Validation of the Clinical Assessment Scale for Autoimmune Encephalitis: A Multicenter Study

Meng-Ting Cai · Qi-Lun Lai · Yang Zheng · Gao-Li Fang ·

Song Qiao · Chun-Hong Shen · Yin-Xi Zhang 🅳 · Mei-Ping Ding

ABSTRACT

Introduction: A new scale, named the Clinical Assessment Scale for Autoimmune Encephalitis (CASE), has recently been developed for rating the severity of autoimmune encephalitis (AE) with a high level of clinimetric properties. In this study, our primary objective was to validate the performance of CASE through a multicenter study in China.

M.-T. Cai · Y. Zheng · C.-H. Shen ·
Y.-X. Zhang (✉) · M.-P. Ding (✉)
Department of Neurology, Second Affiliated Hospital, School of Medicine, Zhejiang University, 88 Jiefang Road, Hangzhou 310009, China
e-mail: zyx-neurology@zju.edu.cn

M.-P. Ding
e-mail: dmp-neurology@zju.edu.cn

Q.-L. Lai · S. Qiao
Department of Neurology, Zhejiang Hospital, Hangzhou, China

G.-L. Fang
Department of Neurology, Zhejiang Chinese Medicine and Western Medicine Integrated Hospital, Hangzhou, China

Methods: Between July 2014 and December 2019, 143 consecutive patients with definite neuronal surface antibody-associated AE from three tertiary hospitals were enrolled in the study. We validated the reliability, internal consistency, and validity of CASE. We further compared CASE with the modified Rankin scale (mRS) among different subtypes of AE in terms of its sensitivity to disease dynamics. Statistical analyses were performed using GraphPad Prism and R software.

Results: Our analyses showed that CASE had good inter- and intraobserver reliability (intraclass correlation coefficient 0.96/0.98) and internal consistency (Cronbach $\alpha = 0.847$) at disease onset. The scores of CASE and mRS remained well correlated in patients at admission and at discharge (both $r = 0.80$, $p < 0.001$). From admission to discharge, the scores of CASE changed in 81 (56.6%) patients, in comparison to changes in mRS in 48 (33.6%) patients ($p = 0.007$ and $p < 0.001$, respectively). The largest changes in scores occurred for non-motor symptoms, including psychiatric, memory, and language dysfunctions (40.6, 26.6, and 23.1% of patients, respectively); in contrast, scores for motor symptoms, such as dyskinesia, weakness and ataxia, changed the least (7.0, 15.4, and 16.1% of patients, respectively).

Conclusion: Based on these results, CASE performed well in assessing the severity of neuronal surface antibody-associated AE. In comparison to mRS, it performed better for

non-motor symptoms and was more sensitive to changes in severity.

Key Summary Points
Why carry out this study?
A new scale, named the Clinical Assessment Scale for Autoimmune Encephalitis (CASE), for rating the severity of autoimmune encephalitis (AE) with a high level of clinimetric properties has recently been developed.
The aim of this study was to validate the performance of CASE through a a multicenter study in China.
What was learned from the study?
CASE is an effective clinical scale that can be utilized to evaluate patients with neuronal surface antibody-associated AE.
CASE has advantages over the modified Rankin scale in terms of better assessing non-motor symptoms and being more sensitive to changes in severity.

INTRODUCTION

Autoimmune encephalitis (AE) is a group of disorders associated with specific antibodies targeting neuronal intracellular proteins, synaptic receptors, ion channels, and/or neuronal surface proteins [1]. Anti-N-methyl-D-aspartate receptor (NMDAR) encephalitis, anti-leucine-rich glioma-inactivated 1 (LGI1) encephalitis, and anti-gamma-aminobutyric acid B receptor (GABA$_B$R) encephalitis are the most common subtypes of neuronal surface antibody-associated AE [2–4]. AE is characterized by a wide spectrum of clinical manifestations mediated by different antibodies, including seizures, loss of memory, psychosis,

changes in behavior, and others. Disease severity in the acute phase is an important prognostic factor in AE [5–7], with severity predominantly assessed according to clinical signs and symptoms. The modified Rankin scale (mRS) [8], initially designed to measure disability after stroke, is commonly used to score severity in AE. However, mRS predominantly captures the impact of motor deficits on functional independence, with less emphasis on non-motor symptoms such as memory and psychiatric symptoms, which are of great importance in AE [7]. Thus, more reliable and accurate clinical scales are needed.

A new scale, named the Clinical Assessment Scale for Autoimmune Encephalitis (CASE), has recently been developed to rate the severity of AE [9]. As the first specialized clinical scale for AE, CASE was designed to address the deficiencies of mRS through the use of a more comprehensive set of scoring items, including motor functions, memory deficit, seizures, speech disorders, abnormal movements, decreased consciousness, and cerebellar ataxia. Nevertheless, there have been problems in the clinical application of CASE in practice [10–13]. For example, the definition of "controlled seizures" was uncertain in the case of seizure-free currently but having seizures previously, or in the condition of a first seizure before medication; in addition, some items, such as the psychiatric and memory symptoms, can be difficult to assess when the patient is in coma. At present, the performance of CASE among different AE subtypes and patient populations remains unknown. Therefore, the aim of this study was to validate the performance of CASE in a cohort of Chinese adults with definite neuronal surface antibody-associated AE. We also explored its sensitivity to changes in severity and performance among patients with AE associated with different neuronal surface antibodies with a focus on improving the application of this tool.

METHODS

Patients

This retrospective study included 143 consecutive patients from three tertiary hospitals (Second Affiliated Hospital School of Medicine

Zhejiang University, Zhejiang Hospital, and Zhejiang Chinese Medicine and Western Medicine Integrated Hospital) in China between July 2014 and December 2019. All patients were at first onset and had been diagnosed with definite neuronal surface antibody-associated AE according to published criteria [14], including 96 with anti-NMDAR encephalitis, 26 with anti-LGI1 encephalitis, 17 with anti-GABA$_B$R encephalitis, two with anti-dipeptidyl-peptidase-like protein-6 (DPPX) encephalitis, one with anti-alpha-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptor (AMPAR) encephalitis, and one with anti-contactin-associated protein-like 2 (CASPR2)-associated encephalitis. The exclusion criteria were: (1) combination with other antibodies against neuronal and glia antigens; (2) disease complicated by central nervous system infections (e.g., neurosyphilis, viral encephalitis, and others); (3) presence of other severe neurological or psychiatric complications (e.g., brain tumor, stroke, myasthenia gravis, and others). The demographic information and clinical characteristics of all patients were collected. Autoimmune antibodies in the cerebrospinal fluid (CSF) and/or collected serum samples were detected by cell-based assays prior to the initiation of immunotherapies. We evaluated the performance of CASE with the total cohort and then performed subgroup analyses (Fig. 1).

This research was approved by the local ethics committee of each participating center (Second Affiliated Hospital School of Medicine Zhejiang University; [approval number: 2019-082], Zhejiang Hospital [approval number: 2021-9K], and Zhejiang Chinese Medicine and Western Medicine Integrated Hospital [approval number: 2021-021]). The study was conducted in accordance with the Helsinki Declaration of 1964 and its later amendments. The requirement to obtain patient consent was waived for this retrospective study.

