Crowdsourcing to identify new algorithms for amyotrophic lateral sclerosis therapies

November 4 2014

Credit: fotolia

Researchers including bioinformaticians from the Helmholtz Zentrum München and Ludwig-Maximilians-Universität München organize crowdsourcing challenge to identify new algorithms that will expedite the search for effective therapies for amyotrophic lateral sclerosis. In the latest issue of *Nature Biotechnology*, the scientists review the outcome of the crowd-sourcing exercise and describe the most effective algorithms that emerged from it.

For the great majority of patients, the diagnosis of amyotrophic lateral sclerosis (ALS) is tantamount to a death sentence. ALS is a neurodegenerative disease which attacks the nerve cells that control

voluntary muscle movements. Consequently, ALS patients suffer from muscular dystrophy, and continuing loss of mobility eventually leads to complete paralysis. Thus ALS progressively destroys the ability to speak, swallow and to breathe. Mean survival time after the appearance of overt symptoms generally ranges between 3 and 5 years. But the course of the disease is highly variable and, in a subset of patients, the disease progresses much more slowly. "The variation in disease progression makes it very difficult to develop effective treatment strategies," says Dr. Robert Küffner, who is at the Research Unit in Bioinformatics at LMU's Institute of Informatics, now affiliated at the Institute of Bioinformatics and Systems Biology at Helmholtz Zentrum München.

To generate new ideas that would facilitate clinical studies of ALS, Küffner, in collaboration with an international team of scientists, the ALS initiative Prize4Life, the IBM DREAM project and pharmaceutical companies, turned to the concept of crowd-sourcing. The challenge solicited participants to develop new algorithms that could more accurately predict the course of the disease in individual patients on the basis of defined clinical tests. This approach exploits computer-assisted procedures that can analyze complex datasets derived from large patient populations. Thus, algorithms that tease out patterns within the clinical data could support doctors to better predict disease progression and improve the precision of diagnosis.

Those who registered for the crowd-sourcing contest were given access to anonymized clinical data obtained from 1822 ALS patients during the first 3 months after diagnosis. The competitors were then asked to design algorithms predicting the condition of each patient 9 months later. Each algorithm submitted was evaluated by comparing the predicted with the actual state of each patient. "The two best algorithms outperformed the models used so far, and were substantially more reliable than the doctors' assessments," says Küffner. "These results will enable us to improve the design of future clinical studies. If the expected

course of the disease can be confidently predicted for each patient, it becomes much easier to evaluate the effects of candidate drugs." In addition, the contestants identified clinical parameters that had not previously been considered informative in this context – such as blood pressure and blood levels of uric acid and creatinine – as useful indicators of the rate of disease progression. "The results demonstrate the great potential of the crowd-sourcing approach for advancing research on ALS," says Küffner, who acted as technical coordinator for the competition.

More data, new contest

The anonymized patient data used had been obtained during various clinical studies, and included analyses of blood composition and assessments of motor function. "The problem is that these data are quite complex and also non-uniform, because they were measured by different laboratories," says Küffner. He therefore harmonized the data to minimize the effects of these variations before sending them out to the contestants. In all, 37 algorithms were submitted for evaluation. "Many of the contributors had had no previous experience with ALS, but their methods can be applied to clinical data. Crowd-sourcing thus opens up entirely new perspectives for clinical research, and can certainly help to identify the most powerful algorithms," says Küffner. Neither of the two prize-winning algorithms had previously been employed in a clinical context. Since both predictive models proved to be more reliable than the prognoses given by experienced ALS clinicians, the consortium is now planning to translate the methods into clinical practice – "so that the benefits are directly passed on to patients," Küffner says.

The crowd-sourcing contest was made possible by the ALS initiative Prize4Life, which created the Open Access Database ALS Clinical Trials (in cooperation with the Neurological Clinical Research Institute at Massachusetts General Hospital) and donated the prize money of

50,000 euros offered for the best algorithm. The platform now comprises clinical data from a total of 8600 patients. "This larger dataset will give us a better understanding of ALS and the factors that determine its clinical course," says Küffner, who is already organizing a second crowd-sourcing challenge which will utilize the extended database.

More information: Küffner, R. et al. (2014). "Crowdsourced analysis of clinical trial data to predict amyotrophic lateral sclerosis progression," *Nature Biotechnology*, DOI: 10.1038/nbt.3051

Provided by Helmholtz Zentrum München