ARD Online First, published on July 29, 2016 as 10.1136/annrheumdis-2016-210002

CONCISE REPORT

Development of a Glucocorticoid Toxicity Index (GTI) using multicriteria decision analysis

Eli M Miloslavsky,[1] Ray P Naden,[2] Johannes W J Bijlsma,[3] Paul A Brogan,[4] E Sherwood Brown,[5] Paul Brunetta,[6] Frank Buttgereit,[7] Hyon K Choi,[8] Jean-Francois DiCaire,[9] Jeffrey M Gelfand,[10] Liam G Heaney,[11] Liz Lightstone,[12] Na Lu,[13] Dedee F Murrell,[14] Michelle Petri,[15] James T Rosenbaum,[16] Kenneth S Saag,[17] Murray B Urowitz,[18] Kevin L Winthrop,[19] John H Stone[20]

Handling editor Tore K Kvien

▶ Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/annrheumdis-2016-210002).

For numbered affiliations see end of article.

Correspondence to
Dr John H Stone, Rheumatology Clinic, Yawkey 2, Massachusetts General Hospital, 55 Fruit Street, Boston, MA 02114, USA; jhstone@mgh.harvard.edu

Received 1 June 2016
Accepted 11 July 2016

To cite: Miloslavsky EM, Naden RP, Bijlsma JWJ, *et al*. *Ann Rheum Dis* Published Online First: [*please include* Day Month Year] doi:10.1136/annrheumdis-2016-210002

ABSTRACT

Objectives To develop a Glucocorticoid Toxicity Index (GTI) to assess glucocorticoid (GC)-related morbidity and GC-sparing ability of other therapies.

Methods Nineteen experts on GC use and outcome measures from 11 subspecialties participated. Ten experts were from the USA; nine from Canada, Europe or Australia. Group consensus methods and multicriteria decision analysis (MCDA) were used. A Composite GTI and Specific List comprise the overall GTI. The Composite GTI reflects toxicity likely to change during a clinical trial. The Composite GTI toxicities occur commonly, vary with GC exposure, and are weighted and scored. Relative weights for items in the Composite GTI were derived by group consensus and MCDA. The Specific List is designed to capture GC toxicity not included in the Composite GTI. The Composite GTI was evaluated by application to paper cases by the investigators and an external group of 17 subspecialists.

Results Thirty-one toxicity items were included in the Composite GTI and 23 in the Specific List. Composite GTI evaluation showed high inter-rater agreement (investigators κ 0.88, external raters κ 0.90). To assess the degree to which the Composite GTI corresponds to expert clinical judgement, participants ranked 15 cases by clinical judgement in order of highest to lowest GC toxicity. Expert rankings were then compared with case ranking by the Composite GTI, yielding excellent agreement (investigators weighted κ 0.87, external raters weighted κ 0.77).

Conclusions We describe the development and initial evaluation of a comprehensive instrument for the assessment of GC toxicity.

INTRODUCTION

Glucocorticoids (GCs) have been a cornerstone of treatment for many diseases since their introduction more than 65 years ago. GC use is associated with considerable treatment morbidity.[1 2] Although the use of these medications is generally reviled by patients and physicians alike, data on the true incidence of GC-associated adverse events remain scarce because until now GC toxicity has simply been a fact of life for patients with immune-mediated diseases.[3] The development of novel immunomodulatory agents offers the potential to reduce GC use and to diminish their adverse effects.[4 5] In order to assess the true benefit of new medications with regard to their steroid-sparing properties, investigators must be able to assess their ability to prevent or reverse GC-related adverse events. Unfortunately, no reliable instrument designed to measure GC-related toxicity both broadly and accurately has been developed.

Measuring GC-related toxicity poses significant challenges.[1 6] Previous studies examining GC-related toxicity have used different combinations of adverse events with varied event definitions.[7–9] We aimed to develop a Glucocorticoid Toxicity Index (GTI) useful across medical disciplines to assess the impact of GC-associated morbidity.

METHODS

Participants and procedures

Twenty-two experts in GC use and outcome measures were invited and 19 agreed to serve on the Scientific Committee (SC). Experts represented multiple specialties (rheumatology (including osteoporosis), paediatrics rheumatology, pulmonology, nephrology, neurology, ophthalmology, dermatology, infectious disease and psychiatry) and had extensive experience in the clinical use and pharmacology of GCs. Ten investigators were from the USA, nine from Canada, Europe or Australia.