Scale Assessment

The CASE and mRS were assessed simultaneously on admission (before treatment) and at discharge. CASE was developed based on diverse syndromes, including definite AE, definite autoimmune limbic encephalitis, autoantibody-negative but probable AE, definite acute disseminated encephalomyelitis, and definite and probable brainstem encephalitis [9]. As a result, nine items were selected for measurement, including seizures (current status), memory dysfunction, psychiatric symptoms (delusion, hallucination, disinhibition, aggression), consciousness, language problem, dyskinesia/dystonia, gait instability and ataxia, brainstem dysfunction, and weakness. The total maximum score was 27. Each item was based on a 3-point grading system, with the exception of the item "brainstem dysfunction", which consisted of gaze paresis, tube feeding, and ventilator care due to hypoventilation [Electronic Supplementary Material (ESM) Table 1] [9]. When clinical applying of CASE was undetermined, we communicated with the original developers for additional details, which are summarized as follows [10–13]: (1) CASE was evaluated at the current time in any disease phase; (2) the items of memory and language performance were evaluated mainly by communications and observations, rather than by neurological examinations; (3) for the item of seizures: controlled seizures with no need of dose-up were scored as 1 (i.e., patients without any seizures over the last month receiving antiepileptic drugs), and intractable seizures with the need of dose-up were scored as 2; (4) in comatose patients, the items of seizure, dyskinesia/dystonia, and brainstem dysfunction could be used for evaluation, whereas all others were scored as 3. Two neurologists (M-TC and Q-LL) who were blinded to the diagnosis evaluated the scales independently by studying the detailed medical records described by independent neurologists and nurses, retrospectively. M-TC repeated the evaluation 1 month later.

Statistical Analysis

Statistical analyses were performed using GraphPad Prism (version 8; GraphPad Software Inc., San Diego, CA, USA) and R (version 4.0.2; https://cran.r-project.org/bin/windows/base/) software. The results were described as
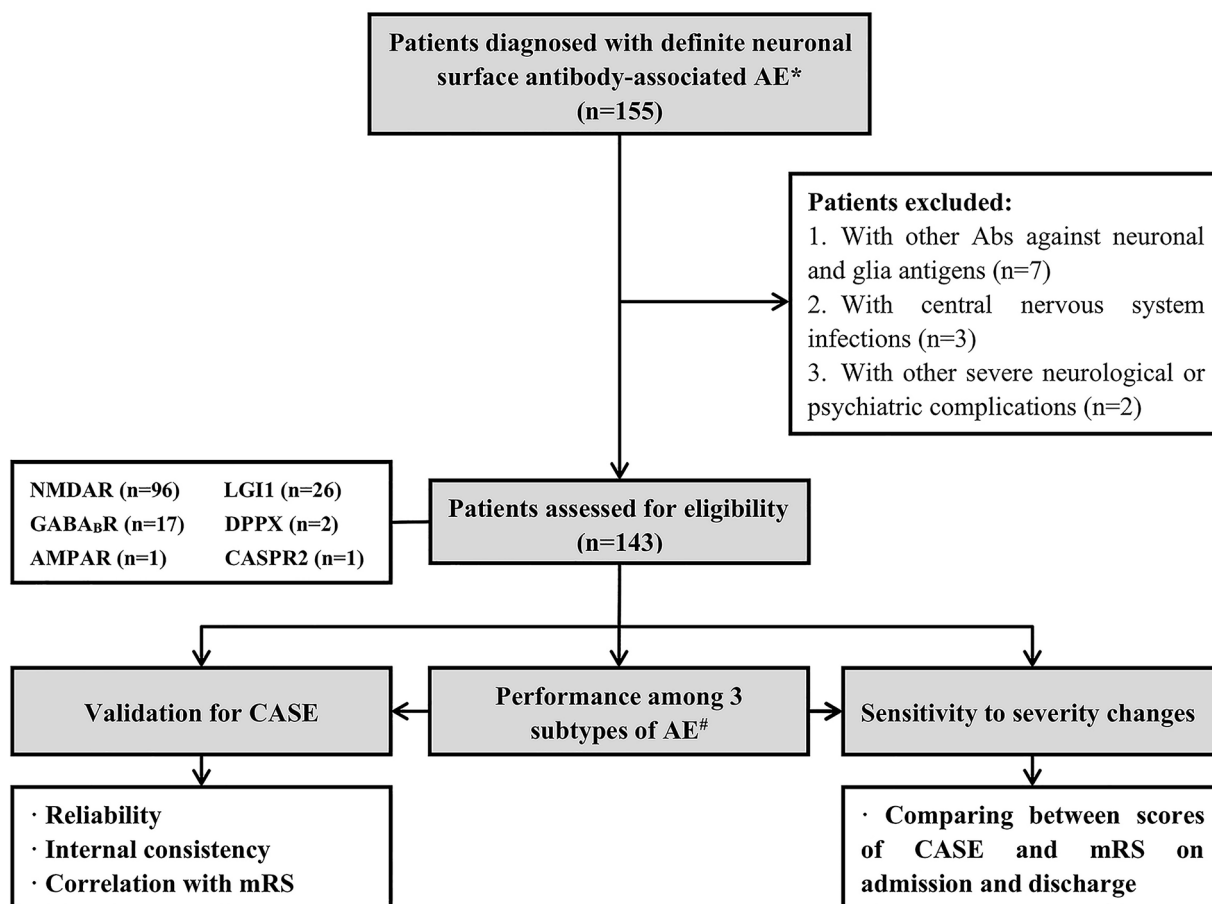
Fig. 1 Flow chart of study design. Asterisk (*) indicates diagnosis based on the criteria by Graus et al. [14]. Hash sign (#) indicates anti-NMDAR encephalitis, anti-LGI1 encephalitis, and anti-GABA$_B$R encephalitis. *Ab* Antibody, *AE* autoimmune encephalitis, *AMPAR* anti-alpha-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptor encephalitis, *CASE* Clinical Assessment Scale for Autoimmune Encephalitis, *CASPR2* anti-contactin-associated protein-like 2-associated encephalitis, *DPPX* anti-dipeptidyl-peptidase-like protein-6 encephalitis, *GABA$_B$R* anti-gamma-aminobutyric acid B receptor encephalitis, *LGI1* anti-leucine-rich glioma-inactivated 1 encephalitis, *mRS* the modified Rankin scale, *NMDAR* anti-*N*-methyl-D-aspartate receptor encephalitis

percentages, medians, and interquartile ranges (IQRs). The Kolmogorov–Smirnov test was performed to evaluate the distribution of the variables. Categorical data were analyzed by the Chi-square test or Fisher's exact test. Analysis of continuous data was by the t test, Mann–Whitney U test, or Wilcoxon matched-pairs signed-rank test between two groups, and by one-way analysis of variance or the Kruskal–Wallis H test among three groups. The interobserver and intraobserver reliability was calculated by weighted kappa (κ) statistics for each item and the intra-class correlation coefficient (ICC) for the total scores, with the strength of agreement indicated by excellent (κ = 0.81–1.00, ICC = 0.90–1.00), substantial (κ = 0.61–0.80, ICC = 0.75–0.90), moderate agreement (κ = 0.41–0.60; ICC = 0.50–0.75), and fair to slight (κ = 0–0.40; ICC = 0–0.50) [15, 16]. Cronbach α was computed to ascertain the internal consistency, with α > 0.70 considered to indicate good reliability [17]. The corrected item-total correlations > 0.40 indicated good homogeneity of each item [18]. As

suggested by Lim et al. [9], we examined the validity of CASE by analyzing its Spearman's rank correlation with mRS given an absence of a gold standard for AE evaluation. The cutoff point > 0.9 indicated a strong relationship [19]. $P < 0.05$ indicated statistical significance.

