BMJ Open
accessible medical research

Upper limb international spasticity study: rationale and protocol for a large, international, multicentre prospective cohort study investigating management and goal attainment following treatment with botulinum toxin A in real-life clinical practice

Lynne Turner-Stokes,[1] Klemens Fheodoroff,[2] Jorge Jacinto,[3] Pascal Maisonobe,[4] Benjamin Zakine[4]

For numbered affiliations see end of article.

Correspondence to
Professor Lynne Turner-Stokes; lynne.turner-stokes@dial. pipex.com

ABSTRACT

Objectives: This article provides an overview of the Upper Limb International Spasticity (ULIS) programme, which aims to develop a common core dataset for evaluation of real-life practice and outcomes in the treatment of upper-limb spasticity with botulinum toxin A (BoNT-A). Here we present the study protocol for ULIS-II, a large, international cohort study, to describe the rationale and steps to ensure the validity of goal attainment scaling (GAS) as the primary outcome measure.

Methods and analysis design: An international, multicentre, observational, prospective, before-and-after study, conducted at 84 centres in 22 countries across three continents.

Participants: 468 adults presenting with poststroke upper limb spasticity in whom a decision had already been made to inject BoNT-A (5–12 consecutive participants recruited per centre).

Interventions: Physicians were free to choose targeted muscles, BoNT-A preparation, injected doses/technique and timing of follow-up in accordance with their usual practice and the goals for treatment.

Primary outcome measure: GAS.

Secondary outcomes: Measurements of spasticity, standardised outcome measures and global benefits. Steps to ensure validity included: (1) targeted training of all investigators in the use of GAS; (2) within-study validation of goal statements and (3) establishment of an electronic case report form with an in-built tracking facility for separation of baseline/follow-up data.

Analysis: Efficacy population: all participants who had (1) BoNT-A injection and (2) subsequent assessment of GAS. Primary efficacy variable: percentage (95% CI) achievement of the primary goal from GAS following one BoNT-A injection cycle.

ARTICLE SUMMARY

Article focus
- To provide an overview of the Upper Limb International Spasticity (ULIS) programme and the rationale and protocol of the ULIS-II study.
- To outline the steps taken in ULIS-II to ensure the quality of goal statements and to support the validity of goal attainment scaling (GAS) as the primary outcome measure for the trial.

Key message
- Evaluation of goals statements part-way through this study has assisted participating centres to improve the quality and function-related focus of goal statements.

Strengths and limitations of this study
- This methodology helps to support the validity of GAS as the primary outcome measure for the efficacy analysis.
- This large international cohort study represents a diverse sample of practice across three continents.
- However, the limited number of participants per centre (5–12) could lead to some selection bias.

Ethics and dissemination: This non-interventional study is conducted in compliance with guidelines for good pharmacoepidemiology practices. Appropriate ethical approvals were obtained according to local regulations. ULIS-II will provide important information regarding treatment and outcomes from BoNT-A in real-life upper limb spasticity management. The results will be published separately.

Registration: ClinicalTrials.gov identifier: NCT01020500.

INTRODUCTION

Spasticity is a common sequela of stroke, with an incidence ranging from 17% to 38%.[1–5] It is more prevalent in younger patients,[2] and most commonly affects the upper limb.[6]

Upper limb spasticity is often painful. It interferes with upper limb movement, and limits use of the limb for active functional tasks. It can cause involuntary movements (associated reactions) that impact on mobility (gait, balance, walking speed, etc). In severe cases, it can also impede 'passive function', such as washing, dressing and caring for the affected limb, thereby increasing the burden on caregivers.[7 8]

There is now a well-established body of evidence demonstrating that botulinum toxin A (BoNT-A) is a well-tolerated and effective focal intervention for the reduction of spasticity, and it is widely recommended for use in standard clinical practice.[7–9] Controlled clinical trials (CCTs)[10–18] have confirmed the benefits of BoNT-A at the level of impairment, but functional change has been harder to demonstrate, particularly where impact on active function is limited by underlying motor dysfunction. Nevertheless, on an individual level, clinical experience suggests that some patients make substantial functional gains.

While CCTs may be helpful for establishing the overall clinical efficacy of an intervention, they do not answer important clinical questions such as which patients are most likely to benefit and in what way; or which treatment approaches work best in real-life clinical care. For these, we need large, multicentre, longitudinal cohort studies conducted in the course of routine clinical practice.[19] If the findings are to be generalisable across different health cultures, these studies need to have wide geographical representation across the international health community.

Measuring the effectiveness of BoNT-A treatment is challenging in the context of upper limb spasticity because of wide diversity in patient presentation, potential for rehabilitation and goals for treatment. Timing of assessment is also important owing to the need for time to practice and develop new skills in the hemiparetic arm after the spasticity has been relieved.[20 21] Current guidelines for the use of BoNT-A in the management of spasticity advocate the application of focused outcome evaluations, targeted on the attainment of priority goals that are both relevant to the treatment intentions and important to the individual.[7 8] Depending on the nature of the goals set, the individual clinical presentation and any underlying trajectory towards recovery or deterioration, the timescale for expected goal achievement will vary from person to person. Cohort study design in this context must therefore be flexible enough to account for all this variation.

Goal attainment scaling (GAS) is a method of assimilating achievement of a number of individually set goals into a single 'goal attainment score', in order to capture outcomes across a diverse range of goal areas. Originally described by Kiresuk and Sherman in the 1960s,[22] it is increasingly recognised as a sensitive method for recording patient-centred outcomes in this context.[23–26] In addition to providing a semiquantitative (ordinal) assessment of goal attainment, GAS offers potentially useful qualitative information regarding the patient's priority goals for treatment. Moreover, the process of goal setting and rating itself offers an opportunity for dialogue and negotiation between the patient and the treating team,[27] which may help to establish mutual agreement of expectations for the outcome. However, this requires knowledge and experience, and it is important to recognise that GAS is not a measure of outcome per se, but a measure of the achievement of intention. It therefore does not replace standardised measures, but is a useful adjunct to use alongside them.[28]

The use of GAS as a primary outcome measure for research is still somewhat controversial. Concerns have been raised in some quarters about the validity of GAS, in particular, the non-linearity of the scaling and lack of unidimensionality,[29] and some authors have proposed the development of standardised goals or 'item banks'.[29 30] The WHO International Classification of Functioning, Disability and Health (ICF)[31] provides a useful common language for categorising goals into different domains of personal experience. A previous secondary analysis of a multicentre CCT from Australia[24] mapped a total of 165 goals onto ICF domains, to identify the key goal areas impacting on quality of life, in order to inform the future development of standardised goal sets. The authors recommended that further research with a priori categorisation of goals in large, prospective, cohort studies is required to describe the full value of BoNT-A in the management of upper limb spasticity.

