incorporation of validated ALS biomarkers in ALS clinical trials (477). Biomarkers are covered in more detail in Section VI.

a. Inflammatory markers. Multiple inflammatory biomarkers have been proposed as markers for diagnosis, disease progression, and disease mechanism however, none have yet been clinically validated. Blood and CSF are primary matrices, though other tissues have been examined. Multiple methodologic and practical concerns exist for which guidelines have been suggested. Ongoing prospective trials are needed to validate proposed inflammatory markers.

b. Markers of oxidative stress. The mitochondrial respiratory chain generates reactive oxygen species and these reactive oxygen species are removed from the cellular environment by antioxidant defenses. Oxidative stress occurs when there is an imbalance between these processes resulting in excess reactive oxygen species that can cause injury and cell death to motor neurons. There is strong evidence supporting the role of oxidative stress created by excess reactive oxygen species in inducing cell death in ALS (478). Although numerous biomarkers measured in blood, urine and CSF, none have been clinically validated. A few examples of biomarkers of oxidative stress being developed in ALS include:

activity of catalase: a reduction in activity of glucose-6-phosphate dehydrogenase, and glutathione reductase was observed in erythrocytes of ALS patients

glutathione: levels of antioxidant, glutathione, was decreased in erythrocytes of ALS patients, and this reduction was correlated with the duration of the disease (479).

uric acid: uric acid also possesses free radical scavenging activity and was decreased in ALS patients. This reduction was shown to be correlated with the rate of disease progression (480).

8-oxo (or 8-hydroxy)deoxyguanosine: one of the principal DNA adducts, derived from mitochondrial DNA and nuclear DNA, this marker can indicate current level of oxidative stress.
**15-F2t-isoprostane**: derived from arachidonic acid via a free radical-catalyzed mechanism, it is used extensively as a clinical biomarker in various diseases including Alzheimer’s disease where it can indicate current level of oxidative stress.

**plasma protein carbonyl**: an end product of intracellular amino acids damaged by excessive ROS; indicates longer term (approximately 3 months) oxidative stress (481).

c. **Markers of cell damage.** The most advanced markers of cell damage in ALS with potential for use as outcomes in ALS clinical trials are neurofilaments (both light chain, NfL and phosphorylated heavy chain, pNfH) and the ratio of pNfH to C3 complement (382). Elevated neurofilaments in CSF and blood may be a marker of axonal loss (482).

The use of neurofilaments as clinical trial markers is underway. CSF pNfH and tau levels have been studied as outcomes in two recent small therapeutic trials (411, 483) and pNfH, pNfH/C3 ratios and tau in CSF and blood will be included as secondary outcomes in an upcoming ALS trial of memantine (ClinicalTrials.gov Identifier: NCT02118727).

d. **Other.** There have been a number of exploratory studies using mass spectrometry of CSF to identify changes in the CSF and plasma proteome, metabolome and neurochemicals in ALS and other neurodegenerative diseases. These studies have identified several candidate biomarkers, but have a sensitivity limitation. Recently ultra-sensitive immunochemical assays (Irenna or Simoa platforms) have been developed identifying and quantitating a number of less abundant intracellular proteins in CSF. These assays have been used to detect fg/ml concentrations of proteins in the CSF. Examples of the use of these technologies is detection of mutant huntingtin protein in CSF, demonstrating a correlation between CSF protein levels and disease progression (484), detecting survival motor neuron (SMN) protein in CSF in response to interventional therapy (485), detecting changes in plasma tau levels in response to brain injury (486, 487). In principle, these platforms can be further optimized to allow multiplexing of samples. The main limitation is that this is a directed approach requiring some knowledge of the pathways being investigated to identify candidate analytes.