The development process, which included 10 milestones (figure 1), was conducted over 10 one-hour conference calls, work between the calls and one daylong, face-to-face meeting.

Instrument characteristics and item inclusion criteria

The SC agreed that the optimal use of the GTI would be in prospective, randomised, controlled clinical trials using GCs, regardless of whether GC therapy is prescribed according to protocol or investigators' best medical judgement. Randomisation and blinding serve the critical purposes of controlling for the background rate of adverse events[10] and prior GC treatment, and also limit the need for attribution.

The SC determined that the GTI would have two components: the Composite GTI and a Specific List. The Composite GTI serves as the primary instrument and is intended to capture common toxicities that are sensitive to differing cumulative GC doses over the period of a typical clinical trial (6 months to 3 years). It is weighted

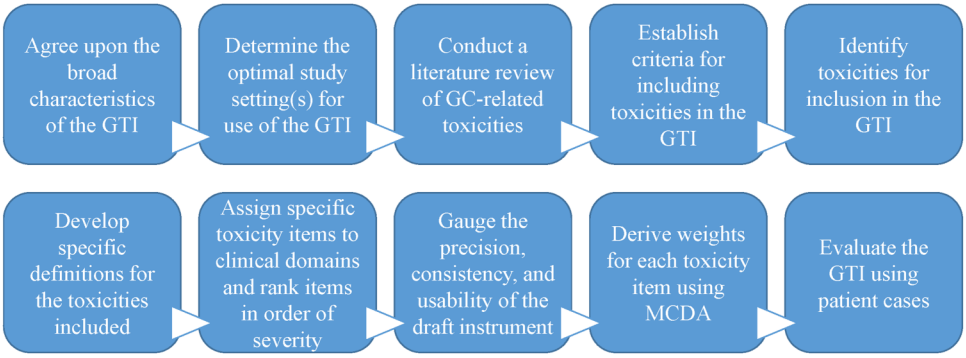

Figure 1 GTI development milestones. GC, glucocorticoid; GTI, Glucocorticoid Toxicity Index; MCDA, multicriteria decision analysis.

and measures both worsening and improvement. The complementary Specific List captures important GC-related adverse events not included in the Composite GTI. The SC agreed to not weigh Specific List toxicities due to the possible skewing that rare but serious events would introduce into the weighting scheme.

Item selection for the Composite GTI was based on the following principles: (1) likelihood of occurrence >5% in patients exposed to GCs; (2) item independence; (3) item equivalence (several GC toxicities could be included within a single item, provided they were within the same clinical domain and were equivalent in their degree of toxicity); (4) toxicity is more likely to be due to the effect of GC therapy than the disease itself; (5) toxicity is unlikely to be the result of GC therapy prior to trial entry (eg, osteoporotic fracture); (6) measurement does not typically require invasive procedures or imaging.

Toxicities that did not meet these criteria but were deemed important and were not confounded by underlying disease or comorbidities were included in the Specific List. Candidate toxicities were generated based on literature review (see online supplementary appendix I) and selected for inclusion by nominal group technique. Definitions for each item, developed by experts from the relevant clinical area, were revised by consensus. Items were grouped by clinical domains in order of increasing toxicity such that only one item within each domain could be assigned to a given patient. The draft GTI was reviewed by the SC for clarity, format, visual design, organisation and navigability. Relative weights were then derived at the face-to-face meeting using multicriteria decision analysis (MCDA) via the 1000Minds software platform (Dunedin, New Zealand) (see online supplementary appendix II).[11 12]

Instrument scoring

The SC agreed that the Composite GTI should measure change in GC toxicity rather than absolute GC toxicity in order to account for the effects of prior GC therapy and background rate of adverse events. Therefore, evaluation at two time points is required for scoring. All domains have the potential for improvement (eg, myopathy can improve from 'mild' to 'none', even though a specific improvement item is not included in the Composite GTI). When a Specific List item occurs (eg, death from infection), the most severe corresponding item in the Composite GTI (ie, Grade III infection) is also scored. The Composite GTI should be scored at 3-month intervals throughout the study, using entry assessment as the baseline. Because bone mineral density studies should generally not be performed more often than every 12 months, the bone domain should be excluded for trials shorter than 1 year in duration. The score should be reported as both a total score and domain-specific scores, to account for scenarios when improvements in certain domains compensate for worsening in others.