From the studies conducted to date, we know that there is considerable individual variation in response. We also know that there is variation in treatment with respect to selection of muscle, injection technique and concomitant therapy interventions (eg, physiotherapy and/or occupational therapy), and that these appear to have more to do with clinician bias and local availability of services than with patient presentation.[32] It is now time to extend the field of investigation in this area to explore how BoNT-A is used in routine clinical practice, in order to gain a better understanding of how to select those most likely to respond and what works best for which types of presentation. Horn and Gassaway[33] argue that this type of 'practice-based evidence' is as important for building the evidence base for clinical practice as the information that derives from CCTs.

To do this, however, we will need to build a consistent body of data that captures clinically important changes at an individual level and is of sufficient size and generalisability to interrogate for future answers to these critical questions. The challenge lies in how to engage clinicians across the globe, to engage in a single common approach to data collection.

The Upper Limb International Spasticity (ULIS) study programme represents the first step towards this aim.

This manuscript provides an overview of the programme and presents the rationale and protocol for the ULIS-II study. The study results for ULIS-II will be presented separately.

OVERVIEW OF THE ULIS PROGRAMME AND RATIONALE

The ULIS programme consists of a series of international observational studies to describe current clinical practice in the application of BoNT-A in the management of upper limb spasticity. The ultimate aim is to work towards the development of a common core dataset for prospective systematic recording of longitudinal outcomes that could inform the development of a large international database of sufficient size and breadth to support future interrogation and subset analysis.

A founding principle of establishing common datasets is to embed them as closely as possible in real-life current clinical practice.

▶ The first stage of the programme (ULIS-I)[34] was an international, cross-sectional survey, which was designed to document clinical practice across four continents with respect to treatment and outcome evaluation.

▶ The second stage (ULIS-II) (described here) is a large, before-and-after, prospective, observational cohort study to record goal attainment as a primary outcome following one cycle of BoNT-A. This has also served to develop expertise in GAS and refine an electronic case report form (e-CRF). ULIS-II has recently completed recruitment and is now in the process of analysis.

▶ Following further refinement of the dataset and tools, a third stage is planned to expand the cohort and to capture the broader benefits of treatment in a fully generalisable sample recorded longitudinally over several cycles of treatment. This is currently in the planning and preinvestigation phase.

How ULIS-I informed the rationale for ULIS-II

ULIS-I was an international, cross-sectional, non-interventional survey to document (from a review of current practice and reported intentions) the clinical profiles, treatment goals and reported outcome evaluation in consecutive adults attending treatment with BoNT-A for upper limb spasticity.[34] Over a 6-month period, a total of 974 consecutive patients were recruited from 122 investigational centres in 31 countries spanning the European Union, Pacific Asia, Eastern Europe, the Middle East and South America.

The findings demonstrated wide diversity in clinical practice with respect to both the method of intervention and the use of assessment and outcome measures.[34] Most frequently recorded were impairment measures, including range of movement (90%) and spasticity—mainly using the Modified Ashworth Scale (MAS) (83%). Although 36% said that they routinely recorded at least one measure of active function, there was wide variation in the instruments used and insufficient commonality to allow pooling of data.

Goal-setting, on the other hand, was very common (78%). However, although the large majority of clinicians reported that they set goals, the way in which they used these to monitor the effects of treatment varied widely. Formal application of GAS was used by only a handful of participating centres (5%).

The findings suggested that goal attainment offered the most widely applicable common outcome measure, and for this reason, it was selected as the principal outcome measure for the next stage of the programme (ULIS-II). However, it was first necessary to establish a consistent approach to the recording of goals and goal attainment. It was also important that the method was simple and practical to apply across the wide international spectrum of clinical practice in upper limb spasticity management.

ULIS-II STUDY PROTOCOL: METHODS AND ANALYSIS
Study objectives

The primary objective of the ULIS-II study was to assess the responder rate (as defined by the achievement of the primary goal from GAS) following one BoNT-A injection cycle delivered in the context of routine clinical practice.

Secondary objectives were to:

▶ Describe the baseline characteristics, including (but not limited to) demographics, duration and pattern of spasticity and concomitant therapies/medication.

▶ Describe injection practices (muscle identification, dosage, dilution and injection points) and additional treatment strategies—including therapy intervention and different types of modality.

▶ Assess the achievement of secondary goals and the overall attainment of treatment goals using the GAS T-score.

▶ Document the use of standardised outcome measures and their results.

▶ Assess the global benefits as perceived by the investigator and the patient (or guardian).

Exploratory objectives, addressed through the analysis plan, were to:

▶ Describe the common goal areas for treatment and to identify those in which goals were most often achieved.

▶ Identify any prognostic factors for response.

Study design and setting

ULIS-II was an 18-month, postmarketing, international, multicentre, observational, prospective, longitudinal study conducted in 84 centres in 22 countries (European Union, Pacific Asia, Eastern Europe and South America). Figure 1 shows the geographical distribution of the participating centres.

Figure 1 World map showing geographical distribution of participating centres in Upper Limb International Spasticity II (ULIS-II).

Study participants

The main inclusion criteria were:

▶ Adults ≥18 years with poststroke upper limb spasticity in whom a decision had already been made to inject BoNT-A.

