ACR Provisional Composite Index for Scleroderma Clinical Trials

<<rrh>>ACR PROVISIONAL COMPOSITE INDEX FOR SCLERODERMA CLINICAL TRIALS
<<lrh>>KHANNA ET AL
<<st2>>SPECIAL ARTICLE
<<title>>The AQ1 American College of Rheumatology Pprovisional Ceomposite Rresponse Iindex for Celinical Ttrials in Eearly Ddiffuse Ceutaneous Ssystemic Ssclerosis
<<aut>>
Dinesh Khanna, ~~MD, MS~~[1], Veronica J. Berrocal, ~~PhD~~[2]PhD[1], Edward H. Giannini, ~~MSc, DrPH~~[3]DrPH[2], James R. Seibold, ~~MD~~[4]MD[3], Peter A. Merkel, ~~MD, MPH~~[5]MPH[4], Maureen D. Mayes, ~~MD, MPH~~[6]MPH[5], Murray Baron, ~~MD~~[7]MD[6], Philip J. Clements, ~~MD, MPH~~[8]MPH[7], Virginia Steen, ~~MD~~[9]MD[8], Shervin Assassi, ~~MD, MS~~[6]MS[5], Elena Schiopu, ~~MD~~[1], Kristine Phillips, ~~MD, PhD~~[1], Robert W. Simms, ~~MD~~[10]MD[9], Yannick Allanore, ~~MD, PhD~~[10]4], Christopher P. Denton, ~~MD, PhD~~[11]2], Oliver Distler, ~~MD~~[123], Sindhu R. Johnson, ~~MD, PhD~~[134], Marco Matucci-Cerinic, ~~MD, PhD~~[145], Janet E. Pope, ~~MD, MPH, FRCPC~~[156], Susanna M. Proudman, ~~MBBS~~[17]MBBS, FRACP[16], Jeffrey Siegel, ~~MD~~[178], Weng Kee Wong, ~~MS, PhD~~[78], Athol U. Wells, ~~MD~~[189], and Daniel E. Furst, ~~MD~~[78]
AQ2

~~*Not sure what text should go in the boxed disclaimer.*~~

This criteria set has been approved by the American College of Rheumatology (ACR) Board of Directors as Provisional. This signifies that the criteria set has been quantitatively validated using patient data, but it has not undergone validation based on an external data set. All ACR-approved criteria sets are expected to undergo intermittent updates.

The ACR is an independent, professional, medical and scientific society which does not guarantee, warrant, or endorse any commercial product or service.

<<ftnts>>

This article is published simultaneously in the February 2016 issue of *Arthritis Care & Research.*

~~*The American College of Rheumatology is an independent, professional, medical and scientific society which does not guarantee, warrant, or endorse any commercial product or service.*~~

~~BOXED TEXT~~

AQ1 Author: I was the staff member who copyedited your article. If you have any questions or comments about the editing or if you would like to discuss specific aspects of your responses to any of the queries, please write your comments on the proofs or contact me by phone (404-633-3777), fax (404-329-7335), or e-mail (jdiamond@rheumatology.org). *Jane Diamond, Managing Editor*

AQ2 Comp.: The two paragraphs in italics are the boxed text that goes below the author names. See ACR criteria or guidelines articles in Oct. '15 issue or Jan. '16 issue for style.