Evaluation process

The performance of the Composite GTI was evaluated by both participating experts and an external, multispecialty group of 17 testers (see online supplementary appendix V and table S1) using paper cases. Each expert submitted four patient cases describing GC toxicity. Fifteen cases were chosen to represent the full range of GC toxicity. Both the experts and external testers then completed an on-line exercise composed of two tasks: (1) rank cases in order of greatest to least GC-toxicity (experts' rankings were then compared with the ranking assigned by the weighted Composite GTI); and, (2) assign Composite GTI items to each case.

Statistical analysis

Inter-rater reliability among raters and agreement between the experts' and external testers' rankings and those of the Composite GTI were assessed using the κ statistic. The overall inter-rater reliability of the ranking agreements was then calculated by averaging pairwise κ values. All statistical analyses were performed on SAS V.9.3 (SAS Institute, Cary, North Carolina, USA).

RESULTS

Nine domains and 31 items were included in the Composite GTI (table 1). Eleven domains and 23 items were included in the Specific List (table 1) (see definitions, online supplementary appendices III and IV). Items reflect severity and account for impact of medications (eg, blood pressure can be stable due to an increase in antihypertensive regimen). Toxicities such as atherosclerosis, myocardial infarction and stroke were not included in the GTI because the SC agreed that all are confounded by comorbid conditions (eg, smoking) or disease effects (eg, systemic lupus erythematosus).[13] Except for bone mineral density, included because of its importance in GC-related toxicity,[14] items requiring imaging were excluded from the Composite GTI.

Fifteen experts participated in the weighting exercise at the face-to-face meeting. Seventeen of 19 experts and 17 independent raters completed this evaluation phase. The inter-rater reliability exercise revealed a high degree of agreement, with a κ of 0.88 ($p<0.01$) for participating experts and a κ of 0.90 ($p<0.01$) for independent raters. The initial validity exercise revealed that both expert and independent rater case rankings had excellent agreement with rankings by the Composite GTI, with a weighted κ of 0.87 ($p<0.01$) and 0.77 ($p<0.01$), respectively.

Table 1 The Glucocorticoid Toxicity Index

Composite GTI	Item weight	Specific List
BMI		
Improvement in BMI	–8	Major increase in BMI
No change in BMI	0	
Moderate increase in BMI	21	
Major increase in BMI	36	
Glucose tolerance		
Improvement in glucose tolerance	–8	Diabetic retinopathy
No change in glucose tolerance	0	Diabetic nephropathy
Worsening of glucose tolerance	32	Diabetic neuropathy
Worsening of glucose tolerance despite treatment	44	
Blood pressure		
Improvement in blood pressure	–10	Hypertensive emergency
No change in blood pressure	0	Posterior reversible encephalopathy syndrome
Worsening hypertension	19	
Worsening hypertension despite treatment	44	
Lipids		
Improvement in lipids	–9	
No change in lipids	0	
Worsening hyperlipidaemia	10	
Worsening hyperlipidaemia despite treatment	30	
Bone density		
Improvement in bone density	–1	Major decrease in bone density
No change in bone density	0	Insufficiency fracture
Decrease in bone density	29	
Steroid myopathy		
No steroid myopathy	0	Severe steroid myopathy
Mild steroid myopathy	9	
Moderate steroid myopathy or greater	63	
Skin toxicity		
No skin toxicity	0	Severe skin toxicity
Mild skin toxicity	8	
Moderate skin toxicity or greater	26	
Neuropsychiatric toxicity		
No neuropsychiatric symptoms	0	Psychosis
Mild neuropsychiatric symptoms	11	GC-induced violence
Moderate neuropsychiatric symptoms or greater	74	Other severe neuropsychiatric symptoms
Infection		
No significant infection	0	Grade IV infection
Oral/vaginal candidiasis or uncomplicated zoster	19	Grade V infection
Grade III infection or greater	93	
Endocrine		Adrenal insufficiency
Gastrointestinal		Perforation
		Peptic ulcer disease
Musculoskeletal		Avascular necrosis
		Tendon rupture
Ocular		Central serous retinopathy
		Intraocular pressure elevation
		Posterior subcapsular cataract
Total	–36 to 439	

BMI, body mass index; GC, glucocorticoid; GTI, Glucocorticoid Toxicity Index.