▶ No previous treatment with BoNT-A or BoNT-B within the last 12 weeks.

▶ Agreement on an achievable goal set and ability to comply with the prescribed treatment.

Exclusions were any contraindications to BoNT-A or failure to consent to participate.

Recruitment

To limit the potential bias that might be introduced by over-recruiting sites, the number of patients was limited to 5–12 per treatment centre. Participating centres represented a range of experience to mirror clinical practice. However, in order to capture the approaches to treatment that are borne of clinical experience, recruitment at less-experienced centres (n=59) was restricted to five patients only, while more experienced centres (n=25) could recruit up to 12 patients. This approach also offers the potential for future subanalysis of the differences between experienced and less-experienced injectors.

Centres were asked to include consecutive patients attending the clinics over a specified period. If consecutive inclusions were not feasible (eg, owing to administrative constraints in a busy clinic setting), investigators were authorised to space the inclusions (eg, one per every two to three patients), until their recruitment target was achieved.

A total of 468 patients were recruited, 226 (48.3%) in more-experienced and 242 (51.7%) in less-experienced centres, confirming more or less equal distribution. Figure 2 shows the disposition of patients and table 1 shows the breakdown by country, including completion rates. The 12-case attrition was primarily in Europe, but did not significantly affect geographic representation as the European countries were the best represented to start with. Table 2 shows the demographics of the efficacy population who completed the study per protocol.
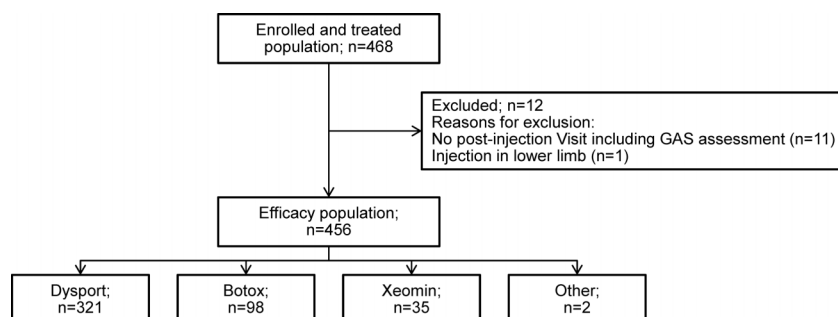
Study schedule

Baseline evaluation (time 1)

On inclusion into the study, the following assessments were recorded by the investigator on the e-CRF:

▶ Demography and history of the stroke including type, location and time since onset (table 2).

Figure 2 Flow chart for recruitment.

Table 1 Recruitment and attrition by country

Country	Number of centres	Total recruits	Attrition	Number of cases completed (efficacy population)	Percentage of total efficacy population*
Europe					
Austria	2	14	0	14	3
Belgium	4	25	1	24	5
Czech Republic	2	10	0	10	2
Denmark	1	5	0	5	1
Finland	3	13	1	12	3
France	14	48	1	47	10
Germany	6	45	2	43	9
Italy	6	33	2	31	7
Portugal	5	27	0	27	6
Russia	7	41	0	41	9
Spain	3	14	0	14	3
Sweden	2	14	0	14	3
The UK	5	45	4	41	9
Sub-total	**60**	**334**	**11 (3.3%)**	**323**	**70**
Pacific Asia					
South Korea	5	29	1	28	6
Singapore	2	10	0	10	2
Taiwan	2	10	0	10	2
Australia	6	44	0	44	10
China	1	5	0	5	1
Malaysia	2	6	0	6	1
Philippines	2	10	0	10	2
Thailand	2	10	0	10	2
Sub-total	**22**	**124**	**1 (0.8%)**	**123**	**27**
South America					
Mexico	2	10	0	10	2
	2	**10**	**0**	**10**	**2%**
Total	**n=22**	**n=84**	**n=468**	**n=12 (2.6%)**	**n=456**

*Owing to rounding, percentages may not total 100%.

Table 2 Demographics of the efficacy population

Parameter	Values	Range	N/missing or untestable
Age (years)			456/0
Mean (SD)	56.7 (13.5)	18–88	
Time since onset of stroke (months)			456/0
Mean (SD)	61.4 (69.1)	1–447	
Gender, n (%)			456/0
Male	266 (58.3%)		
Female	190 (41.7%)		
Aetiology, n (%)			456/0
Infarct	320 (70.2%)		
Haemorrhage	139 (30.5%)		
Both	3 (0.7%)		
Location of CVA, n (%)			456/0
Left hemisphere	215 (47.1%)		
Right hemisphere	235 (51.5%)		
Bilateral*	4 (0.9%)		
Posterior circulation	13 (2.9%)		

*CVAs affecting both hemispheres.
CVA, cerebrovascular accident.

- The pattern of impairment in the affected upper limb (motor function, sensation and contractures) and the presence of any generalised impairments that may affect outcome (including cognitive, emotional and behavioural function) were recorded using a modified version of the Neurological Impairment Scale.[35]
- Previous/concomitant treatments for upper limb spasticity, including therapies and medication.
- Clinical examination, including a list of measurements of spasticity and other standardised outcome measures, as routinely performed in that centre, for example:
    - Measurements of spasticity (MAS and Tardieu scale);
    - Visual analogue and verbal rating scales for symptoms such as pain, ease of care, etc;
    - Standardised scales for active and/or passive function (eg, the Leeds Adult Spasticity Impact Scale[11] or the Arm Activity scale[36 37]);
    - Measures of involuntary movement, for example, the Associated Reaction Rating Scale.[38]
- Goal setting and GAS were applied using the 'GAS-light' method[28] as detailed below with emphasis on setting SMART (specific, measurable, achievable, realistic and timed) goals agreed between the investigator, the patient and the treating team. One primary and up to three secondary goals were set and assigned to one of seven goal categories.

Injection of BoNT-A

To reflect real-life practice in this non-interventional; observational study, physicians were free to choose targeted muscles, BoNT-A preparation, injected doses, number of points and volume for each point and use of EMG/electrical stimulation in accordance with their usual practice, local summary of product characteristics and therapeutic guidelines.