RESULTS

Clinical Manifestations of Neuronal Surface Antibody-Associated AE

The characteristics of patients are summarized in Table 1. The median age of the 143 patients (58.7% males) with neuronal surface antibody-associated AE patients included in the study was 36.0 (IQR 23.5–57.0, range 16.0–81.0) years. The median duration between disease onset and admission was 14.0 (IQR 6.0–30.0, range 1.0–365.0) days, indicating that all patients were in the acute phase of the disease. The median interval from admission to diagnosis was 3.0 (IQR 2.0–5.0, range 0.0–24.0) days. Subsequent immunotherapies, including intravenous methylprednisolone pulse therapy, intravenous immunoglobulin, and plasma exchange, were initiated at a median interval of 4.0 (IQR 2.0–6.0, range 0.0–19.0) days after admission. Second-line immunotherapies, such as rituximab, were considered for patients with poor response to the first-line treatment. However, ten (7.0%) patients rejected these immunotherapies due to their poor financial situation, the minor severity of their disease, or the side-effects of drugs. Fifty-six (39.2%) of all patients received empiric treatment prior to diagnosis given the high suspicion of AE. The median time from treatment to discharge was 12.0 (IQR 7.0–19.0, range 2.0–114.0) days. Of the nine clinical symptoms described in CASE, upon admission the most common symptom was seizures (93/143, 64.3%), followed by psychiatric symptoms (80/143, 55.9%) and memory dysfunction (74/143, 51.7%); dyskinesia/dystonia was the least common presenting symptom (11/143, 7.7%). Comparison of the most common subtypes of neuronal surface antibody-associated AE, i.e., $GABA_BR$, LGI1, and NMDAR, showed that patients with anti-

NMDAR encephalitis presented with a lower frequency of seizures ($p = 0.023$, $p = 0.002$) and memory dysfunction ($p = 0.040$, $p = 0.016$) and that those with anti-LGI1 encephalitis presented with a lower frequency of language problems ($p = 0.015$, $p = 0.042$), as shown in Fig. 2. Among patients who underwent paraclinical examinations, 51.0% (73/143) had abnormal results on magnetic resonance imaging, 93.8% (90/96) had abnormal electroencephalogram results, and 49.7% (71/138) had CSF white blood cell count > 5/μL.

Validation and Explanation of CASE

We validated the performance of CASE at admission, including the reliability, internal consistency, and correlation with mRS. The interobserver and intraobserver reliability of the total scores demonstrated excellent agreement (0.96 and 0.98, respectively) (details shown in ESM Table 2). The internal consistency of CASE was good (Cronbach $\alpha = 0.847$), and the corrected item-total correlations of all items, with the exception of the items seizure and dyskinesia, were > 0.40 (details shown in ESM Table 3). The correlation between the total CASE score and mRS score at admission was only fair (both $r = 0.80$, $p < 0.001$). Change in the CASE score relative to mRS is shown in Fig. 3a.

CASE as a Sensitive Scale for Changes in Severity

The score distribution corresponding to CASE and mRS at discharge and admission are shown in Fig. 3. We used Sankey diagrams to visualize patients with different scores between admission and discharge (Fig. 4). These resulted in the scores of some patients assessed by these two scales remaining unchanged from admission to discharge and those of others becoming even worse. To better illustrate the changes in severity, we used the terms deterioration, stabilization, and remission to denote an increase, no change, and a decrease in assessment scores, respectively. As shown in Table 2, CASE identified 22 (15.4%), 62 (43.4%), and 59 (41.2%)

Table 1 Demographic and clinical characteristics of patients with autoimmune encephalitis enrolled in the study

Demographic and clinical manifestations	Total patient population (*n* = 143)	Anti-NMDAR encephalitis (*n* = 96)	Anti-LGI1 encephalitis (*n* = 26)	Anti-GABA$_B$R encephalitis (*n* = 17)	Other subtypes[a] (*n* = 4)
Gender, female:male (*n*)	59:84	50:46	7:19	2:15	4:0
Age at onset (years), median (IQR)	36.0 (23.5–57.0)	27.5 (21.8–42.0)	59.5 (45.3–64.0)	56.0 (54.0–64.0)	61.5 (52.5–65.5)
Disease duration (days), median (IQR)					
From onset to admission	14.0 (6.0–30.0)	14.0 (6.0–21.5)	20.0 (10.0–60.0)	10.0 (6.0–20.0)	13.5 (5.5–22.5)
From admission to diagnosis	3.0 (2.0–5.0)	3.0 (2.0–5.0)	3.0 (2.0–3.8)	3.0 (2.0–4.0)	3.5 (2.3–4.3)
From admission to treatment	4.0 (2.0–6.0)	3.0 (2.0–5.0)	4.0 (3.0–7.0)	2.5 (1.0–4.3)	7.0 (6.0–11.0)
From treatment to discharge	12.0 (7.0–19.0)	14.0 (8.0–21.0)	9.0 (6.0–12.0)	12.0 (9.0–16.0)	6.0 (5.0–14.0)
Clinical features, *n* (%)					
Seizure	92 (64.3)	54 (56.3)	21 (80.8)	16 (94.1)	1 (25.0)
Memory dysfunction	74 (51.7)	41 (42.7)	17 (65.4)	13 (76.5)	3 (75.0)
Psychiatric symptoms	80 (55.9)	57 (59.4)	12 (46.2)	8 (47.1)	3 (75.0)
Consciousness	23 (16.1)	17 (17.7)	2 (7.7)	2 (11.8)	2 (50.0)
Language problem	40 (28.0)	30 (31.3)	2 (7.7)	6 (35.3)	2 (50.0)
Dyskinesia/dystonia	11 (7.7)	8 (8.3)	2 (7.7)	0 (0.0)	1 (25.0)
Gait instability and ataxia	35 (24.5)	26 (27.1)	2 (7.7)	4 (23.5)	3 (75.0)
Brainstem dysfunction	10 (7.0)	8 (8.3)	0 (0.0)	0 (0.0)	2 (50.0)
Weakness	29 (20.3)	23 (24.0)	1 (3.8)	3 (17.6)	2 (50.0)
Admission to ICU, *n* (%)	16 (11.2)	14 (14.6)	0 (0.0)	2 (11.8)	0 (0.0)
Presence of antibody, *n* (%)					
Serum	36 (25.2)	24 (25.0)[b]	7 (26.9)	4 (23.5)	1 (25.0)
CSF	131 (91.6)	89 (92.7)	23 (88.5)	16 (94.1)	3 (75.0)