A second approach is to monitor exosome contents (RNA and protein) in CSF. Exosomes are extracellular vesicles secreted by cells that can reflect changes in intracellular proteins and RNAs occurring in tissues. The most advanced application is quantifying changes in microRNAs in CSF (488). Although methods are still being developed to reproducibly isolate and quantitate changes in exosome contents in CSF, this is could be an important source of material to detect changes in neurons and glial cells.

4. **Time-to-failure endpoints**

It is possible to define a progression endpoint in ALS that would be defined as a 6-point drop in ALSFRS. This was proposed as an “endpoint” in a trial of lithium (489). Its advantage was that patients who progressed could go off study and be treated with a different therapy, thus this trial was considered to be more ethical or at least more acceptable to patients. However, if one analyzes the actual ALSFRS measurements using a random effects model with the data generated by such a trial, the power would be greater than the power of a proportional hazard model using the time to progression. Both analyses are valid as the data after progression would be Missing at Random (MAR). Given that the random effects model is more powerful, it is a
better analysis plan. The design where patients are allowed to leave the study on progression can be used whenever it is ethically or practically necessary.

5. Safety outcomes

How safety is monitored and analyzed in a clinical trial depends on the severity of the untreated disease. For instance, in some oncology trials potentially life-threatening side effects of therapy are allowable because the untreated disease is fatal in a relatively short time frame. For other diseases, such as hypertension, only minimal side effects are allowable. ALS is intermediate as survival is 3 to 5 years but the disease is nearly uniformly fatal. Given this, patients would tolerate some level of discomfort for an effective treatment and life-threatening side effects might be acceptable if they were rare. This consideration could be used to shorten the preclinical and early clinical testing of new drugs.

In general, the International Conference on Harmonization (ICH) E1 Note for guidance on population exposure applies when monitoring safety in a clinical trial setting.

When monitoring safety in a clinical trial, adverse events (AE) should be characterized in relation to the duration of treatment, the dose and/or plasma level, the recovery time, age and other relevant variables. Assessment of adverse events, especially those predicted by the pharmacodynamic properties of the investigational drug should be performed using a systematic and planned methodology.

All adverse events occurring during the course of a clinical trial should be fully documented with separate analysis of adverse drug reactions, drop-outs and patients who died while on therapy. Depending on the investigations drug being studied, relevant guidelines with specific safety topics should be taken into account.

D. Statistical issues and analysis challenges in ALS

1. Missing data

Missing data are ubiquitous in clinical trials. In ALS clinical trials, missing data arise mainly from participants who die while engaged in the study, withdraw from the study due to hardship or perceived lack of response, are unable to complete outcome measures due to difficulties in conducting evaluations as the disease progresses, or withdraw due to treatment specific adverse events. While a variety of strategies have been proposed to ameliorate the effect of missing data, all have the potential to introduce bias. Whatever strategy is chosen should be specified prior to study initiation.

Methods for accommodating missing data in the statistical analyses should be pre-specified and implemented in these trials. The FDA-commissioned report from the National Research Council (490) entitled “The Prevention and Treatment of Missing Data in Clinical Trials” provides highly useful set of recommendations pertaining to this issue.

Ad hoc methods of accommodating missing data such as carrying forward the last available observation (LOCF) are no longer acceptable, particularly in trials of progressive
conditions such as ALS. Analyses that use so-called single imputation methods that substitute a single value for the missing datum and treat such data as if they were observed, in addition to possibly introducing bias, have the major limitation that they fail to account for the uncertainty associated with the imputation, leading to false precision that is manifested in narrower confidence intervals and smaller p-values than are appropriate.

Many other methods have been developed to accommodate missing data, but their validity depends critically on their underlying assumptions concerning the missing data mechanism. Methods that assume that the data are missing completely at random (MCAR), such as complete case analysis (i.e., an analysis that eliminates subjects without complete data) and marginal models fit using unweighted generalized estimating equations (GEE), are rarely appropriate unless the missingness is independent of both the observed and unobserved values. More commonly, methods that assume that the data are missing at random (MAR), including direct likelihood methods (e.g., mixed effects models), multiple imputation, and marginal models fit using weighted GEE (inverse probability weighting), are used for the primary analysis of clinical trials. Data are considered MAR if the missingness is independent of the unobserved values given the observed values. The essence of the MAR assumption in the context of longitudinal data with subject dropout is that reasonable predictions of future values for subjects who drop out at a given time can be made from those who have observed data at or after that time.