Follow-up evaluation (time 2)

The timing of follow-up was at the discretion of the investigator, based on their usual practice and the nature of the goals set—usually between months 3 and 5 after injection. At time 2 (end of study), the following data were recorded:

- Achievement of primary and secondary GAS goals rated on a six-point verbal rating scale, and transcribed within the computer software to the five-point numerical scale (range −2 to +2), and the GAS T-score (see details below).
- Any concomitant treatments for upper limb spasticity given since baseline.
- Clinical examination including measurements of spasticity as was routinely performed.
- Global assessment of benefits were rated by the investigator and patient as either: 'great benefit'; 'some benefit'; 'same'; 'worse' or 'much worse'.
- Change on any standardised measures performed was recorded on the same five-point scale. (This was for

pragmatic reasons as it was not possible to accommodate the wide range of standardised measures that were used by different centres on the e-CRF without making it unwieldy).

- Related adverse events (AEs): as this was a non-interventional study, this followed the standard regulations for reporting of related spontaneous AEs within each country.
- The next therapeutic strategy—including information on any planned reinjection with BoNT-A—if so, whether to use the same agent and protocol or an adjusted one.

Sample size calculation

The sample size calculation was based upon an estimate that 60% of patients would achieve their primary goal following their first BoNT-A injection cycle. Using a 0.05 two-sided significance level, with a power of 80%, 450 patients were needed to allow estimation of this proportion with a precision of 4.5%. This sample size also allowed the detection of potential prognostic factors to response (based on detection of OR larger or equal to two).

Analysis

Statistical evaluations will be performed using Statistical Analysis System (SAS) (V.8 or later versions).

All analyses will be conducted on the efficacy population that includes all participants who received one BoNT-A injection and who underwent a postinjection visit including an assessment of the GAS. For the primary statistical analysis, 'responders' are those who achieve their primary goal (GAS score 0, 1 or 2) (primary statistical analysis). Patient demographics, baseline characteristics and efficacy evaluations for secondary variables will be presented as descriptive statistics, including 95% CIs where relevant.

Relationships between GAS T-scores and other measures of outcome (eg, measures of spasticity, global benefit and other standardised measures) will be examined using Spearman rank correlation coefficients. Stepwise logistic regression modelling will be used to identify potential prognostic factors for response including the duration of spasticity, time interval to follow-up, presence of confounding factors (including contractures, impaired cortical function, etc) and provision of concomitant therapy.

Descriptive analysis will also be carried out to evaluate any possible country effect, if considered relevant.

Method for recording GAS

We used the simplified 'GAS-light' approach described by Turner-Stokes,[28] which is based on the original method described by Kiresuk and Sherman,[22] but designed to be timely and practical for use in a busy clinic setting. The GAS-light method uses a verbal rating scale, which reflects the way clinicians usually record goals in routine practice. Verbal ratings are then translated into the five-point numerical scales as shown in figure 3.

Figure 3 Converting the goal attainment scaling-light verbal scoring system to a numerical five-point scale (−2 to +2). The verbal descriptions align with how clinicians normally think about and describe goal attainment. They allow goal attainment to be recorded without reference to the numeric scores, and so avoid the perceived negative connotations of zero and minus scores.

Verbal rating				Numerical conversion	
At Baseline	With respect to this goal do they have?	Some function	☐	-1	
		No function (as bad as they could be)	☐		-2
At Outcome: Was the goal achieved?	**Yes**	A lot more	☐	+2	+2
		A little more	☐	+1	+1
		As expected	☐	**0**	**0**
	No	Partially achieved	☐	-1	-1
		No change	☐	-1	-2
		Got worse	☐	-2	

▶ At the baseline visit, the investigator (in conjunction with their multidisciplinary team where possible) interviewed the patient and identified the main problem areas. An agreed set of goals (one primary and up to three secondary goals) was defined.

▶ A single SMART goal statement was recorded in the free text box of the e-CRF to describe the intended outcome for each goal (as opposed to the predefining levels for each score).

▶ Each primary and secondary goal was assigned to one of seven predefined goal categories ('pain', 'passive function', 'active function', 'mobility' (balance, gait), 'involuntary movement' (associated reaction), 'impairment' (eg, range of movement) and 'other').

▶ There was an option to weight goals for importance to the patient and/or family on a scale of 0=not at all, 1=a little, 2=moderately and 3=very important.[28] If no goal weighting was recorded, a default value of 1 was entered.
  – A baseline score was chosen for each goal as either 'some function' (−1) or 'no function' (as bad as they could be; −2).

▶ At the follow-up visit, goal attainment for each goal was recorded by the treating team, in conjunction with the patient, based on the review of the detailed goal statement.

▶ Attainment was rated on a 6-point verbal rating scale and transcribed to the five-point GAS numerical scale (−2 to +2), depending on the baseline rating as shown in figure 3.[39]

▶ A composite GAS T-score for each patient was then derived from the product of their individual goal achievement scores multiplied by goal weighting, using the following standard formula described by Kiresuk and Sherman:[22]

▶ The overall GAS T-score was calculated by the following formula:

$$\text{T-score} = 50 + \frac{10 \times \Sigma(w_i x_i)}{\sqrt{(0.7\Sigma w_i^2 + 0.3(\Sigma w_i)^2)}}$$

RIGOUR: STEPS TO ENSURE VALIDITY OF ULIS-II RESULTS

In view of the concerns raised about the use of GAS as noted in the introduction, a number of steps were taken to ensure the validity of GAS as the primary outcome measure for ULIS-II, which included:

▶ Selection of contributing centres;

▶ Training in SMART goal setting and the use of GAS prior to the study start;

▶ Within-study validation of goal statements;

▶ Establishment of an e-CRF with an inbuilt auditing facility for separation of time 1 (baseline, before the BoNT-A treatment) and time 2 (assessment after the treatment cycle) data.