DISCUSSION

A useful measurement of the steroid-sparing ability of new treatment agents requires a reliable outcomes-based instrument of GC-related toxicity.[15 16] We describe a multispecialty effort to develop the GTI, a comprehensive measure of change in GC-toxicity over time. The initial evaluation of the Composite GTI by participating experts and a multispecialty group of external testers demonstrated excellent reliability and validity.

The development of two complementary assessment instruments within the GTI—the Composite GTI and the Specific List—was crucial in addressing several challenges in measuring GC toxicity. The creation of the Specific List permits documentation of certain important and often severe toxicities, leaving the Composite GTI as a relatively concise and easy-to-administer tool intended to detect differences between patients receiving divergent GC amounts. The inclusion of rare toxicities and those that may reflect prior GC use in the Specific List allowed us to simplify the usability, limit weight skewing and minimise the effect of pretrial GC therapy on the Composite GTI.

An important strength of the Composite GTI is the assignment of relative weights to each toxicity item in a systematic manner using MCDA.[11] The MCDA approach greatly enhances the feasibility of this complex task in a way that group consensus methods struggle to approach. Further, the MCDA approach allows us to perform modifications of the Composite GTI as new data become available, including the addition and weighting of new items, without disrupting the validity of the method.

The next phase in GTI development includes the development of a web-based interface, prospective use in clinical trials and input from patient support groups. Our initial evaluation exercise of the Composite GTI, including testing by an external group of GC experts, implies excellent performance characteristics. The development of a web-based interface should further increase the instrument's reliability. For the GTI to be truly valid, it must be assessed in clinical trials and compared with doses of GCs administered, quality-of-life measures and damage indices that include GC toxicity.[17 18]

In conclusion, we describe the development and initial evaluation of the GTI, a comprehensive GC toxicity assessment instrument. The GTI can be used across disciplines to assess the clinical value of steroid-sparing therapies, as well as to measure the impact of GC toxicity. Given the widespread use of GCs and the accelerating pace of immunological drug discovery, this instrument represents a considerable advance in our ability to assess the utility of new pharmacological agents.

Author affiliations
[1]Rheumatology, Allergy and Immunology Division, Massachusetts General Hospital, Boston, Massachusetts, USA
[2]Maternal-Fetal Medicine, McMaster University Faculty of Health Sciences, Hamilton, Ontario, Canada
[3]Department of Rheumatology, UMCUtrecht, Utrecht, Netherlands
[4]Institute of Child Health, University College London, UCL Inst of Child Health, London, UK
[5]Department of Psychiatry, University of Texas Southwestern Medical Center at Dallas, Dallas, Texas, USA
[6]Late Stage Immunology Product Development, Genentech, Inc., South San Francisco, USA
[7]Department of Rheumatology and Immunology, Charité University Medicine Berlin, Berlin, Germany
[8]Department of Rheumatology, Harvard Medical School, Boston, Massachusetts, USA
[9]Pinnacle, Inc., Montreal, Quebec, Canada
[10]University of California-San Francisco, San Francisco, USA
[11]Queen's University of Belfast, Belfast, UK
[12]Section of Renal Medicine and Vascular Inflammation, Division of Immunology and Inflammation, Department of Medicine, Imperial College London, Imperial College London, London, UK

[13]Department of Rheumatology, Massachusetts General Hospital, Boston, USA
[14]University of New South Wales, Sydney, New South Wales, Australia
[15]Department of Rheumatology, Johns Hopkins University, Baltimore, Maryland, USA
[16]Casey Eye Institute, Oregon Health and Science University, Portland, Oregon, USA
[17]UAB Division of Clinical Immunology/Rheumatology, University of Alabama at Birmingham, Birmingham, Alabama, USA
[18]Center for Prognosis Studies in the Rheumatic Diseases, Toronto Western Hospital, University of Toronto, Lupus Clinic, Toronto, Canada
[19]Oregon Health Sciences University, Portland, Oregon, USA
[20]Massachusetts General Hospital Rheumatology Unit, Harvard Medical School, Rheumatology Clinic, Boston, Massachusetts, USA

Twitter Follow Liz Lightstone at @kidneydoc101

Contributors All of the named authors have contributed to the design, conduct, and analysis of this study. All have contributed to writing and editing the manuscript. All fulfil ICJME criteria for authorship, and all have approved the final manuscript.

Funding This study was funded by an investigator-initiated grant from Genentech.

Competing interests None declared.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement Data from the study (published and unpublished) are available upon written request to the corresponding author.