While the MAR assumption may be more realistic than the MCAR assumption in many cases, this assumption is untestable. For the reason, the National Research Council report (490) recommends that sensitivity analyses be performed that relax this assumption. For example, analyses can be performed that make various assumptions about the difference in outcome between those who do and do not have missing data, separately in each treatment group, with the MAR assumption as a special case. Pattern mixture models and selection models have been proposed to facilitate these sensitivity analyses; in particular, multiple imputation can be useful in the context of pattern mixture models for this purpose (491).

Sample size planning for these trials ideally needs to account for both the increase in variability and the increase in bias (attenuation of treatment effect) due to missing data and non-compliance, based on reasonable assumptions that are appropriately justified.

2. Incorporating deaths in the analysis of longitudinal outcomes

The National Research Council report states that missing data (e.g., on the ALSFRS-R) due to death are not considered missing data because it is not logical to consider what someone's function would have been had they not died. There are several approaches to incorporating death in the analysis of longitudinal outcomes. These approaches have been compared in a simulation study (419).

a. Approach 1: Analyze data on longitudinal outcomes using mixed effects models ignoring death as a cause for missing data. Missing data due to death are treated as “missing at random”. This assumption is unlikely in that assumes that the information that a patient will die is present in the measured ALSFRS values prior to their death. However, the bias with using this method may be quite small in short trials where the number of deaths is relatively small. In addition, there is surely some information in the sequence of ALSFRS values that are observed
before death. Of note is that the patients who drop out before death also present the same problem, as the mortality after a patient has dropped out is very high, indicating that the principal cause of a patient dropping out is disease progression.

This method loses the possibility of recovering an effect of treatment on death that is not mediated by its effect on the ALSFRS (function). In simulations the approach may yield the most statistical power in a study of short duration as mortality will be rare and will have a high coefficient of variation.

**b. Approach 2:** Use a joint model for death and the longitudinal outcomes; for instance, a shared parameter model where the random slope for the ALSFRS is considered to be a frailty in a survival model (492). This gives very similar results to those that treat death as withdrawal from the study (Approach 1) in shorter studies. The model includes a parameter for the direct effect of treatment on death not mediated through the frailty, but it is unclear how to combine this parameter with the treatment effect on the longitudinal outcome itself.

**c. Approach 3:** Use a combined endpoint that includes both death and the longitudinal outcome; for example, use the Joint Rank Test, a version of which, developed for ALS, is the Combined Assessment of Function and Survival (CAFS) (493). If the effect of treatment on death is mediated through its effect on function or the death rate is low, this will have lower power than the methods in Approaches 1 and 2; however it will have higher power if there is enough mortality and there is a direct effect of treatment on mortality.

The CAFS gives a p-value, but without an associated estimate and confidence interval. The use of the “win ration,” -- the proportion of patient who had a better outcome, as defined by the CAFS, over the proportion who had a worse outcome -- has been suggested; however, this statistic will depend on when the data are analyzed, typically ALSFRS early in the study and mortality later on. One should report the difference in both mortality rates over time and mean rates of ALSFRS decline, with the understanding that the significance of these statistics may not be consistent with the significance of the CAFS.

3. **Use of less stringent significance levels for hypothesis testing**

The traditional evidentiary standard for confirmatory clinical trials incorporates the use of an overall 5% significance level (two-tailed) in hypothesis testing. A lower standard (e.g., significance levels of 10%-20%) has sometimes been used in early-phase trials (436, 494) and can be justified by the need to eventually confirm the findings in Phase III trials.