Selection of contributing centres

Contributing centres were rigorously selected on the basis that they:

▶ Demonstrated a reasonably high and consistent level of data recording and outcome measurement as part of their routine clinical practice.

▶ Routinely collected at least one standardised measure of spasticity (Ashworth scale or Tardieu scale).

▶ Were formally trained in the use of GAS prior to recruitment.

Training in SMART goal setting and the use of GAS

In preparation for the ULIS-II study, a comprehensive GAS training programme was carried out across all participating centres. This programme not only educated clinicians who were unfamiliar with the use of GAS but also formed an essential part of the validation process for the use of GAS in ULIS-II.

First, a common training programme was established with a set of training tools, including:

▶ The practical guide to GAS, outlining the GAS-light method described previously[28]
  – A set of standard training slides in the form of a Microsoft PowerPoint presentation.

- A digital versatile/video disc (DVD) of three case examples to illustrate goal setting and recording of goal attainment in different scenarios.
- A standard Microsoft Excel spreadsheet for recording goal ratings applying the formula to derive a GAS T-score (as part of the learning process, this offered the opportunity for the teams to calculate their mean T-scores, providing feedback on whether they were over/under ambitious in their goal setting during the practice period).

Further information about the GAS-light method, including copies of the practical guide and GAS T-score calculator, may be found on the Cicely Saunders Institute website http://www.csi.kcl.ac.uk/gas-tool.html.

The training was organised through a series of national and regional workshops. A network of regional trainers was established as the leads for GAS training within their country/area. A series of initial 'Train-the-trainer' sessions was held to familiarise the local trainers with the training materials and to ensure that themselves were competent in using GAS before training others. Thereafter, all investigators at the participating centres attended GAS training (a total of 35 workshops) before recruiting patients.

Quality checks on goal setting

The ultimate focus of treatment for spasticity should be towards functional improvement, even though this may be at the level of passive function (ie, ease of caring for the affected limb), as opposed to active function (ie, active use of the limb in functional tasks). It is therefore expected that the primary goal statements should be both SMART and function-related in the majority of cases.

If goals are set in an unbiased fashion so that the results exceed and fall short of expectations in roughly equal proportions, the GAS T-score should be normally distributed at around a mean of 50 with an SD of around ±10.[28] If a team attempts to inflate its results by setting goals overcautiously, the mean score will be >50. Similarly, if it is consistently overambitious, the mean score will be <50. This provides a means for checking the overall quality of goal setting.

Validation of GAS goal statements

As a further step to ensure the validity of GAS, an interim validation process was undertaken part-way through the recruitment phase of the ULIS-II study. The purpose of this was to check that clinicians were setting SMART function-related goals in accordance with the training.

Goal statements for the primary goal in each patient were independently evaluated by three lead clinical investigators (LTS, KF and JJ) in two rounds.[40] In round 1 (September 2010), 345 goal statements from 67 centres and in round 2 (December 2010), 438 goal statements from 79 centres were evaluated.

Goal statements were examined on a centre-by-centre basis and investigators were blinded to country and centre. They were assessed on two criteria: (1) the WHO ICF domain and (2) the quality of the SMART description (table 3). It was accepted that for some patients, goal statements would be impairment related, for example, prevention of contractures. However, investigators expected that at least some goals from each centre would be function-related. Some examples of comparative SMART and non-SMART goals are shown in table 4.

After independent evaluation, ratings were compared and an overall centre rating for the WHO ICF domain

Table 3 Quality rating criteria for primary goal statements—WHO ICF domain and SMART description used during ULIS-II validation process

Rating	WHO 'ICF' domain, disability and health	Example
A	Some goal statements contain reference to functional activities at the level of disability or participation—may be 'active' or 'passive' function*	Reference to meaningful activities such as ease of self-care, reduced care burden, mobility, community-based activities, work-related function, etc
B	Goal statements contain reference to impairment only	Reference to movement, range, grip strength, spasticity, clonus, etc
C	Goal statements contain reference to anatomical structures only	Reference to extension, flexion, pronation, etc

Rating	'SMART' description	Example
++	There is a SMART goal description, sufficiently detailed and specific to make accurate GAS rating	'To be able to type a four-word sentence with only a single typing error using index fingers in 15 s'
+	There is some clear goal description sufficient to support GAS rating, but still reliant on subjective interpretation	'To be able to open and close hand, as well as use fingers more in household chores'
–	No clear goal description	'To use the hand more easily'

*'Active' function refers to using the affected limb in some motor activity, preferably for an identified functional purpose. 'Passive' function includes tasks related to caring for the affected limb (whether by a carer or by the person him/herself).
GAS, goal attainment scaling; ICF, International Classification of Functioning; SMART, specific, measurable, achievable, realistic and timed; ULIS, Upper Limb International Spasticity.

Table 4   SMART and function-related goals for GAS analysis

SMART goal statements	Non-SMART goal statements
To ease passive upper body dressing (<25% assistance) 1 month after injection	To improve ease of dressing upper limb
Improve carry angle from 25° to 0° when walking 1 month after injection	Elbow extension
To reduce upper limb pain during rest and passive range of motion (<4/10 on VAS) 1 month after injection	To improve pain
To relieve thumb in palm and ease nail clipping (taking less than 20 min) 1 month after injection	Easier thumb and finger extensions
Sitting at the table, to grip fork and spoon to move them to mouth and eat 1 month after injection	To improve grasp and release function of the hand

SMART, specific, measurable, achievable, realistic and timed; GAS, goal attainment scaling; VAS, visual analogue scale.

(A, B or C) and SMART description (++, + or −) was reached by consensus. Results were then fed back to the centres. The goal was to have a rating of A+ or A++. In order to improve the quality of goal setting, centres with lower rates (B, C or −) were invited to submit revised goal statements for those patients who had not yet received their follow-up evaluation.

The results from the interim validation of GAS goal statements are shown in figure 4A. In round 1, 62.7% recorded function-related statements rated ('A' or 'AB') and 40.3% received a SMART quality rating of '++'/'+'. In round 2, these figures rose to 70.9% and 46.8%, respectively.