Table 1 continued

Demographic and clinical manifestations	Total patient population ($n = 143$)	Anti-NMDAR encephalitis ($n = 96$)	Anti-LGI1 encephalitis ($n = 26$)	Anti-GABA$_B$R encephalitis ($n = 17$)	Other subtypes[a] ($n = 4$)
Both	24 (16.8)	17 (17.7)	4 (15.4)	3 (17.6)	0 (0.0)

CSF cerebrospinal fluid, *GABA$_B$R* gamma-aminobutyric acid B receptor, *ICU* intensive care unit, *IQR* interquartile range, *LGI1* leucine-rich glioma-inactivated 1, *NMDAR* N-methyl-D-aspartate receptor

[a] Including 2 anti-dipeptidyl-peptidase-like protein-6 (DPPX) encephalitis, 1 anti-alpha-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptor (AMPAR) encephalitis, and 1 anti-contactin-associated protein-like 2 (CASPR2)-associated encephalitis

[b] Of these 24 patients, seven had only serum-positive antibody, and the results were all confirmed by tissue immunohistochemistry, in addition to cell-based assay

patients in the categories of deterioration, stabilization, and remission, respectively; in comparison mRS identified nine (6.3%), 95 (66.4%), and 39 (27.3%) patients in the same categories, respectively. The scores of CASE and mRS at discharge were both significantly lower than those at admission ($p = 0.007$, $p < 0.001$). The scores of CASE and mRS remained well correlated at discharge ($r = 0.80$, $p < 0.001$).

During the disease course, the scores of non-motor item of psychiatric symptoms changed most frequently (40.6%), followed by changes in memory and language symptoms (26.6 and 23.1%, respectively), with the majority of changes indicating remission (32.2, 21.7, and 16.8%, respectively) (Table 2), although only the scores of the first two items were statistically different ($p = 0.002$, $p = 0.001$). On the contrary, the scores for motor items, such as dyskinesia, weakness, and ataxia, changed only slightly (7.0, 15.4, and 16.1%, respectively) (Table 2). Analyses aimed at exploring the potential influencing factors on the total scores of CASE and mRS, including gender, onset age, disease duration, treatment time, and length of stay (LOS), showed that only the scores at discharge were significant. Both scales had low correlations with LOS (both $r = 0.30$, $p < 0.001$), and mRS scores also showed a slightly negative correlation with the onset age ($r = -0.20$, $p = 0.034$) and initial treatment time ($r = -0.20$, $p = 0.012$).

Performance of CASE and mRS Among Three Subtypes of AE

CASE was developed on the basis of various syndromes in addition to definite AE, but its performance in different subtypes has remained unknown. We therefore explored the utility of CASE in the three most common subtypes of neuronal surface antibody-associated AE: anti-NMDAR, anti-LGI1, and anti-GABA$_B$R encephalitis.

Comparison of the scores of CASE among these three subtypes revealed that their respective median total score on admission was 3 (IQR 2.0–6.3), 3 (IQR 2.0–4.8), and 4 (IQR 3.0–5.0) ($p = 0.413$). As shown in Fig. 5, the scores of each item of CASE at admission were generally low. Seizures, memory dysfunction, and psychiatric symptoms were the most frequently occurring items, with the latter two accounting for the largest proportion of top scores; in contrast, dyskinesia/dystonia and brainstem dysfunction were rare presentation. Further analyses among these three subtypes showed that only the severity of seizures had a statistical difference between anti-NMDAR encephalitis and anti-LGI1 encephalitis ($p = 0.036$) and anti-GABA$_B$R encephalitis ($p = 0.016$) (ESM Table 4).

We also compared the efficacy of CASE and mRS as an assessment scale for AE mediated by different antibodies. As shown in Table 2, the largest change in total CASE and mRS scores occurred for anti-NMDAR encephalitis (60.4
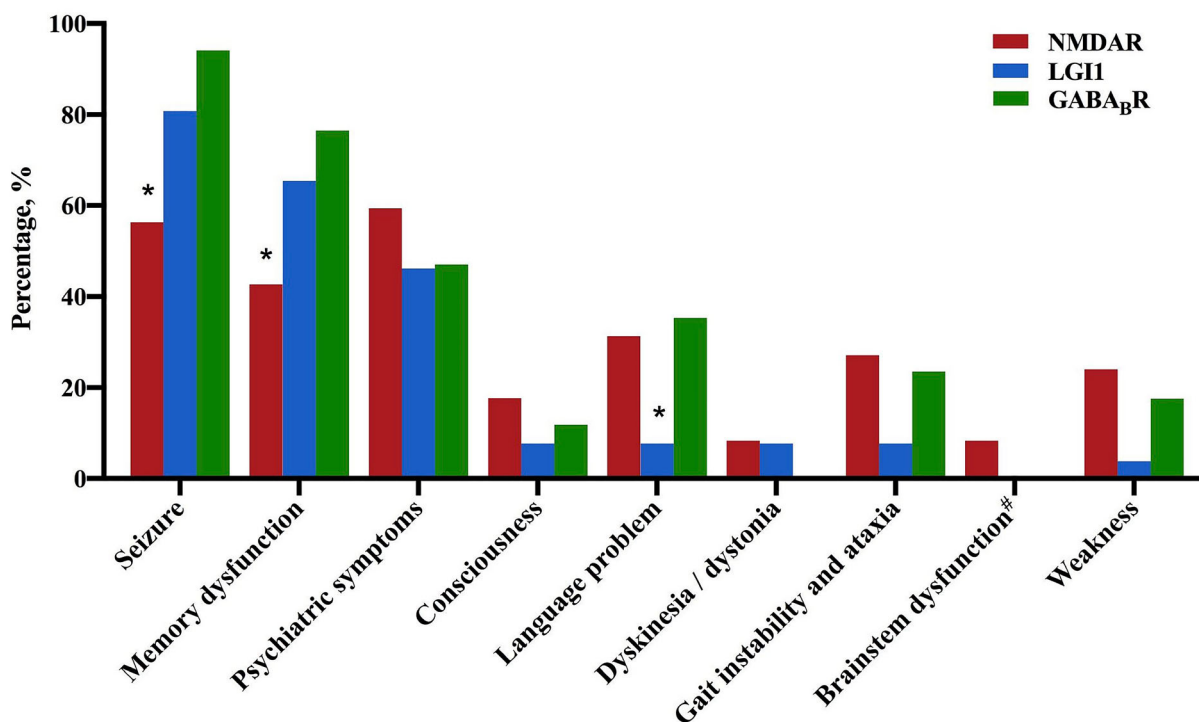
P value of the 3 significant symptoms by Chi-square test (^ Fisher's exact test).

