**1. 연구의 필요성**

Amyotrophic lateral sclerosis(ALS)는 유럽통계로 10만명당 2~3명에 발병하는 비교적 드문 질환으로 upper motor neuron과 lower motor neuron이 degeneration이 되어 motor, extra-motor symptom을 나타내는 특징을 보이는 질환임. ALS의 임상적인 단계를 나누는 것은 환자를 치료하는데 자원을 효율적으로 사용할 수 있게 해줌. 예를 들면 초기에는 진단을 위한 service가 필요하고 중간단계에서는 관련된 여러과의 협업이 필요하며 후기에는 증상 완화를 위한 치료가 필요함. 그리고 환자와 보호자에게 설명할 때에도 이해를 돕는데 사용될 수 있으며 ALS에 관한 임상시험을 할 때 secondary endpoint로 사용될 수 있다.

기존에 발표되어 현재 널리 쓰이고 있는 clinical staging system의 대표적인 예로 King’s staging system(이하, King)과 ALS-Milano Torino staging(이하 MiToS)가 있음. 전자는 anatomical disease spread에 초점을 맞추고 있고 초기, 중기 진환을 감별하는 데에는 효과적이나 후기의 환자들은 상대적으로 명확하게 잘 구별해내지 못하며 후자는 질환의 진행에 따라서 기능을 상실하는 것에 초점을 맞추고 후기의 환자들을 세분화해서 구분하는데 상대적인 이점을 보임. 이러한 점은 ALS의 staging system에 좀 더 보완이 필요함을 의미한다고 볼 수 있음.

Revised ALS functional rating scale (ALSFRS-r)은 기존부터 임상현장에서 오래 사용되어 임상의 의사들에게 익숙하다는 장점이 있고 오랜기간 자료가 누적되어 있어 retrospective study시에도 사용될 수 있는 장점이 있음.

King, MiToS sensitivity, specificity

한국의 ALS 유병율, 사망률, care burden

## 2. 연구 목표

ALS의 진행에 따른 환자의 상태를 반영한 staging system을 만드는것.

King’s staging (이하 King), MiTos staging(이하 MiT)의 단점을 보완할 수 있는 새로운 staging system을 만드는 것이 목표.

ALSFRS-r에 기반하여 staging의 결과를 봤을 때 MiT이나 King보다 환자의 상태를 조금 더 잘 반영하는 결과값을 도출해 낼 수 있도록 하는것이 목표. 진료시에 ALSFRS를 f/u하면서 진행여부를 잘 보여줄 수 있는지 평가

## 3. 연구 내용 및 방법

hidden markov model의 개념을 이용한 machine learning을 통해서 stage를 예측하는 model을 구축.

ALSFRS의 데이터를 7:2:1정도의 비율로 training data set,  validation data set, evaluation data set로 나누고 training data set을 통해 model을 만듦.

overfitting등의 문제가 있는지를 검증하기 위해 validation data set을 통해서 예측률을 확인해서 적절한 model이 만들어졌는지 **검증**.

검증이 끝나고 적절한 model이 만들어졌다고 판단되면 evaluation data set을 model에 대입하여 stage를 잘 예측해 내는지 확인함.

1. 현재 초안은 0개월(초기상태)에서부터 18개월까지 매달 ALSFRS-r의 subscore(12항목)을 observable data로 하고 hidden state는 6개로 해서 분석을 하는 방향으로 구상. 이를 위해서 PROACT database의 ALSFRS revised 데이터셋을 정리해야 함.
   1. right censored data: 마지막으로 존재하는 ALSFRS-r score값과 동일한 값을 보인 다른 환자들 중 f/u score data가 있는 환자들의 결과를 참고. 그 중에서 최빈값으로 채워넣기가 잠정적 결론.
   2. f/u을 매달 하지 않아서 중간중간 score가 없는 환자들은 다음 f/u전까지는 동일한 score로 전부 채워넣기
      1. initial state probability는 1/6로 할 것인지? (hidden state가 6개이므로)
      2. transition probability를 1/(6-number of state+1) (예를 들면 2번째 stage라면 1/(6-2+1) 이므로 1/5로 할 것인지?: 2번째 stage에서 1번째 stage로 다시 좋아지는 것은 불가능 하므로 배제하고 2번째에서 2번째 stage로 그대로 유지되는 2→2와 악화되는 2→3, 2→4, 2→5, 2→6 을 모두 고려)
      3. emission probability는 HMM model을 통해 구할 수 있을지? 12차원의 변수들이 hidden state에 미치는 영향의 정도를 산술적으로 구하는 것은 굉장히 어려울 것으로 생각됨.