After the goal refinement process, 37 centres submitted revised goal statements (of which 21 (56.8%) had improved their rating), and 12 centres (24%) had achieved a maximum possible combined rating of 'A++'. Even after this process, however, there was residual heterogeneity between countries in the quality of goal setting, especially with respect to 'SMARTness', as illustrated in figure 4B.

In any multicentre study reflecting real-life practice, one would expect a range of quality in goal setting and we do not think that ULIS-II is unique in this respect. Therefore, we do not plan to discard the goals that were not SMART from the primary analysis, but we will perform a secondary analysis to examine the relationship between GAS and the quality of goal setting, testing the hypothesis that goal achievement rates are lower for SMART goals (see below). If a significant difference is found, subset analyses will be conducted to correct for this.

Development of the e-CRF

A further purpose of ULIS-II was to develop and refine the e-CRF for capturing a standardised dataset. In ULIS-I, we identified the most commonly used approaches to record assessment, treatment and outcome evaluation. Learning from this experience, we addressed ways of reducing the information and making it understandable by investigators in all countries. This was achieved through small international group workshop discussions with key investigators (including authors LTS, KF and JJ), followed by circulation and

feedback on the draft (case report form, prior to the start of development of the electronic version).

The e-CRF was presented in one language only (English) to avoid introducing variance through translation. Investigators therefore needed to be adequately proficient in the English language to complete the form, although goal statements could be completed in
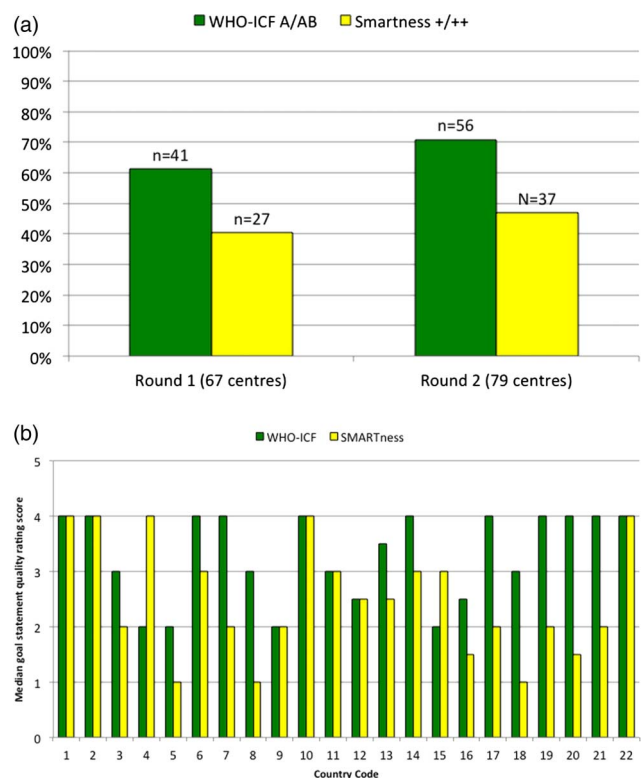
Figure 4   (A) Percentage of centres achieving high quality ratings in rounds 1 and 2 during validation of goal attainment scaling statements. See table 1 for description of WHO International Classification of Functioning (ICF) domains A, B and C, and specific, measurable, achievable, realistic and timed (SMART) descriptor ratings ++, + and −. (B) Median quality rating of final goal statements by participating country. Goal quality ratings were derived from the WHO ICF domain rating (A=4, A/B=3, B=2 and C=1) and the SMART rating ('++'=4, '+'=3, '+/−'=2 and '−'=1). Each centre was assigned two goal quality ratings and the graph shows the medians for each participating country.

the native language. Site monitors provided support for any interpretational queries. The e-CRF was posted on the internet to facilitate access to all participating countries. It was developed on a customised software that supported electronic and interactive data collection, as well as online edit checks to ensure data accuracy and quality. Wherever possible, it included dropdown value lists and check-box options to minimise ambiguity, and it also provided automated calculation of the GAS T-score.

Separation of time 1 and 2 data

In this before-and-after study, time 1 acts as a 'within patient' control. GAS rating requires the investigator to be able to view the goals set at time 1, so it was important to maintain the independence of time 1 and time 2 data as far as possible. The e-CRF software did not have the facility to lock the time 1 data prior to completion of the follow-up form, which would have prevented the opportunity for retrospective changes. However, it did have an inbuilt tracking facility. Site monitors were responsible for reviewing the tracking log to ensure that no post hoc changes were made that could have influenced outcome evaluation.

STRENGTHS AND LIMITATIONS

The study builds on others that have used goal attainment to evaluate outcome from interventions for upper limb spasticity.[23–26] It will expand our understanding of the types of goals that are and are not likely to be achieved following treatment with BoNT-A. Importantly, it sets in train a methodology that is practical for use in routine clinical practice, which will be used in future studies to expand the clinical dataset to one of sufficient size to interrogate for answers to the important questions including:

1. The identification and selection of patients most likely to benefit from treatment.
2. The most effective approaches to muscle selection, injection technique, etc.
3. The most useful approaches to outcome measurement.

Goal setting is a complex process, which requires skill and experience on the part of investigators—both to coin the SMART goal description and also to be able to predict the likely outcome of the intervention and its timescale. This study has emphasised the need for training in consistent goal-setting techniques as highlighted in other studies.[24] Our interim goal validation study was an important step to ensure the rigour of goal setting in participating centres. In future ULIS studies, demonstration of competency in high quality, unbiased goal setting will be an essential prerequisite for centre participation. We recommend that future trials using GAS in this context should adopt a similarly robust approach.

The strengths of our approach include the following:

▶ The wide international representation of participating centres captures the experience of clinicians from around the globe, ensuring the generalisability of results.

▶ The use of goal setting and GAS in this context emphasises the assessment of outcomes that are important to the patient.

▶ The simplified approach to GAS enabled its application in clinical practice and supported the negotiation of realistic expectations for outcome.