	3 Cohorts	NMDAR v.s. LGI1	NMDAR v.s. GABA$_B$R	LGI1 v.s. GABA$_B$R
Seizure	0.002	0.023	0.002^	0.377^
Memory dysfunction	0.010	0.040	0.016^	0.513^
Language problem	0.042	0.022^	0.742	0.042^

* Disease with lower frequency comparing with the others.
\# Including gaze paresis, tube feeding, and ventilator care due to hypoventilation.

Fig. 2 Frequency of distributions of symptoms among three subtypes of autoimmune encephalitis: GABA$_B$R encephalitis, LGI1 encephalitis, and NMDAR encephalitis

and 34.4%, respectively), and the smallest change occurred for anti-LGI1 encephalitis (42.3 and 34.6%, respectively). Overall, CASE showed more obvious changes in disease severity than mRS among the three subtypes, especially in the items of psychiatric, language, and memory symptoms. Among these latter symptoms, more patients achieved remission than deterioration. The correlation between CASE and mRS scores differed in the three subtypes, with a descending order from anti-NMDAR encephalitis ($r = 0.84$, $p < 0.001$), anti-GABA$_B$R encephalitis ($r = 0.74$, $p = 0.001$) to anti-LGI1 encephalitis ($r = 0.64$, $p = 0.006$), possibly due

to the different distribution of symptoms in each subtype.

DISCUSSION

As the first severity assessment scale specifically designed for AE, CASE compensated for the deficiences of mRS in non-motor symptoms and became a more comprehensive and promising tool. However, its clinical application has been limited. In the study reported here, we attempted to validate CASE in a Chinese multicenter cohort of patients with neuronal surface antibody-associated AE and to provide
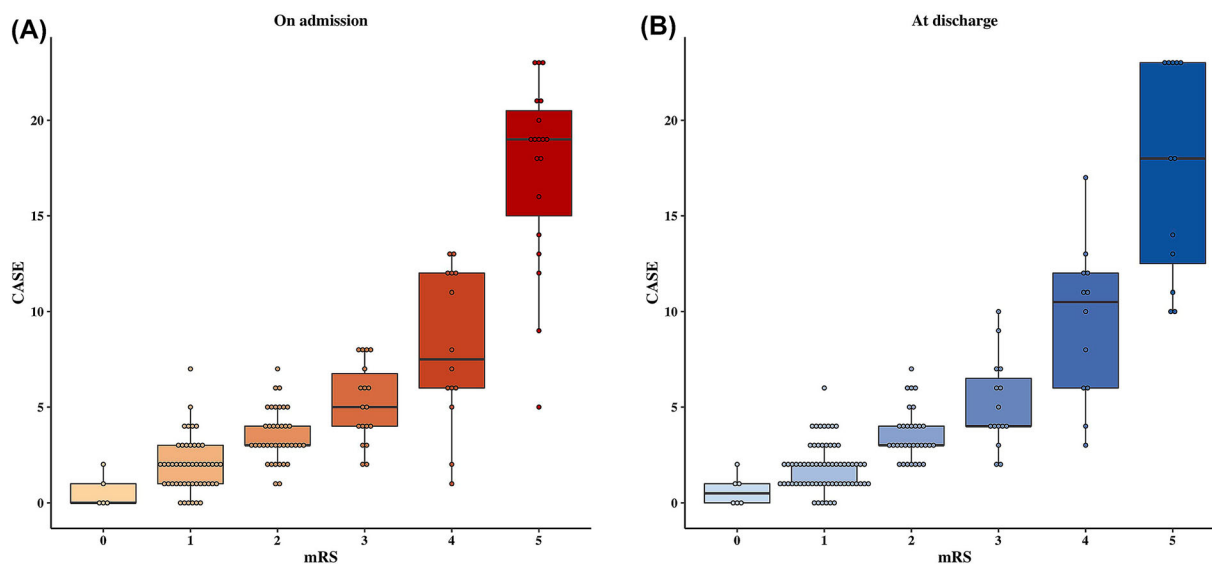
Fig. 3 The total scores of the Clinical Assessment Scale for Autoimmune Encephalitis (CASE) according to the modified Rankin scale (mRS) at admission (**a**) and discharge (**b**). Each box plot shows the distribution of CASE scores in a group with given mRS score. The construction is based around a dataset's quartiles, as the ends of each box mark the locations of the first and third quartiles (the colored "box"), respectively, and the center line represents the median. The upper and lower whiskers represent scores outside the middle 50%, and a score beyond the whiskers is considered to be an outlier. The gradient colors are inconsistent with the mRS scores

more details on CASE, with the aim to increase its readability and applicability. We noted that CASE performed well even in patients with mild symptoms, and in different subtypes of AE, and was more sensitive to changes in severity than mRS. The items of memory dysfunction, psychiatric symptom, and language problem showed a higher utility in terms of disease assessment, while dystonia and brainstem symptoms had lower applicability.

The results of our validation of CASE differ slightly from those of the original study, with a higher interobserver and intraobserver reliability of most items despite a lower internal consistency and correlation with mRS [9]. The explanation for these differences can be divided into subjective and objective factors.

Subjective aspects depend on individual differences among observers, reflecting directly on the reproducibility. On the basis of individual discrepancies in understanding and execution, a good reproducibility demands detailed instructions for a more accurate evaluation [20], especially in the case of frequent and controversial AE symptoms such as seizures, memory, and language symptoms. Our study provides more details that will enable observers to reach an agreement by reducing misunderstandings of the items, especially for these latter three symptoms. Above all, the reliability of all items, with the exception of seizure, dyskinesia, brainstem dysfunction, and weakness, were very good (> 0.80), with the lower reliability for the latter four symptoms possibly linked to the lower number of patients [15]. Larger cohorts and more observers will be required for further validation.

