

Botulinum toxin treatment

Optimising the methods of evaluation of the effectiveness of botulinum toxin treatment of post-stroke muscle spasticity

A M O Bakheit

The correct research design is crucial in establishing whether treatments for complex neurological disorders are effective

The effectiveness of botulinum toxin type A (BtxA) in reducing post-stroke muscle hypertonia and its complications has been demonstrated in adult patients in at least eight large randomised controlled trials (RCTs).^{1–8} The clinical benefits were reported when diverse treatment protocols and different dose schedules and injection techniques were used for the management of the same clinical indication. Nonetheless, some doubt as to the value of this treatment remain.⁹ What is the explanation for this apparent paradox? The contradiction may be explained by the inadequate design of these studies and the inherent difficulties with the interpretation of the results of RCTs of a drug treatment that is carried out in the context of rehabilitation for a person with chronic neurological disability. The WHO classification of functioning, disability, and health¹⁰ provides a useful framework for understanding the complex relation between the statistical significance and the clinical significance of a given treatment and helps to explain the complexity of rehabilitation research. A statistically significant change at the level of impairment (as a result of the intervention) may not alter the subject's functional abilities or potential for social participation. The outcome of treatment with BtxA is clinically significant only if it results in the full or partial attainment of functional goals that are meaningful to the recipient of this treatment. An example of such outcomes is the achievement of safe walking and prevention of falls in a patient with dynamic foot equinus after BtxA injections into the calf muscles. Successful treatment reduces muscle tone in the ankle plantar flexors and allows correct placement of the foot in stance (which ensures stability) and adequate foot clearance of the ground in the swing phase of the gait cycle (which prevents tripping over).

The major RCTs on the effectiveness of BtxA in spasticity that have been published to date suffer from methodological problems, especially in relation to the choice of the outcome measures and the data analysis and interpretation. The Ashworth scale was often chosen as the primary outcome measure, although the value of this scale as a measure of spasticity is doubtful.¹¹ Furthermore, the Ashworth scale measures change at the level of impairment. It does not inform the observer about the effect of treatment on functional abilities. Another drawback of the Ashworth scale is that the measurement technique is not standardised. The scale also does not reliably distinguish spasticity from a fixed contracture (which is not amenable to medical treatment). It is clear from the above thesis that the Ashworth scale is not a suitable outcome measure of treatment effectiveness, although it may be used as a screening tool for inclusion for RCTs. Similarly, measurements of the joint range of motion on passive muscle stretch have little or no value in the assessment of the outcomes of treatment with BtxA.

Another frequently encountered weakness of the hitherto published RCTs is the inappropriate statistical analysis of the data. For example, the scores derived from the Ashworth scale (which is an ordinal scale) were often analysed with parametric statistics, as if they were the result of an interval level outcome measure.

The treatment of muscle spasticity is usually undertaken as part of a holistic rehabilitation programme that seeks to reduce the subject's disability and to promote their social participation (that is, reduce handicap). Under these circumstances the overall rehabilitation care is likely to be more important in producing functional change than a single specific intervention, such as BtxA injections. Consequently, inadequately

designed RCTs may be misleading when used to examine whether treatment with BtxA improves motor function. This is because antispasticity treatment is directed at the level of impairment and reversal or improvement in the impairment—that is, the muscle hypertonia—does not necessarily translate into better functional abilities. Furthermore, the measurement of the treatment outcomes is confounded by the effects of other therapeutic interventions, including physiotherapy, the use of orthotic devices, and so on. Other important confounding factors are the variability in the treatment goals between subjects and the influences of the patient's personal attributes, personal rehabilitation objectives, physical environment, and social context. The standardised outcome measures used in RCTs do not address the variability in the personal experience or treatment objectives of individual patients which are important determinants of the clinical outcome.

It is considered a major advantage of RCTs that the data generated are easily amenable to statistical analysis and statistical significance testing. However, often there is no direct relation between statistical and clinical significance and this is particularly relevant for rehabilitation research. This is because statistical significance does not measure the size or importance of the treatment effect that is deemed “significant.” It only excludes the probability of it occurring by chance. Therefore, a statistically significant result of a given intervention—for example, the reduction of muscle spasticity with BtxA—may not correspond to or reflect a meaningful functional gain to the patient. A statistically significant reduction in muscle tone may even be detrimental in some patients although it may be useful in others with an identical diagnosis or impairment.

RCTs are more suited to the study of the effectiveness of treatment at the level of impairment. However, intervention at this level is seldom useful in the context of complex neurological disability except occasionally when the successful treatment of the impairment has a direct linear relation with the desired functional outcome. In addition, the results of RCTs demonstrate the general trends of the study variable in a given study population. The averaged group data of RCTs obscure the poor response to treatment in individual patients. As poor response to treatment may reflect inadequate control for the confounding factors described above, efforts to minimise this effect are necessary before randomisation for group trials. The

design of future RCTs on the effectiveness of BtxA in spasticity can be improved only when goal oriented outcome measures are incorporated in the study design and when the design of RCTs is informed by data from qualitative and single case studies.