Formatted: Left, Don't add space between paragraphs of the same style, Tab stops: Not at 0.06"
Formatted: Left: 1.25", Right: 1.25"
Formatted: Font: (Default) Times New Roman, 12 pt, Not Bold
Formatted: Highlight
Formatted: Highlight
Formatted: Font: (Default) Times New Roman, 12 pt, Not Bold
Formatted: Font: (Default) Times New Roman, 12 pt, Not Bold
Formatted: Font: (Default) Times New Roman, 12 pt, Not Bold
Formatted: Font: (Default) Times New Roman, 12 pt, Not Bold
Formatted: Font: (Default) Times New Roman, 12 pt, Not Bold
Formatted: Font: (Default) Times New Roman, 12 pt, Not Bold
Formatted: Font: (Default) Times New Roman, 12 pt, Not Bold
Formatted: Font: (Default) Times New Roman, 12 pt, Not Bold
Formatted: Font: (Default) Times New Roman, 12 pt, Not Bold
Formatted ... [1]
Formatted ... [2]
Formatted ... [3]
Formatted ... [4]
Formatted ... [5]
Formatted: Not Superscript/ Subscript
Formatted: Highlight
Formatted: Highlight
Formatted ... [6]
Formatted ... [7]
Formatted ... [8]
Formatted ... [9]
Formatted ... [10]
Formatted: Font: Univers
Formatted: Font: Univers, Bold
Formatted ... [11]
Formatted ... [12]
Formatted: Font: Univers
Formatted: Font: Univers, Bold

ACR Provisional Composite Index for Scleroderma Clinical Trials

T >><<ftnts>>The contents herein isare solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Research reported in this publication was s Supported by the NIH (National Institute of Arthritis and Musculoskeletal and Skin Diseases of the National Institutes of Health under Award Numbergrant U0O1–AR-055057). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Drs. Khanna and Berrocal's work were alsoas supported by the NIH/NIAMS (National Institute of Arthritis and Musculoskeletal and Skin Diseases grant K24–AR-063120). Dr. Johnson's work wasis supported by the Canadian Institutes of Health Research (Clinician Scientist Award).

[1]Dinesh Khanna, MD, MS1, Veronica J. Berrocal, PhD1, Elena Schiopu, MD1, Kristine Phillips, MD, PhD1,; University of Michigan, Ann Arbor;
[2]Edward H. Giannini, MSc, DrPH2: Cincinnati Children's Hospital, Cincinnati, Ohio,;
[3]James R. Seibold, MD3,: Scleroderma Research Consultants, Litchfield, Connecticut;
[4]Peter A. Merkel, MD, MPH4,: University of Pennsylvania, Philadelphia;
[5]Maureen D. Mayes, MD, MPH5, Shervin Assassi, MD, MS5,: University of Texas Health Science Center at Houston;
[6]Murray Baron, MD6: Jewish General Hospital and McGill University, Montreal, Quebec, Canada;,
[7]Philip J. Clements, MD, MPH7, Weng Kee Wong, MS, PhD7, Daniel E. Furst, MD7: University of California, Los Angeles,;
[8]Virginia Steen, MD8,: Georgetown University, Washington, DC; Rob
[9]Shervin Assassi, MD, MS5, Elena Schiopu, MD1, Kristine Phillips, MD, PhD1,; Robert W. Simms, MD9,: Boston University, Boston, Massachusetts;
[10]Yannick Allanore, MD, PhD10,: Paris Descartes University and Cochin Hospital, AP-HP, Paris, France; [11]Chri
[11]Christopher P. Denton, MD, PhD, FRCP11: Royal Free and University College London Medical School, London, UK;,
[12]Oliver Distler, MD12: University Hospital Zurich, Zurich, Switzerland;,
[13]Sindhu R. Johnson, MD, PhD13: Toronto Western Hospital and University of Toronto, Toronto, Ontario, Canada;,
[14]Marco Matucci-Cerinic, MD, PhD14: Azienda Ospedaliero-Universitaria Careggi (AOUC) and University of Florence, Florence, Italy;, [15]J
[15]Janet E. Pope, MD, MPH, FRCPC15:; Schulich School of Medicine, Western University, London Ceampus, and St. Joseph's Health Care, London, Ontario, Canada;
[16]Susanna M. Proudman, MBBS, FRACP16: Royal Adelaide Hospital and University of Adelaide, Adelaide, South Australia, Australia;,
[17]Jeffrey Siegel, MD17: Genentech/Roche, San Francisco, California;,
[18]Weng Kee Wong, MS, PhD7, Athol U. Wells, MD18:, Royal Brompton Hospital, London, UK and Daniel E. Furst, MD7