환자들마다 ALSFRS 기록 간격이 다름.

환자들마다 ALSFRS 기록 횟수도 다름.

환자들마다 추적 기간이 다름. 이런 문제점들을 극복할 방법이 필요.

Statistical analysis

HMM staging system의 validity(accuracy)와 reliability(concordance with other staging system)를 평가

1. accuracy: Kaplan-Meier survival analysis and log-rank test
2. accuracy: Cox regression test for homogeneity (differences in survival among same stage patients within each system), Cochran-Armitage test for discriminatory ability of each staging system (differences in survival among patients in different stages within each system) (reference: Ferraro, D., et al., Eur J Neurol, 2016)
3. reliability: intraclass correlation coefficient (ICC) between our method and other methods such as King’s staging method, MiTos staging method (reference: Balendra, R., et al., Amyotroph Lateral Scler Frontotemporal Degener, 2014)

## 4. 선행 연구 결과

## 기존에 HMM을 사용한 다른 질환 progression의 연구들의 결과, ALS에 머신러닝을 이용하여 stage를 나눴던 연구결과

## 5. 기대 효과 및 활용 방안

## Table 1. basic, demographical characteristics of patients (age, sex, ALS onset site, mean ALSFRS-r score, median ALSFRS-r score)

## Figure 1. baseline stage bar plot

## Figure 2. flowchart of ALS staging system using HMM and their definitions

## Table 2. median number of months and standardized median time (SMT) from onset to each stage (A) HMM stage, (B) King (C) MiToS

## Figure 3. box plot for SMT from onset to each stage in each staging system (A) HMM stage, (B) King (C) MiToS

## Table 3. Kaplan-Meier analysis of survival from each milestone (eg. Involvement of 2nd region, 3rd region)

## Figure 4. probability of transition between each stage in HMM staging

## Figure 5. emission probability of each ALSFRS-r subscore

## Table 4. discriminatory ability and homogeneity for prediction

임상현장에서 staging을 나누는데 사용해보고 transition probability를 환자와 보호자에게 설명하는데 사용. 예를 들면 stage 1,2라면 현재 상태는 초기상태이고 stage3,4/ stage 5,6으로 진행할 가능성이 ~%정도 입니다. 라는 식으로 활용가능. 임상시험에 환자를 포함시킬 때 초기, 중기, 후기 환자로 나눠서 등록하고 secondary end point로 활용하여 치료에 따른 호전이나 진행 속도를 늦췄는지를 평가하는데 사용 가능.

Pooled Resource Open-Access ALS Clinical Trials Database (PRO-ACT)

Estimated to be

Conducted by

On the basis of

Exert an action

Be implicated in

Impact

Proxy

Validate the ability

Confirm the robustness 견고성을 확인하다

Proxy of long-term survival

To be consistent with

Comparison was undertaken using system

For simplicity

Encoded with prefix A

Figure depict Kaplan-Meier curve for ~ survival

Comparison was undertaken using ~ system

Regression analyses were depicted on directed acyclic graphs

Analysis yielded ~ findings

Other approach drawing upon degree of general loss of function

We herein suggest

Relative trajectory(상대궤적) of subscore

Approach is explitation of topographical spread of neurodegeneration in ALS

Conceptually

In the context of clinical encounter

Low-burden applicability in clinical settings

Landscape

Era

Relentlessly 가차없이

Ensues 계속

Arbitrarily 임의로

Blurring of boundary

Apparent 명백한

Entities

Key difficulty

Comprise

Difficulty arising from

Overarching 압도적

We highlight the challenges raised by

Convey relevant information

Propose a systematic approach

Idiosyncratic 특이한

These problems are compounded by the use of phenotypic terms for diagnosis

Notwithstanding 그럼에도불구하고

Arguably

Perspective

Contiguous

Replicated 복제된

Propagated

Plausible 그럴듯한

Hierarchical cluster

Ascertain 명백히하다

Straightforward 똑바로

Discrete 이산

Ascribe ~에 돌리다

Signposted 표지판

Substantially 대체로

Admixed 혼합된

Antecedent 전례

Discernible 보고 알 수 있는

Ancestral 조상의

Seemingly 겉보기에