Single case study design has several advantages.¹² The subjects act as their own controls and all confounding variables, including personal, social, and environmental factors, remain constant. In addition, this research method permits the use of individualised outcome measures (for example, goal attainment scales) and allows flexibility in the treatment protocol (such as variations in the dose of BtxA and the choice of muscles for injection). This approach also deals better with the broad range of sensory and cognitive symptoms in the stroke population that have an impact on motor function, such as unilateral neglect, anosognosia for hemiplegia, and so on. Although single case studies eliminate the confounding factors that are often seen in a rehabilitation setting, they do not take into account the patients' variability. This difficulty may be overcome by conducting a series of carefully designed single case studies to identify subgroups of patients who are likely to benefit from the specific treatment. The group of patients identified in this way can then be studied in an RCT.

In conclusion, a research method in rehabilitation is clinically relevant only if it addresses the impact of the intervention on functional abilities and social

participation. Although RCTs are considered the gold standard in biomedical research, they may not be suitable as the most valid research method in cases of complex neurological disability unless the treatment effect is assessed with goal oriented outcome measures. Partial or full attainment of these goals would be the yardstick by which treatment success is confirmed. Single case studies and qualitative observational research are more likely to reflect the complexities of clinical practice and to help answer the research question more completely. One could argue that a series of single case studies is a useful prerequisite for stratification for RCTs in the context of rehabilitation research. The importance of choosing the correct research design cannot be overemphasised and this sometimes means combining different research methods.¹³

J Neurol Neurosurg Psychiatry
2004;75:665-666.
doi: 10.1136/jnnp.2003.026096

Correspondence to: Professor A M O Bakheit, Neurological Rehabilitation, Peninsula Medical School and Plymouth Primary Care Trust, Mount Gould Hospital, Plymouth PL4 7QD, UK; magid.bakheit@pcs-tr.swest.nhs.uk

REFERENCES

- 1 **Burbaud P**, Wiart L, Dubos JL, *et al*. A randomised double blind placebo controlled trial of botulinum toxin in the treatment of spastic foot in hemiparetic patients. *J Neurol Neurosurg Psychiatry* 1996;61:265-9.
- 2 **Simpson DM**, Alexander DN, O'Brien CF, *et al*. Botulinum toxin type A in the treatment of upper extremity spasticity: a randomized, double-blind,

placebo-controlled trial. *Neurology* 1996;46:1306-10.

- 3 **Bakheit AMO**, Thilmann AF, Ward AB, *et al*. A randomized, double-blind, placebo-controlled, dose-ranging study to compare the efficacy and safety of three doses of botulinum toxin type A (Dysport) with placebo in upper limb muscle spasticity after stroke. *Stroke* 2000;31:2402-6.
- 4 **Bakhta BB**, Cozens JA, Bamford JM, *et al*. Impact of botulinum toxin on disability and carer burden due to arm spasticity after stroke: a randomised double blind placebo controlled trial. *J Neurol Neurosurg Psychiatry* 2000;69:217-21.
- 5 **Richardson D**, Sheehan G, Werring D, *et al*. Evaluating the role of botulinum toxin in the management of focal hypertonias in adults. *J Neurol Neurosurg Psychiatry* 2000;69:499-506.
- 6 **Bakheit AMO**, Pittock S, Moore AP, *et al*. A randomised, double-blind, placebo-controlled study of the efficacy and safety of botulinum toxin A (Dysport) in upper limb spasticity in patients with stroke. *J Neurology* 2001;8:559-65.
- 7 **Brashear A**, Gordon MF, Elovic E, *et al*. Intramuscular injection of botulinum toxin for the treatment of wrist and finger spasticity after a stroke. *N Engl J Med* 2002;347:395-400.
- 8 **Pittock SJ**, Moore AP, Hardiman O, *et al*. A double blind randomised placebo controlled evaluation of three doses of botulinum toxin type A (Dysport) in the treatment of spastic equinovarus deformity after stroke. *Cerebrovasc Dis* 2003;15:289-300.
- 9 **Moore AP**. Botulinum toxin A (BoNT-A) for spasticity in adults. What is the evidence? *Eur J Neurol* 2002;9(suppl 1):42-7.
- 10 **WHO**. *International classification of functioning, disability and health*. Geneva: World Health Organisation, 2001; <http://www3.who.int/icf/icfemplate.cfm>.
- 11 **Bakheit AMO**, Maynard VA, Curnow J, *et al*. The relationship between Ashworth scale scores and the excitability of the alpha motor neurones in patients with post stroke muscle spasticity. *J Neurol Neurosurg Psychiatry* 2003;74:646-8.
- 12 **Ottensbacher KJ**. Clinically relevant designs for rehabilitation research: the idiographic model. *Am J Phys Med Rehabil* 1990;69:286-92.
- 13 **Sackett DL**, Wennberg JE. Choosing the best research design for each question. *BMJ* 1997;315:1636.