▶ The comprehensive GAS training programme allowed investigators unfamiliar with this method to be educated in goal setting prior to the start of the study.

▶ The interim validation process assisted participating centres to improve the quality and function-related focus of goal statements, supporting the eventual validity of GAS as the primary outcome measure for this and future ULIS studies. At first sight, this might be regarded as 'cheating'. However, SMARTening goal statements would, if anything, make goals harder to achieve. For example, the broad goal statement 'to improve pain' would be achieved with some improvement, while the SMART equivalent 'to reduce upper limb pain during rest and passive range of motion (<4/10 on the visual analogue scale (VAS)) 1 month after injection' would only be achieved if the specified targets were achieved.

Recognised limitations:

1. Limiting recruits to 5–12 per centre may have introduced some selection bias through under-representation of less common presentations of spasticity.
2. The e-CRF was not locked after time 1 data entry, allowing for the possibility of retrospective alteration (of the goal statement only) at time 2. Using the e-CRF auditing facility, no cases were found in which goal statements were retrospectively adjusted at time 2. In future studies, it will be important to use software with a locking facility.

In this study, it was not feasible to record raw data for outcome measures other than GAS and spasticity (MAS). Aside from this, investigators were simply asked to record which measures they had used and whether the results were 'the same', 'better' or 'worse' than at baseline. Given the long list of measures used, it was not feasible or appropriate to compute all of them within the e-CRF. However, this information will allow us to identify those measures that are sensitive to change in the responder population, which will assist in the selection of an appropriate and feasible battery of standardised measures to record alongside GAS in future ULIS studies.

SUMMARY

The importance of the ULIS programme lies in its staged approach to development, which both educates participating centres and ensures that the final core dataset will capture the cultural diversity of worldwide clinical practices in which it will be used. ULIS-II is not the final step, but marks an important phase in the development of the programme.

Despite the recognised limitations at this stage of development, ULIS-II will provide a unique and important set

of information regarding the treatment and outcomes from BoNT-A in real-life management of poststroke upper limb spasticity worldwide. Through the process described above, we have developed a methodology that will not only help to ensure credibility of results from ULIS-II, but will also underpin future studies and inform other clinical trials and cohort studies in this context.

ETHICS AND DISSEMINATION

▸ Marketing authorisation for the use of BoNT-A in this context was ensured for each participating country prior to the start of the study.
▸ As the study was non-interventional, it did not fall under the scope of the EU Directive 2001/20/EC of the European Parliament. It was conducted in compliance with guidelines for Good Pharmacoepidemiology Practices (http://www.pharmacoepi.org/resources/guidelines_08027.cfm).
▸ To reflect real-life practice in this observational study, physicians were free to choose BoNT-A treatment (targeted muscles, preparation, injected doses and number of points and volume for each point in accordance with their usual practice, and with their local Summary of Product Characteristics and therapeutic guidelines).
▸ Ethical approval and written informed consent were obtained prior to anonymous data collection in countries where this was required and the study protocol was approved by an independent ethics committee at each participating site.
▸ Data protection: All personal information was collected at the investigational site and protected according to local data protection law. Patients were identifiable only by patient ID at the investigational site.
▸ The results will be presented at international meetings and published in peer-reviewed journals.

Author affiliations
[1]Department of Palliative Care, Policy and Rehabilitation, Cicely Saunders Institute, School of Medicine, King's College London, London, UK
[2]Department of Neurorehabilitation, Gailtal-Klinik, Hermagor, Austria
[3]Serviço de Reabilitação de adultos 3, Centro de Medicina de Reabilitaçãode Alcoitão, Estoril, Portugal
[4]Departments of Biostatistics & Data Management, Medical Affairs, Ipsen Pharma, Boulogne-Billancourt, France

REFERENCES

1. Leathley MJ, Gregson JM, Moore AP, *et al*. Predicting spasticity after stroke in those surviving to 12 months. *Clin Rehabil* 2004;18:438–43.
2. Lundstrom E, Terent A, Borg J. Prevalence of disabling spasticity 1 year after first-ever stroke. *Eur J Neurol* 2008;15:533–9.
3. Sommerfeld DK, Eek EU, Svensson AK, *et al*. Spasticity after stroke: its occurrence and association with motor impairments and activity limitations. *Stroke* 2004;35:134–9.
4. Watkins CL, Leathley MJ, Gregson JM, *et al*. Prevalence of spasticity post stroke. *Clin Rehabil* 2002;16:515–22.
5. Welmer AK, von Arbin M, Holmqvist L Widen, *et al*. Spasticity and its association with functioning and health-related quality of life 18 months after stroke. *Cerebrovasc Dis* 2006;21:247–53.
6. Welmer AK, Holmqvist L Widen, Sommerfeld DK. Location and severity of spasticity in the first 1–2 weeks and at 3 and 18 months after stroke. *Eur J Neurol* 2010;17:720–5.
7. Spasticity in adults: Management using Botulinum Toxin. *National guidelines*. London: Royal College of Physicians, 2009.
8. Sheean G, Lannin NA, Turner-Stokes L, *et al*. Botulinum toxin assessment, intervention and after-care for upper limb hypertonicity in adults: international consensus statement. *Eur J Neurol* 2010; (Suppl 2):74–93.
9. Ward AB. Spasticity treatment with botulinum toxins. *J Neural Transm* 2008;115:607–16.
10. Bakheit AM, Sawyer J. The effects of botulinum toxin treatment on associated reactions of the upper limb on hemiplegic gait–a pilot study. *Disabil Rehabil* 2002;24:519–22.
11. Bhakta BB, Cozens JA, Chamberlain MA, *et al*. Impact of botulinum toxin type A on disability and carer burden due to arm spasticity after stroke: a randomised double blind placebo controlled trial. (erratum appears in *J Neurol Neurosurg Psychiatry* 2001 Jun;70(6):821). *J Neurol Neurosurg Psychiatr* 2000;69:217–21.
12. Bhakta BB, Cozens JA, Bamford JM, *et al*. Use of botulinum toxin in stroke patients with severe upper limb spasticity. *J Neurol Neurosurg Psychiatr* 1996;61:30–5.
13. Brashear A, Gordon MF, Elovic E, *et al*. Intramuscular injection of botulinum toxin for the treatment of wrist and finger spasticity after a stroke. *NEJM* 2002;347:395–400.
14. Hesse S, Jahnke MT, Luecke D, *et al*. Short-term electrical stimulation enhances the effectiveness of Botulinum toxin in the treatment of lower limb spasticity in hemiparetic patients. *Neurosci Lett* 1995;201:37–40.
15. Richardson D, Edwards S, Sheean GL, *et al*. The effect of botulinum toxin on hand function after incomplete spinal cord injury at the level of C5/6: a case report. *Clin Rehabil* 1997;11:288–92.
16. Rodriquez AA, McGinn M, Chappell R. Botulinum toxin injection of spastic finger flexors in hemiplegic patients. *Am J Phys Med Rehabil* 2000;79:44–7.
17. 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.
18. Smith SJ, Ellis E, White S, *et al*. A double-blind placebo-controlled study of botulinum toxin in upper limb spasticity after stroke or head injury. *Clin Rehabil* 2000;14:5–13.
19. DeJong G, Horn SD, Conroy B, *et al*. Opening the black box of post-stroke rehabilitation: stroke rehabilitation patients, processes, and outcomes. (see comment). *Arch Phys Med Rehabil* 2005;86:S1–7.
20. Francis HP, Wade DT, Turner-Stokes L, *et al*. Does reducing spasticity translate into functional benefit? An exploratory meta-analysis. *J Neurol Neurosurg Psychiatr* 2004;75:1547–51.
21. Turner-Stokes L, Ashford S, Nair A. Physical therapy and botulinum toxin-A (BoNT-A)—the temporal relationship between spasticity