REFERENCES

1 Mcdonough AK, Curtis JR, Saag KG. The epidemiology of glucocorticoid-associated adverse events. *Curr Opin Rheumatol* 2008;20:131–7.
2 Sarnes E, Crofford L, Watson M, *et al*. Incidence and US costs of corticosteroid-associated adverse events: a systematic literature review. *Clin Ther* 2011;33:1413–32.
3 Da Silva JA, Jacobs JW, Kirwan JR, *et al*. Safety of low dose glucocorticoid treatment in rheumatoid arthritis: published evidence and prospective trial data. *Ann Rheum Dis* 2006;65:285–93.
4 Bel EH, Wenzel SE, Thompson PJ, *et al*. Oral glucocorticoid-sparing effect of mepolizumab in eosinophilic asthma. *N Engl J Med* 2014;371:1189–97.
5 Fortunet C, Pers YM, Lambert J, *et al*. Tocilizumab induces corticosteroid sparing in rheumatoid arthritis patients in clinical practice. *Rheumatology (Oxford)* 2015;54:672–7.
6 Van Der Goes MC, Jacobs JWG, Boers M, *et al*. Monitoring adverse events of low-dose glucocorticoid therapy: EULAR recommendations for clinical trials and daily practice. *Ann Rheum Dis* 2010;69:1913–19.
7 Rutgeerts P, Löfberg R, Malchow H, *et al*. A comparison of budesonide with prednisolone for active Crohn's disease. *N Engl J Med* 1994;331:842–5.
8 Capell HA, Madhok R, Hunter JA, *et al*. Lack of radiological and clinical benefit over two years of low dose prednisolone for rheumatoid arthritis: results of a randomised controlled trial. *Ann Rheum Dis* 2004;63:797–803.
9 Wassenberg S, Rau R, Steinfeld P, *et al*. Very low-dose prednisolone in early rheumatoid arthritis retards radiographic progression over two years: a multicenter, double-blind, placebo-controlled trial. *Arthritis Rheum* 2005;52:3371–80.
10 Black S, Eskola J, Siegrist C-A, *et al*. Importance of background rates of disease in assessment of vaccine safety during mass immunisation with pandemic H1N1 influenza vaccines. *Lancet (London, England)* 2009;374:2115–22.
11 Johnson SR, Naden RP, Fransen J, *et al*. Multicriteria decision analysis methods with 1000Minds for developing systemic sclerosis classification criteria. *J Clin Epidemiol* 2014;67:706–14.
12 Neogi T, Jansen TL, Dalbeth N, *et al*. 2015 Gout classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Ann Rheum Dis* 2015;74:1789–98.
13 Nikpour M, Urowitz MB, Gladman DD. Epidemiology of atherosclerosis in systemic lupus erythematosus. *Curr Rheumatol Rep* 2009;11:248–54.
14 Weinstein RS. Clinical practice. Glucocorticoid-induced bone disease. *N Engl J Med* 2011;365:62–70.
15 Bell GM, Reynolds G, Isaacs JD. Biologic therapies in non-rheumatic diseases: lessons for rheumatologists? *Nat Rev Rheumatol* 2011;7:507–16.
16 United States Food and Drug Administration. *Guidance for Industry Systemic Lupus Erythematosus — Developing Medical Products for Treatment*.
17 Ruiz-Arruza I, Ugarte A, Cabezas-Rodriguez I, *et al*. Glucocorticoids and irreversible damage in patients with systemic lupus erythematosus. *Rheumatology (Oxford)* 2014;53:1470–6.
18 Exley AR, Bacon PA, Luqmani RA, *et al*. Development and initial validation of the Vasculitis Damage Index for the standardized clinical assessment of damage in the systemic vasculitides. *Arthritis Rheum* 1997;40:371–80.

Development of a Glucocorticoid Toxicity Index (GTI) using multicriteria decision analysis

Eli M Miloslavsky, Ray P Naden, Johannes W J Bijlsma, Paul A Brogan, E Sherwood Brown, Paul Brunetta, Frank Buttgereit, Hyon K Choi, Jean-Francois DiCaire, Jeffrey M Gelfand, Liam G Heaney, Liz Lightstone, Na Lu, Dedee F Murrell, Michelle Petri, James T Rosenbaum, Kenneth S Saag, Murray B Urowitz, Kevin L Winthrop and John H Stone

Ann Rheum Dis published online July 29, 2016