Ideally, the choice of significance level and power for any trial should be based on the implications of making each type of error (Type I: rejecting the null hypothesis when the treatment has no effect; Type II: failing to reject the null hypothesis when the treatment has an effect). In confirmatory trials in lethal diseases with limited existing treatments, such as ALS, an argument can be made for the use of a less stringent significance level than the traditional 5% level, particularly for a treatment that does not appear to be associated with any major safety concerns. In this case, the consequences of making a Type I error may be small compared to the consequences of failing to identify an effective treatment.
E. Preclinical requirements – should they be different than for less serious illnesses?

The U.S. FDA has published several guidance documents on nonclinical studies needed to support initiation of clinical trials. Regulatory authorities in other jurisdictions generally follow these guidelines. Briefly, the FDA guidelines state that toxicology studies be performed in two species, one of which is non-rodent, with duration equal to or longer than the anticipated clinical studies and with maximum length of studies up to 6 months in rodents and 9 months in non-rodent species. Depending on the design of the phase 1 clinical studies, the initial toxicology studies could be from 6 to 12 weeks in duration. It is desired that the pharmacological agent be pharmacologically active in at least one of the toxicology studies to evaluate on-target toxicology. If the agent is not active in the preclinical species, a surrogate molecule that is active and has a similar profile as the therapeutic agent may need to be included toxicology studies. Chronic toxicology studies (9 to 12 months) are generally completed prior to start of phase 3 studies. In addition to standard toxicology studies, safety pharmacology studies as well as genotoxicity studies will likely be required prior to initiation of first in man studies. In some cases, reproductive toxicology studies may be deferred until later in development. As each drug candidate has unique properties and challenges, it is advisable to meet with appropriate regulatory authorities to clarify what non-clinical studies will be required prior to initiation of first in man studies.

Some therapeutic classes of molecules do not allow for standard nonclinical toxicology studies and therefore some flexibility in the design of the non-clinical toxicology program is granted. Examples include gene therapy, stem cells, protein and nucleic acid based therapeutics. There are guidance documents for each of these novel classes of therapeutics that should be reviewed.

Currently the regulatory agencies require similar non-clinical toxicology packages for severe disease with unmet medical need such as ALS as they do for less severe diseases, although exceptions may be granted if there is strong scientific rationale. The one exception is for anticancer agents, in which regulatory agencies commonly allow duration of exposures in patients to exceed duration of exposure in non-clinical toxicology studies. Given the severity of ALS and the desperate need for effective therapies, it could be argued that a similar position should be adopted for ALS.

The most effective way of expediting start of clinical trials is to hold a pre-IND meeting with the FDA or the regulatory body which has jurisdiction in the country in which the first-in-human study is being planned. The goal of the meeting is to discuss design of the initial clinical study and the planned non-clinical toxicology to support the clinical program. Achieving agreement on the non-clinical toxicology studies avoids regulatory delays in the program. Some countries such as the UK allow initiation of clinical studies based on draft non-clinical toxicology reports rather than finalized reports. This can save several months in starting a clinical study.
VIII. PUBLIC POLICY

A. General Comments

Drug development for a low-prevalence, rapidly progressive, and devastating disease such as ALS requires flexibility from regulatory and other governmental agencies. Policies that are particularly germane to ALS drug development include the FDA’s Expanded Access programs, as outlined in the Guidance on Expedited Programs for Serious Conditions (507); their support for the use of patient registries, as outlined in the Guidance on Good Pharmacovigilance Practices and Pharmacoepidemiologic Assessment (508); and the increasing commitment to patient-focused drug development, as embodied by the fifth authorization of the Prescription Drug Free User Fee Act (PDFUA V).