Dr. Khanna has received consulting fees from Actelion, Bristol-Myers Squibb, Cytoria, Genentech/Roche, InternMune, Lycera, EMD Serono, and Seattle Genetics (less

Formatted: Space After: 0 pt, Don't add space between paragraphs of the same style, Line spacing: single
Formatted: Font: Not Bold
Formatted: Not Superscript/ Subscript
Formatted: Not Superscript/ Subscript
Formatted: Not Superscript/ Subscript
Formatted: Indent: First line: 0.5", Don't add space between paragraphs of the same style, Line spacing: single
Formatted: Not Superscript/ Subscript
Formatted: Not Superscript/ Subscript
Formatted: Not Superscript/ Subscript
Formatted: Not Superscript/ Subscript
Formatted: Not Superscript/ Subscript
Formatted: Not Superscript/ Subscript
Formatted: Not Superscript/ Subscript
Formatted: Not Superscript/ Subscript
Formatted: Not Superscript/ Subscript
Formatted: Not Superscript/ Subscript
Formatted: Not Superscript/ Subscript
Formatted: Not Superscript/ Subscript
Formatted: Not Superscript/ Subscript
Formatted: Not Superscript/ Subscript
Formatted: Not Superscript/ Subscript
Formatted: Not Superscript/ Subscript
Formatted: Not Superscript/ Subscript
Formatted: Font: (Default) Times New Roman, 12 pt
Formatted: Font: (Default) Times New Roman, 12 pt
Formatted: Font: (Default) Times New Roman, 12 pt
Formatted: Superscript
Formatted: Not Superscript/ Subscript
Formatted: Not Superscript/ Subscript
Formatted: Not Superscript/ Subscript
Formatted: Don't add space between paragraphs of the same style, Line spacing: single

than $10,000 each) and Bayer (more than $10,000) and has received research funding from Actelion, Bristol-Myers Squibb, Cytori, Genentech/Roche, InterMune, Lycera, EMD Serono, Seattle Genetics, Bayer, ~~those companies and/or Bayer,~~ Biogen Idec, Celgene, Forward, Gilead, GlaxoSmithKline, Medac, and Sanofi-Aventis/Genzyme. AQ3

Dr. Giannini has received consulting fees, speaking fees, and/or honoraria from ●●● ● (less than $10,000). AQ4

Dr. Seibold has received consulting fees from ~~consultancies relevant to the present work with~~ Bayer, Boehringer-Ingelheim, Biogen Idec, FibroGen, Novartis, Sanofi-Aventis, Celgene, and DART, (less than $10,000 each) and EMD Serono, InterMune, and Sigma Tau (more than $10,000 each). ~~FibroGen, Novartis, Sanofi-Aventis, Celgene, DART, InterMune and Sigma Tau.~~ AQ5

Dr. Merkel has ~~had consultancy relationships with~~ received consulting fees from Actelion, ChemoCentryx, Glaxo-Smith-Kline, and Sanofi (less than $10,000 each) AQ6 and has received research funding from Actelion, Bristol-Myers Squibb, Celgene, GlaxoSmithKline, and Genentech/-Roche.

Dr. Mayes has received consulting fees, speaking fees, and/or honoraria from Medtelligence and Cytori (less than $10,000 each).

Dr. Steen has received consulting fees ~~/had consultancy relationship with and/or has received research funding~~ from Actelion, Bayer, Bristol-Myers Squibb, Celgene, CSL Behring, Cytori, Genentech/-Roche, Gilead, InterMune, Sanofi-Aventis/Genzyme, and United Therapeutics (less than $10,000 each) and research funding from these companies. AQ7

AQ3 Author: I have tried to combine the information from the author disclosure forms and the manuscript for the various authors, including you, ~~Dr. Khanna~~. In many cases I am not sure which companies paid consulting fees/other fees (and whether these were in the "less than $10,000 category" or "more than $10,000 category") and which companies provided research funding. And, in some cases, a company may have paid fees to an author and also provided research funding. Please help with sorting this out in my queries. In your disclosures, I'm not sure if any of the companies listed for fees also provided research funding and should be listed there as well. Please advise.