Rapid test for dementia

Detecting dementia in just 12 minutes: the seven minute screen

V W Henderson

There could be a useful niche for this neuropsychological test

The 11 item mini-mental state examination (MMSE) was devised as a brief yet "thorough" and "practical" assessment of the cognitive mental status.¹ It is often used to screen for dementia caused by Alzheimer's disease and to grade dementia severity. From humble origins three decades ago, the MMSE has become the benchmark against which other cognitive instruments are routinely compared. Alternatives have been proposed to correct perceived limitations of the original;

among these psychometric pretenders is the seven minute neurocognitive screening battery, or seven minute screen (7MS).²

The 7MS is a rational concatenation of four tests, each sensitive to the presence of dementia (temporal orientation, cued recall, category fluency, and clock drawing).² Envisioning a role for the 7MS in clinical practice, Meulen *et al* compared Dutch translations of the 7MS and MMSE in large outpatient convenience samples in Amsterdam and

Maastricht (this issue, pp 700-705).³ The participants, who were referred to specialty clinics for memory complaints, may be different from patients in primary care settings. Exclusion criteria included psychoactive drug use, recent stroke, active neurological or medical disease, and current alcohol abuse. Primary comparisons were with a combined "normal" group: referrals with memory complaints where cognition proved to be intact, plus patients' relatives and friends without memory symptoms. Normal individuals were similar to demented patients in terms of age and education but probably differed from the general population from which referral cases arose.

The 7MS took about 12 minutes to administer (eight minutes for normal subjects, 16 minutes for patients with Alzheimer's disease), with additional time required for scoring.³ Analyses were unadjusted for confounders of age, education, or sex, but receiver operator characteristic curves (a useful method of displaying sensitivity and

specificity information) favoured the 7MS over the MMSE in distinguishing Alzheimer's disease patients from normal individuals. Small numbers precluded comparisons between other specific diagnoses and normal subjects, but the 7MS was also superior when applied to an amalgamated non-Alzheimer's disease dementia group.³ Non-demented patients with depression or other psychiatric diagnoses were often misclassified by both tests. Performance characteristics of the 7MS would certainly be reduced in the complex, messy world inhabited by most clinicians. Here, target populations often include people using psychoactive drugs and patients with depression, neurological or medical disease, or mild cognitive impairment without dementia.

How might the practitioner use the 7MS? This instrument is not intended to assess cognition comprehensively. If the goal is to preselect older patients who might benefit from detailed clinical assessment for dementia, the 7MS clearly outperforms the benchmark MMSE. Although not addressed by

Meulen *et al*, it is likely that the 7MS would also provide a more precise measure of dementia severity in following longitudinal progression. Better sensitivity, specificity, and predictive value should be weighed against a longer administration time and reduced familiarity. It is not known how the 7MS might compare with other validated instruments—for example, the cognitive abilities screening instrument⁴ or neuropsychological tasks from CERAD (the consortium to establish a registry for Alzheimer's disease).⁵

Practitioners should carefully consider the purpose of cognitive screening. For example, clinical assessment for dementia is more commonly triggered by patient or caregiver complaints of cognitive impairment than by screening results. In some settings, however, 12 minutes could be a rewarding investment, suggesting a useful niche for the optimistically named 7MS.

J Neurol Neurosurg Psychiatry
2004;**75**:666–667.
doi: 10.1136/jnnp.2003.032359

Correspondence to: V W Henderson, Central Arkansas Veterans Healthcare System, and Departments of Geriatrics, Neurology, Pharmacology, and Epidemiology, University of Arkansas for Medical Sciences, Little Rock, AR, USA and Department of Psychiatry, University of Melbourne, Parkville, VIC, Australia; vhenderson@uams.edu

REFERENCES

- 1 Folstein MF, Folstein SE, McHugh PR. "Mini-mental state." A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975;**12**:189–98.
- 2 Solomon PR, Hirschhoff A, Kelly B, *et al*. A 7 minute neurocognitive screening battery highly sensitive to Alzheimer's disease. *Arch Neurol* 1998;**55**:349–55.
- 3 Meulen EFJ, Schmand B, van Campen JP, *et al*. The seven-minute screen: a neurocognitive screening test highly sensitive to various types of dementia. *J Neurol Neurosurg Psychiatry* 2004;**75**:700–5.
- 4 Teng EL, Hasegawa K, Homma A, *et al*. The Cognitive Abilities Screening Instrument (CASI): a practical test for cross-cultural epidemiological studies of dementia. *Int Psychogeriatr* 1994;**6**:45–58.
- 5 Welsh K, Butters N, Hughes J, *et al*. Detection of abnormal memory decline in mild cases of Alzheimer's disease using CERAD neuropsychological measures. *Arch Neurol* 1991;**48**:278–81.