B. Expanded Access

1. Patient eligibility for expanded access

ALS is generally acknowledged to be an *immediately life threatening disease or condition*, as defined in 21 CFR 312.300 (b). For patients who have access to clinical research trials and who meet the enrollment criteria, participating in those trials is the best way to explore investigational medicines. Patients who don’t fit this category can be accommodated in *treatment-use* expanded access trials, which may be authorized after adequate human safety evidence and preliminary signs of efficacy are established for the particular drug. At the same time, there are important requirements including defining dose, safety and possibly efficacy that must be in place prior to granting expanded access and that relate to the need to have defined preliminary safety and efficacy data, such that the preliminary benefit-risk balance can be determined.

FDA acknowledges the advanced stage of disease suffered by many expanded access participants, which naturally indicates high mortality and co-morbidity rates. When reviewing safety data, FDA will make every effort to differentiate *serious adverse reactions* related to the investigational treatment from the large number of expected *serious adverse events* expected across a cohort of ALS sufferers participating in an expanded access trial.

2. Drug eligibility for expanded access

Due to the large scope of unmet medical need in ALS, and the high proportion of patients who are not candidates for research trials, an investigational drug that appears to be well tolerated in controlled phase 2 studies may warrant an expanded access trial of several hundred patients or more. Investigational products in clinical development for a different disease or condition, as well as products whose clinical development has occurred outside the U.S., may be eligible for expanded access trial in ALS, under 312.315(a).

Drug companies with products in clinical development for ALS will be asked in advance by FDA if they plan to explore group-level expanded access in the eventuality that the product becomes eligible for such a program.

The sponsor of an FDA authorized expanded access trial may utilize a third party such as a disease specific foundation, a collaboration of physicians, or a multi-disciplinary platform that specializes in responsibly designed group-level access programs.
3. Benefits of expanded access protocols for drug development

Group level expanded access trials under intermediate-sized investigational new drug (IND) protocols or treatment IND protocols can be implemented in ways that generate valuable data that can be used to identify possible responder/non-responder subpopulations and that can support the design of more targeted, high-powered phase 3 registration trials. Expanded access programs can also serve to gather long term safety and efficacy data while providing treatment access to patients during a drug’s registration process, continuing until pricing and reimbursement is established by payers.

C. National ALS Registry

The National ALS Registry is congressionally mandated and the first and only population-based registry for ALS patients in the United States. The Registry is being used for the recruitment of patients for clinical trials and epidemiological studies. Institutions in the United States and abroad have effectively used the Registry’s Research Notification Mechanism for enrolling patients in research. Over 96% of Registry enrollees have elected to be notified about new ALS research opportunities for which they may be eligible. ALS researchers currently can recruit patients on a state, regional, or national level with qualifiers such as age, sex, as well as date of diagnosis.

The agency encourages sponsors to explore the use of patient registries, including existing registries such as the National ALS Registry, as tools to enhance research and development and support regulatory decision-making. Registries can be used to:

- Improve recruitment of participants in clinical trials
- Identify possible cohorts for studies
- Assess benefit-risk and patient preferences
- Assist in conducting natural history studies
- Collect patient and caregiver reported outcomes, clinical data and post marketing data
- Collect biosamples
- Identify geographic locations for trial sites based on proximity of larger patient populations that meet enrollment criteria
- Stimulate research on the causes, treatments and outcomes of ALS
- Accelerate knowledge discovery and gain new insights from patients living with ALS

The immediate contribution of the Registry for clinical research is the recruitment of potential participants on behalf of clinical trials, other research notifications, and launch of the National ALS Biorepository in the fall of 2016 for the collection of pre and post-mortem bio-specimens. In the future, the registry is positioned to obtain data on patient preferences and to support benefit/risk assessment of potential treatments in ALS. Patient outcomes and post-marketing data also could be pursued via the Registry.