AQ4 Author: Dr. Giannini indicated on his author disclosure form that he received fees and/or honoraria of less than $10,000, but did not list the name of the company or companies ~~from which he received fees and/or honoraria~~. Please provide that information.

AQ5 Author: Dr. Siebold listed some companies on his disclosure form in the "less than $10,000" or "more than $10,000" categories for fees and/or honoraria, but there were several companies listed in the manuscript that are not on his form: Bayer, FibroGen, Novartis, Sanofi-Aventis, Celgene, and DART. I have included them in the less than $10,000 category. Please make any corrections necessary.

AQ6 Author: Regarding Dr. Merkel's disclosure information, please advise if any of the companies from which he received consulting fees should be listed in the "more than $10,000" category.

AQ7 Author: Regarding Dr. Steen's disclosure information, I could not tell which companies paid consulting fees and whether these fees were less than or more than $10,000. Also, which provided research funding? Please advise ~~so I can edit accurately~~. We are not required to disclose the dollar amount for research funding.

Formatted: Font color: Auto
Formatted: Font: Not Bold, Highlight
Formatted: Font: Not Bold, Highlight
Formatted: Font: Not Bold, Highlight
Formatted: Font: Not Bold, Highlight
Formatted: Normal, Line spacing: single
Formatted: Font color: Auto
Formatted: Font color: Auto
Formatted: Font color: Red, Highlight
Formatted: Highlight
Formatted: Font: Not Bold, Highlight
Formatted: Font: Not Bold, Highlight
Formatted: Don't add space between paragraphs of the same style, Line spacing: single
Formatted: Highlight
Formatted: Highlight
Formatted: Font: Not Bold, Not Highlight
Formatted: Font: Univers, Not Superscript/ Subscript
Formatted: Font: Univers
Formatted: Font: Univers, Highlight
Formatted: Not Highlight
Formatted: Font: Univers, Bold, Not Superscript/ Subscript
Formatted: Font: Univers
Formatted: Font: Univers, Highlight
Formatted: Font: Univers, Bold, Not Superscript/ Subscript
Formatted: Font: Univers
Formatted: Font: Not Bold
Formatted: Font: Univers, Bold
Formatted: Font: Univers, Bold, Not Superscript/ Subscript
Formatted: Font: Univers
Formatted: Font: Univers
Formatted: Font: Univers, Bold, Not Superscript/ Subscript
Formatted: Font: Univers
Formatted: Font: Univers

Dr. Simms ~~is on the~~has received speaking fees from Gilead ~~Speakers bureau~~(less than $10,000) and~~, has received~~ consulting fees from ~~has/had consultancies with~~ Actelion~~,~~ and Cytori (less than $10,000 each)~~,~~ ~~Has/had~~and has received ~~has received~~ research ~~grants~~funding ~~support~~ from Actelion, Gilead, Medimmune, and InterMune. AQ8

Dr. Allanore has~~/had consultancy relationship and/or~~ ~~has~~ received consulting fees from ~~research funding with~~ Actelion, Bayer, Behring, Biogen Idec, Bristol-Myers Squibb, Genentech/-Roche, Inventiva, Medac, Pfizer, Sanofi/Genzyme, Servier, and UCB (less than $10,000 each) and research funding from these companies. AQ9

Dr. Denton has~~/had~~ received consulting fees from~~ancy relationship~~ Actelion, Bayer, GlaxoSmithKline, and Roche (less than $10,000 each) and has received~~/or has received~~ research funding from Actelion, Bayer, GlaxoSmithKline, Roche~~Actelion~~, Genentech/ Roche, Pfizer, GlaxoSmithKline, Bristol-Myers ~~M~~Squibb, CSL Behring, Novartis, Sanofi-Aventis, Inventiva, and Biogen-Idec. AQ10