Objective factors can be attributed to differences in the disease profile. CASE was developed based on diverse syndromes, including acute disseminated encephalomyelitis and brainstem encephalitis which contribute to the major component of gait instability and brainstem dysfunction [9]. Moreover, differences in the structure of healthcare services in China should be considered: the choice of hospital of our Chinese patients was directly associated to national healthcare policy, medical resources,
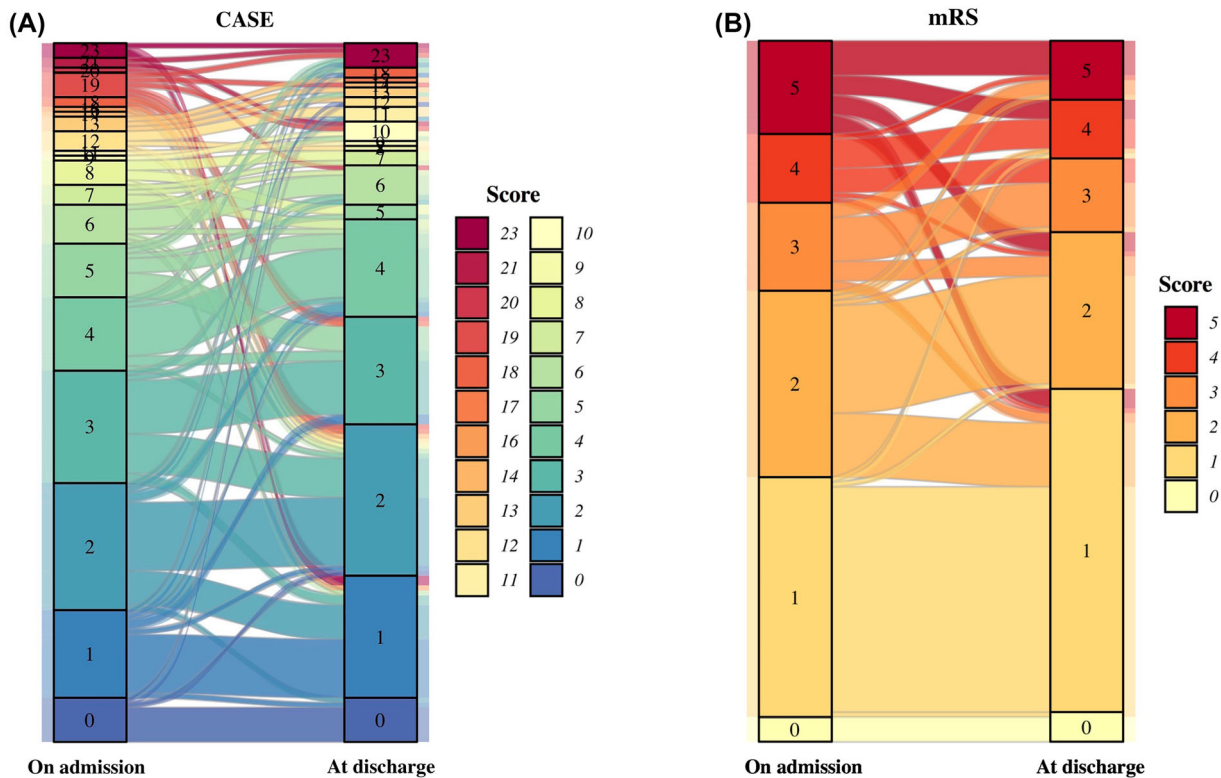
Fig. 4 Sankey diagrams for visualizing the change in CASE scores (**a**) and mRS scores (**b**) from admission to discharge. For each diagram, the lodes in gradient colors represent the different scores in the two sidebars (admission and discharge). The flow lines between the two sidebars represent changes in individual scores. The height of the colored lodes and line width of the flow lines are proportional to the numbers of patents

individual trust and satisfaction in institutions, and other rather complex factors [21–24]. In contrast, the referral system might be stricter in developed countries where tertiary hospitals are inclined to accept critically ill patients [25]. As a result, the initial phase of the disease may have been missed in patients in the original Korean cohort, with the patients referred to the tertiary hospital only after significant disease progression [9]. In comparison, the disease severity of our patients at admission was varied, and a large proportion of the cohort was admitted directly into hospital without a referral. Hence, these patients commonly started with mild symptoms and the disease score indicated a relatively lower severity. Indeed, only 16 (11.2%) patients in our study required intensive care unit (ICU) admission. Overall, CASE, with its multidimensional items, demonstrated values that were able to discriminate patients with various manifestations. However, CASE need to be used cautiously in AE patients with mono-symptomatic presentation, as this is an inherent limitation of a scale across multiple domains.

Compared with mRS, the advantages of CASE in assessing disease severity are reflected in the following two aspects. Firstly, CASE provided a comprehensive assessment with various symptoms by using a more detailed and specific score system. CASE had widely distributed scores within the same mRS score and was especially useful in assessing the non-motor symptoms in AE, such as memory and psychiatric dysfunctions, which occurred frequently or were sometimes the only clinical presentation. Secondly, CASE could be more helpful for comparing the change in severity at different stages of the disease, as supported by the

Table 2 An overview of the Clinical Assessment Scale for Autoimmune Encephalitis and modified Rankin scale scores at admission and discharge and changes in the respective scores between admission and discharge

CASE and mRS	CASE score at admission, median (IQR)	CASE score at discharge, median (IQR)	p^a	Change in severity by deviation of scores, n (%)		
				Deterioration	Stabilization	Remission
Total cohort ($n = 143$)						
mRS	2.0 (1.0–3.0)	1.0 (1.0–3.0)	< 0.001	9 (6.3)	95 (66.4)	39 (27.3)
CASE (total score)	3.0 (2.0–6.0)	3.0 (2.0–4.5)	0.007	22 (15.4)	62 (43.4)	59 (41.3)
Seizure	1.0 (0.0–1.0)	1.0 (0.0–1.0)	0.130	14 (9.8)	125 (87.4)	4 (2.8)
Memory dysfunction	1.0 (0.0–2.0)	0.0 (0.0–1.5)	0.001	7 (4.9)	105 (73.4)	31 (21.7)
Psychiatric symptoms	2.0 (0.0–2.0)	0.0 (0.0–2.0)	0.002	12 (8.4)	85 (59.4)	46 (32.2)
Consciousness	0.0 (0.0–0.0)	0.0 (0.0–0.0)	0.144	5 (3.5)	123 (86.0)	15 (10.5)
Language problem	0.0 (0.0–1.0)	0.0 (0.0–0.0)	0.060	9 (6.3)	110 (76.9)	24 (16.8)
Dyskinesia/dystonia	0.0 (0.0–0.0)	0.0 (0.0–0.0)	0.227	7 (4.9)	133 (93.0)	3 (2.1)
Gait instability and ataxia	0.0 (0.0–0.0)	0.0 (0.0–0.0)	0.168	9 (6.3)	120 (83.9)	14 (9.8)
Brainstem dysfunction	0.0 (0.0–0.0)	0.0 (0.0–0.0)	0.164	11 (7.7)	127 (88.8)	5 (3.5)
Weakness	0.0 (0.0–0.0)	0.0 (0.0–0.0)	0.408	9 (6.3)	121 (84.6)	13 (9.1)
Anti-NMDAR encephalitis ($n = 96$)						
mRS	2.0 (1.0–4.0)	2.0 (1.0–3.0)	0.039	8 (8.3)	63 (65.6)	25 (26.0)
CASE (total score)	3.0 (2.0–6.3)	3.0 (2.0–6.0)	0.088	18 (18.8)	38 (39.6)	40 (41.7)
Seizure	1.0 (0.0–1.0)	1.0 (0.0–1.0)	0.074	9 (9.4)	85 (88.5)	2 (2.1)
Memory dysfunction	0.0 (0.0–2.0)	0.0 (0.0–1.0)	0.033	5 (5.2)	74 (77.1)	17 (17.7)
Psychiatric symptoms	2.0 (0.0–2.0)	1.50 (0.0–2.0)	0.016	8 (8.3)	57 (59.4)	31 (32.3)
Consciousness	0.0 (0.0–0.0)	0.0 (0.0–0.0)	0.238	4 (4.2)	81 (84.4)	11 (11.5)
Language problem	0.0 (0.0–2.0)	0.0 (0.0–1.0)	0.207	8 (8.3)	70 (72.9)	18 (18.8)