reduction and functional gain (AAPMR Annual Assembly Poster 14). *Phys Med Rehabil* 2010;2:S13–14.

22. Kiresuk T, Sherman R. Goal attainment scaling: a general method of evaluating comprehensive mental health programmes. *Community Ment Health J* 1968;4:443–53.
23. Ashford S, Turner-Stokes L. Goal attainment for spasticity management using botulinum toxin. *Physiother Res Int* 2006;11:24–34.
24. Turner-Stokes L, Baguley I, De Graff S, *et al*. Goal attainment scaling in the evaluation of treatment of upper limb spasticity with botulinum toxin: a secondary analysis from a double blind placebo controlled randomised clinical trial. *J Rehabil Med* 2010;42:81–9.
25. McCrory P, Turner-Stokes L, Baguley IJ, *et al*. Botulinum toxin A for treatment of upper limb spasticity following stroke: a multi-centre randomised placebo-controlled study of the effects on quality of life and other person-centred outcomes. *J Rehabil Med* 2009;41:536–44.
26. Borg J, Ward AB, Wissel J, *et al*. Rationale and design of a multicentre, double-blind, prospective, randomized, European and Canadian study: evaluating patient outcomes and costs of managing adults with post-stroke focal spasticity. *J Rehabil Med* 2011;43:15–22.
27. Scobbie L, Dixon D, Wyke S. Goal setting and action planning in the rehabilitation setting: development of a theoretically informed practice framework. *Clin Rehabil* 2011;25:468–82.
28. Turner-Stokes L. Goal attainment scaling in rehabilitation; a practical guide. *Clin Rehabil* 2009;23:362–70.
29. Tennant A. Goal attainment scaling: current methodological challenges. *Disabil Rehabil* 2007;29:1583–8.
30. Yip AM, Gorman MC, Stadnyk K, *et al*. A standardized menu for Goal Attainment Scaling in the care of frail elders. *Gerontologist* 1998;38:735–42.
31. *International classification of functioning, disability and health*. Geneva: World Health Organisation, 2002: http://www.who.int/classifications/icf/en/ (accessed Jan 2013)
32. Baguley IJ, Nott MT, Turner-Stokes L, *et al*. Investigating muscle selection for botulinum toxin-A injections in adults with post-stroke upper limb spasticity. *J Rehabil Med* 2011;43:1032–7.
33. Horn SD, Gassaway J. Practice-based evidence study design for comparative effectiveness research. *Med Care* 2007;45:S50–7.
34. Bakheit AM, Zakine B, Maisonobe P, *et al*. The profile of patients and current practice of treatment of upper limb muscle spasticity with botulinum toxin type A: an international survey. *Int J Rehabil Res* 2010;33:199–204.
35. Turner-Stokes L, Siegert RJ, Thu A, *et al*. The neurological impairment scale: a predictor of functional outcome in patients with severe complex disability (AAPMR Annual Assembly Poster 38). *Phys Med Rehabil* 2011;3:S189.
36. Ashford S, Slade M, Turner-Stokes L. Conceptualisation and development of the arm activity measure (ArmA) for assessment of activity in the hemiparetic arm. *Disabil Rehabil* 2013. In press. Published Online January 2013.
37. Ashford S, Turner-Stokes L, Siegert RJ, *et al*. Initial psychometric evaluation of the Arm Activity measure (ArmA)—a measure of active and passive function in the hemiparetic arm. *Clin Rehabil*. Published online before print 20 February 2013 doi: 10.1177/0269215512474942.
38. Macfarlane A, Turner-Stokes L, De Souza L. The associated reaction rating scale: a clinical tool to measure associated reactions in the hemiplegic upper limb. *Clin Rehabil* 2002;16:726–35.
39. Turner-Stokes L, Williams H. Goal attainment scaling: a direct comparison of alternative rating methods. *Clin Rehabil* 2010;24:66–73.
40. Turner-Stokes L, Fheodoroff K, Jacinto J. Validation of individual person-centered goal statements for goal attainment scaling in a large international cohort study of botulinum toxin type A for upper limb spasticity after stroke. *Phys Med Rehabil* 2011;3:s338.