Psychogenic non-epileptic seizures

Assessment of patients with psychogenic non-epileptic seizures

L H Goldstein

Psychogenic non-epileptic seizures pose a management problem

Psychogenic non-epileptic seizures (PNES), while superficially resembling epileptic seizures, are not accompanied by the abnormal electrical discharges associated with epilepsy and cannot be explained by other medical conditions. Instead they are psychologically determined and, as patients with PNES may be misdiagnosed as having epilepsy, this disorder poses a considerable management problem. The development of a better understanding of the psychiatric characteristics of PNES patients and other risk factors associated with having PNES may not only assist, along with medical investigations, in the correct diagnosis of patients' attacks, but may also inform effective treatment.

Though personality disorder has been diagnosed in very variable proportions of PNES patients, and research in this

area is complicated by the use of different psychiatric classification systems and measures, the most common diagnoses associated with PNES have been borderline, histrionic, avoidant, and antisocial personality disorder.¹ Given that not all PNES patients satisfy formal diagnostic criteria for personality disorder, a different approach to understanding personality pathology is adopted in the paper by Reuber *et al* in this issue,² who used the dimensional assessment of personality pathology—basic questionnaire (DAPP-BQ) to identify personality profiles in PNES patients. While indicating that PNES patients as a group have greater pathology of personality than people with epilepsy or healthy controls, Reuber *et al* also found that the PNES patients' scores on the DAPP-BQ did not form a single profile of personality

traits. Instead they identified two large and one much smaller clusters of patients accounting for 84 of their sample of 85 PNES patients. The DAPP-BQ scores of the two larger clusters resembled the characteristics of borderline personality disorder and compulsive personality disorder, respectively. The third cluster of PNES patients had traits resembling the characteristics associated with avoidant personality; this cluster contained only four individuals.

While a diagnosis of personality disorder per se has been considered a poor prognostic factor in people with PNES,¹ Reuber *et al*'s approach to personality measurement in individuals with PNES offers the opportunity of moving away from the requirement to make a formal diagnosis of personality disorder before considering personality as one of the factors predictive of outcome. Reuber *et al* report, on the basis of their retrospective data, that those PNES patients with personality traits resembling borderline personality disorder had worse outcome on follow up in terms of seizure freedom than did the cluster with traits resembling compulsive personality disorder, despite the fact that more of the former than of the latter group had received (unspecified) psychiatric treatment. Reuber *et al* also address the possible use of different psychotherapeutic techniques for patients

with different personality profiles. In addition to dialectical behaviour therapy,³ it may be significant in this context that cognitive therapy has also been developed for use with patients with borderline personality disorder.⁴ However, the use of these personality related psychotherapeutic approaches with people who have not received a formal diagnosis of personality disorder but who have PNES and other psychiatric comorbidities¹ requires careful evaluation alongside other cognitive behavioural methods⁵ before clear

indications of their specific value and effectiveness with PNES patients can be obtained.

J Neural Neurosurg Psychiatry
2004;**75**:667–668.
doi: 10.1136/jnnp.2003.033753

Correspondence to: L H Goldstein, Department of Psychology, PO77, Institute of Psychiatry, De Crespigny Park, London SE5 8AF, UK; l.goldstein@iop.kcl.ac.uk

REFERENCES

- 1 **Bowman ES.** Psychopathology and outcome in pseudoseizures. In: Ettinger AB, Kanner AM,

eds. *Psychiatric issues in epilepsy. A practical guide to diagnosis and treatment.* Philadelphia: Lippincott Williams and Wilkins, 2001:355–77.

- 2 **Reuber M,** Pukrop R, Bauer J, *et al.* Multidimensional assessment of personality in patients with psychogenic non-epileptic seizures. *J Neural Neurosurg Psychiatry* 2004;**75**:743–8.
- 3 **Linehan M.** *Cognitive-behavioural treatment of borderline personality disorder.* New York: Guilford, 1993.
- 4 **Davidson KM,** Tyrer P. Cognitive therapy for antisocial and borderline personality disorders: single case study series. *Br J Clin Psychol* 1996;**35**:413–29.
- 5 **Chalder T.** Non-epileptic attacks; a cognitive behavioural approach in a single case with a four year follow-up. *Clin Psychol Psychotherapy* 1996;**3**:291–7.