Table 2 continued

CASE and mRS	CASE score at admission, median (IQR)	CASE score at discharge, median (IQR)	p^a	Change in severity by deviation of scores, n (%)		
				Deterioration	Stabilization	Remission
Dyskinesia/dystonia	0.0 (0.0–0.0)	0.0 (0.0–0.0)	0.182	6 (6.2)	88 (91.7)	2 (2.1)
Gait instability and ataxia	0.0 (0.0–1.0)	0.0 (0.0–0.0)	0.459	8 (8.3)	78 (81.2)	10 (10.4)
Brainstem dysfunction	0.0 (0.0–0.0)	0.0 (0.0–0.0)	0.149	9 (9.4)	83 (86.5)	4 (4.2)
Weakness	0.0 (0.0–0.0)	0.0 (0.0–0.0)	0.736	7 (7.3)	79 (82.3)	10 (10.4)
Anti-LGI1 encephalitis (n = 26)						
mRS	1.0 (1.0–2.0)	1.0 (1.0–1.0)	0.006	0 (0.0)	17 (65.4)	9 (34.6)
CASE (total score)	3.0 (2.0–4.8)	2.0 (1.0–3.0)	0.003	0 (0.0)	15 (57.7)	11 (42.3)
Seizure	1.0 (1.0–1.0)	1.0 (1.0–1.0)	0.850	3 (11.5)	22 (84.6)	1 (3.8)
Memory dysfunction	1.50 (0.0–2.0)	1.0 (0.0–1.0)	0.005	0 (0.0)	17 (65.4)	9 (34.6)
Psychiatric symptoms	0.0 (0.0–2.0)	0.0 (0.0–1.0)	0.007	0 (0.0)	17 (65.4)	9 (34.6)
Consciousness	0.0 (0.0–0.0)	0.0 (0.0–0.0)	0.346	0 (0.0)	24 (92.3)	2 (7.7)
Language problem	0.0 (0.0–0.0)	0.0 (0.0–0.0)	0.346	0 (0.0)	24 (92.3)	2 (7.7)
Dyskinesia/dystonia	0.0 (0.0–0.0)	0.0 (0.0–0.0)	1.000	0 (0.0)	25 (96.2)	1 (3.8)
Gait instability and ataxia	0.0 (0.0–0.0)	0.0 (0.0–0.0)	0.346	0 (0.0)	24 (92.3)	2 (7.7)
Brainstem dysfunction	0.0 (0.0–0.0)	0.0 (0.0–0.0)	NA	0 (0.0)	26 (100.0)	0 (0.0)
Weakness	0.0 (0.0–0.0)	0.0 (0.0–0.0)	1.000	0 (0.0)	25 (96.2)	1 (3.8)
Anti-GABA$_B$R encephalitis (n = 17)						
mRS	2.0 (2.0–3.0)	2.0 (1.0–3.0)	0.203	1 (5.9)	12 (70.6)	4 (23.5)
CASE (total score)	4.0 (3.0–5.0)	3.0 (3.0–4.0)	0.548	3 (17.6)	8 (47.1)	6 (35.3)
Seizure	1.0 (1.0–1.0)	1.0 (1.0–1.0)	0.346	2 (11.8)	15 (88.2)	0 (0.0)

Table 2 continued

CASE and mRS	CASE score at admission, median (IQR)	CASE score at discharge, median (IQR)	p^a	Change in severity by deviation of scores, n (%)		
				Deterioration	Stabilization	Remission
Memory dysfunction	2.0 (1.0–2.0)	1.0 (1.0–2.0)	0.490	1 (5.9)	12 (70.6)	4 (23.5)
Psychiatric symptoms	0.0 (0.0–2.0)	0.0 (0.0–2.0)	1.000	3 (17.6)	10 (58.8)	4 (23.5)
Consciousness	0.0 (0.0–0.0)	0.0 (0.0–0.0)	1.000	1 (5.9)	15 (88.2)	1 (5.9)
Language problem	0.0 (0.0–1.0)	0.0 (0.0–1.0)	0.713	1 (5.9)	13 (76.5)	3 (17.6)
Dyskinesia/ dystonia	0.0 (0.0–0.0)	0.0 (0.0–0.0)	NA	0 (0.0)	17 (100.0)	0 (0.0)
Gait instability and ataxia	0.0 (0.0–0.0)	0.0 (0.0–0.0)	1.000	1 (5.9)	15 (88.2)	1 (5.9)
Brainstem dysfunction	0.0 (0.0–0.0)	0.0 (0.0–0.0)	0.371	2 (11.8)	15 (88.2)	0 (0.0)
Weakness	0.0 (0.0–0.0)	0.0 (0.0–0.0)	1.000	1 (5.9)	15 (88.2)	1 (5.9)

[a] Wilcoxon matched-pairs signed-rank test

CASE Clinical Assessment Scale for Autoimmune Encephalitis, *mRS* modified Rankin scale, *NA* not available
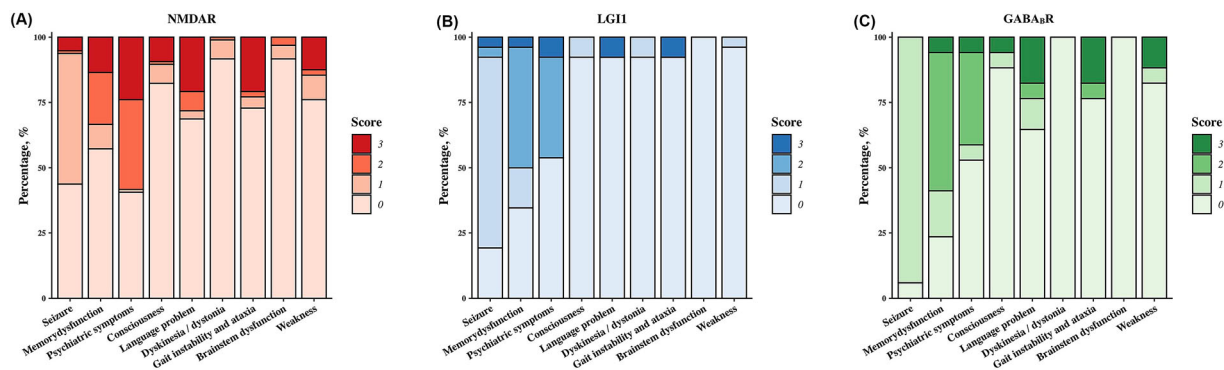
Fig. 5 The proportion of each score in items of CASE at admission among NMDAR encephalitis (**a**), LGI1 encephalitis (**b**), and GABA$_B$R encephalitis (**c**)

observed change in scores at admission and discharge. Indeed, the higher remission rate might be related to the fact that patients were typically discharged after some clinical improvement. Considering the variation in the scores of each item during hospitalization of the patient, adjusting their weightings might be helpful to improve the sensitivity of CASE on

tracking disease severity in the acute phase. Admittedly, there were some difficulties in the application of CASE. As the performance of each item varies with their frequencies, a low frequency of symptoms, such as dyskinesia and brainstem dysfunction, might affect the overall performance in terms of diagnosing definite neuronal surface antibody-associated AE. The analyses of our patient cohort revealed that the items of seizure, memory dysfunction, and psychiatric symptoms may be of potential efficacy to rate the severity of the disease, while those of dyskinesia/dystonia and brainstem dysfunction may be of limited value. However, it should also be noted that as a diagnostic tool, CASE was more complex and time-consuming than the mRS. Notably, mRS was merely used as a coarse and simplistic assessment in the case of no "gold standard" for assessing AE severity. The construct validity of CASE therefore remains unestablished. Relevant additional clinically meaningful measures, such as medication burden, ICU admission rate, formal neuropsychological assessments, formal psychiatric assessments, and seizure burden (not exhaustive), may provide further support or external validation for some items of CASE.