Additional information on the National ALS Registry is available at www.cdc.gov/als.
D. FDA Stakeholder Communication

ALS is usually a rapidly progressive, disabling and life shortening disease. Patients and families affected by ALS want and deserve communication about developing treatments that is accurate, timely and consistent across sponsors and regulatory agencies. This has not always been the case, as evidenced by some online discussions (e.g., http://blogs.wsj.com/pharmalot/2015/04/20/a-dispute-flares-over-data-for-a-nascent-als-drug-and-an-fda-review/). The clinicaltrials.gov website is an important means for communicating on these opportunities to the ALS community.

Recognizing the devastating nature of ALS, the FDA has a number of programs designed to speed drug development including Orphan Drug Designation, Fast Track Designation, Breakthrough Therapy Designation, Priority Review, Accelerated Approval and Expanded Access. The FDA encourages sponsors of ALS products to utilize these programs, and to accurately communicate both the application date and approval date for each program.

The FDA recognizes that patients with ALS want a greater voice in the drug development process. We encourage sponsors to work with patient groups, such as the NEALS Research Ambassadors, at all stages of drug development including protocol design.

E. FDA Accountability in Using Patient-oriented tools

1. Introduction

The 2012 Food and Drug Administration Safety and Innovation Act (FDASIA) provided new direction to the agency to expand the integration of patient perspectives in regulatory decision-making.

- Structured Benefit-Risk Assessment
- Patient Focused Drug Development
- Section 1137

2. Patient perspectives

This guidance and existing efforts establish positive direction for the evolution of patient-focused development and regulation of medical products for ALS.

Existing programs are not sufficient to inform FDA stakeholders about the range of opportunities for patient perspectives and data that could be useful in the development and regulation of medical products for ALS. However, in the fall of 2015 FDA signaled its intention to draft guidance for use by patient communities, researchers and drug developers to outline pragmatic and methodologically sound strategies, pathways, and methods to gather and use patient input. Such guidance will be helpful for advancing patient-focused drug development of therapies for ALS and other serious conditions with unmet medical need.

The manner in which FDA assesses and the degree to which patient information is utilized in the review of medical product development programs, regulatory approval and post-market surveillance will provide an important stimulus for external stakeholders to generate such
data and submit it to the agency as a supplement to specific product applications or through independent avenues

To encourage patient-focused medical product development and integration of patient perspectives in regulatory decision-making, FDA will include in its documentation of product reviews an assessment of patient engagement efforts, including an explanation of whether patient preference and patient-reported outcomes data were reviewed or examined as part of the decision to approve the product or issue a complete response letter.

FDA will also consider this type of data when reviewing sponsors’ submissions of regulatory packages for ALS-related therapies including but not limited to: target product profiles, clinical trial designs and data, qualification of drug development tools (including PROs), Risk Evaluation and Mitigation Strategies (REMS), labelling changes, annual safety reports and other regulatory submissions.

**F. Regulatory implications of targeted drug development, and development Incentives, including evidence for reimbursement**

ALS is a highly heterogeneous disease with multiple variants that remain unidentified. Therefore, the FDA recognizes that population-based trials may fail to measure a drug’s impact on distinct high-response subgroups. Drug developers may increasingly pursue adaptive clinical research strategies in an attempt to target and show therapeutic benefit in these patients. It is important to note, however, that a drug shown to be effective in one subpopulation of ALS may also benefit other ALS patients. Should a targeted trial generate compelling evidence of benefit in an ALS subgroup, the agency does not believe it is necessary to narrow the marketing indication solely to that subpopulation of patients unless the drug’s pharmacology is believed to exclusively benefit that subpopulation.

**XI. REFERENCES**


101 | Page


449. Papakostas GI, Fava M. Does the probability of receiving placebo influence clinical trial outcome? A meta-regression of double-blind, randomized clinical trials in MDD.


460. NINDS. Summary of Core Recommendations: Amyotrophic Lateral Sclerosis CDEs. 2015.

461. NINDS. Edinburgh Cognitive and Behavioural ALS Screen (ECAS) . 2015.