There are several limitations to our study. First, there might be an inevitable recall bias due to the retrospective nature of the study. For example, the symptom of dyskinesia might be related to confounding from the paroxysmal and confusing features and thus easily be ignored without records or marked as seizure [26, 27]. Secondly, our study included a limited number of subtypes of AE that are known to be associated with neuronal surface antibodies, with the majority of these presenting with a mild severity. The performance of the scale should be validated by including more categories, including AE associated with antibodies against intracellular antigens (onconeural proteins). Moreover, this study is limited to adult Chinese patients. The utility of CASE needs worldwide validation with studies including patient poulations of different ethnicity and all ages. Further, our study did not take account of adjustments for multiple comparisons. More prospective studies with a larger sample size from different categories of AE and diverse populations are needed to illustrate the feasibility of CASE for better clinical practice.

CONCLUSION

CASE is an effective clinical scale that can be utilized to evaluate patients with neuronal surface antibody-associated AE with high intra- and interobserver reliability. CASE had advantages over mRS in terms of its assessment of non-motor symptoms and greater sensitivity to changes in severity. However, its prognostic utility remains unknown, which merits further validation studies from international centers involving different patient populations.

ACKNOWLEDGEMENTS

committee of each participating center (Second Affiliated Hospital School of Medicine Zhejiang University [approval number: 2019-082], Zhejiang Hospital [approval number: 2021-9K], and Zhejiang Chinese Medicine and Western Medicine Integrated Hospital [approval number: 2021-021]). The study was conducted in accordance with the Helsinki Declaration of 1964 and its later amendments. The requirement to obtain patient consent was waived for this retrospective study

Data Availability. Anonymized data not published within this article will be made available upon reasonable request from any qualified investigator within 5 years after publication.

REFERENCES

1. Esposito S, Principi N, Calabresi P, Rigante D. An evolving redefinition of autoimmune encephalitis. Autoimmun Rev. 2019;18(2):155–63.

2. Lu Y, Wang Y, Wu X, Wang J, Zhu G, Hong Z. Clinical features and long-term outcomes of seizures associated with autoimmune encephalitis: a follow-up study in East China. J Clin Neurosci. 2019;68:73–9.

3. Gu Y, Zhong M, He L, et al. Epidemiology of antibody-positive autoimmune encephalitis in Southwest China: a multicenter study. Front Immunol. 2019;10:2611.

4. Dubey D, Pittock SJ, Kelly CR, et al. Autoimmune encephalitis epidemiology and a comparison to infectious encephalitis. Ann Neurol. 2018;83(1):166–77.

5. Qiu X, Zhang H, Li D, et al. Analysis of clinical characteristics and poor prognostic predictors in patients with an initial diagnosis of autoimmune encephalitis. Front Immunol. 2019;10:1286.

6. Broadley J, Seneviratne U, Beech P, et al. Prognosticating autoimmune encephalitis: a systematic review. J Autoimmun. 2019;96:24–34.

7. Balu R, McCracken L, Lancaster E, Graus F, Dalmau J, Titulaer MJ. A score that predicts 1-year functional status in patients with anti-NMDA receptor encephalitis. Neurology. 2019;92(3):e244–52.

8. van Swieten JC, Koudstaal PJ, Visser MC, Schouten HJ, van Gijn J. Interobserver agreement for the assessment of handicap in stroke patients. Stroke. 1988;19(5):604–7.

9. Lim JA, Lee ST, Moon J, et al. Development of the clinical assessment scale in autoimmune encephalitis. Ann Neurol. 2019;85(3):352–8.

10. Spagni G, Iorio R. Grading the severity of autoimmune encephalitis: advances and pitfalls. Ann Neurol. 2019;86(1):150.

11. Lim JA, Lee ST, Chu K, Lee SK. Reply to "Grading the severity of autoimmune encephalitis: advances and pitfalls." Ann Neurol. 2019;86(1):150–1.

12. Cai MT, Zheng Y, Zhang YX. Grading the severity of autoimmune encephalitis: when to evaluate? Ann Neurol. 2020;87(6):989–90.

13. Lee ST, Lim JA, Chu K, Lee SK. Reply to "Grading the severity of autoimmune encephalitis: when to evaluate?" Ann Neurol. 2020;87(6):990.

14. Graus F, Titulaer MJ, Balu R, et al. A clinical approach to diagnosis of autoimmune encephalitis. Lancet Neurol. 2016;15(4):391–404.

15. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33(1):159–74.

16. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med. 2016;15(2):155–63.

17. Bland JM, Altman DG. Cronbach's alpha. BMJ. 1997;314(7080):572.

18. Arah OA, Hoekstra JB, Bos AP, Lombarts KM. New tools for systematic evaluation of teaching qualities of medical faculty: results of an ongoing multi-center survey. PLoS ONE. 2011;6(10):e25983.

19. Schober P, Boer C, Schwarte LA. Correlation coefficients: appropriate use and interpretation. Anesth Analg. 2018;126(5):1763–8.

20. Vetter TR, Schober P. Agreement analysis: what he said, she said versus you said. Anesth Analg. 2018;126(6):2123–8.

21. Hu Y, Zhang Z. Skilled doctors in tertiary hospitals are already overworked in China. Lancet Glob Health. 2015;3(12):e737.

22. Yu W, Li M, Ye F, Xue C, Zhang L. Patient preference and choice of healthcare providers in Shanghai, China: a cross-sectional study. BMJ Open. 2017;7(10):e016418.

23. Tu J, Wang C, Wu S. Skilled doctors in tertiary hospitals are already overworked in China—Authors' reply. Lancet Glob Health. 2015;3(12): e738.

24. Lu C, Zhang Z, Lan X. Impact of China's referral reform on the equity and spatial accessibility of healthcare resources: a case study of Beijing. Soc Sci Med. 2019;235:112386.

25. Papanicolas I, Woskie LR, Jha AK. Health care spending in the United States and other high-income countries. JAMA. 2018;319(10):1024–39.

26. Fasano A, Di Bonaventura C, Bove F, et al. Movement disorders phenomenology in focal motor seizures. Parkinsonism Relat Disord. 2019;61:161–5.

27. Freitas ME, Ruiz-Lopez M, Dalmau J, et al. Seizures and movement disorders: phenomenology, diagnostic challenges and therapeutic approaches. J Neurol Neurosurg Psychiatry. 2019;90(8):920–8.