Dr. Distler has~~/had consultancy relationship and/or has~~ received ~~research funding in the area of SSc and related conditions~~consulting fees from Actelion, Pfizer, Ergonex, Bristol -Myers Squibb, Sanofi-Aventis, United BioSource ~~Corporation~~, Genentech/-Roche, Medac, Biovitrium, Boehringer Ingelheim Pharma, Novartis, ~~4 D Science,~~ Active Biotec, Bayer-Schering, Sinoxa~~/~~, Serodapharm, EpiPharm, Biogen Idec, Inventiva, GlaxoSmithKline, and Pharmacyclics (less than $10,000 each) and 4D Science and Bayer (more than $10,000 each) and has received research funding from Actelion, Pfizer, Ergonex, Bristol-Myers Squibb, Sanofi-Aventis, United BioSource, Genentech/Roche, Medac, Biovitrium, Boehringer Ingelheim Pharma, Novartis, Active Biotec, Bayer-Schering, Sinoxa, Serodapharm, EpiPharm, Biogen Idec, Inventiva, GlaxoSmithKline, Pharmacyclics, 4D Science and Bayer;~~-~~
he holds a patent for the use of microRNA-29 in the treatment of systemic sclerosis. AQ11

AQ8 Author: Were the speaking fees and consulting fees that Dr. Simms ~~has~~ received less than $10,000, as edited?

AQ9 Author: With regard to Dr. Allanore's disclosure information, I have the same questions as for Dr. Steen: which companies paid consulting fees and which provided research funding? He did not include Behring on his disclosure form, but it was listed in the manuscript – OK to include here? Dr. Allanore noted on his disclosure form that none of the fees were over $10,000, so I edited accordingly.

AQ10 Author: Dr. Denton listed on his disclosure forms ~~the~~ four companies (Actelion, Bayer, GlaxoSmithKline, and Roche) from which he received fees (all under $10,000). Did he receive fees from any of the other companies listed? Is it correct to list ~~those four companies, plus the others~~all of the companies~~,~~ as having provided research funding? Please make any corrections necessary.

AQ11 Author: Regarding Dr. Distler's disclosure information, I could not tell which companies paid consulting fees. Several companies were listed on his hard copy disclosure form as being in the less than $10,000 category and 4D Science and Bayer were listed as being more than $10,000, but I wasn't sure whether Ergonex, United BioSource, Biovitrium, Novartis, Biogen Idec, and Inventiva paid consulting fees (and if the fees were less than or more than $10,000). Also, which companies provided research funding? Please advise so we~~l~~ can ~~edit~~make any needed corrections ~~accurately~~.

Formatted: Highlight
Formatted: Highlight
Formatted: Font color: Auto
Formatted: Highlight
Formatted: Highlight
Formatted: Highlight
Formatted: Highlight
Formatted: Font: Univers, Bold, Not Superscript/ Subscript
Formatted: Font: Univers
Formatted: Font: Univers
Formatted: Font: Univers, Bold, Not Superscript/ Subscript
Formatted: Font: Univers
Formatted: Font: Univers
Formatted: Font: Univers, Bold, Not Superscript/ Subscript
Formatted: Font: Univers
Formatted: Font: Univers
Formatted: Font: Univers, Bold, Not Superscript/ Subscript
Formatted: Font: Univers
Formatted: Font: Univers

~~Dr. Johnson is supported by the Canadian Institutes of Health Research Clinician Scientist Award.~~

Dr. Matucci Cerinic has~~/had consultancy relationship and/or has~~ received consulting fees from ~~research funding with~~GlaxoSmithKline, Actelion, ~~Bayer, Behring,~~ Bristol-Myers Squibb, and MSD Pfizer~~, and UCB~~ (less than $10,000 each) and has received research funding from GlaxoSmithKline, Actelion, Bayer, Behring, Bristol-Myers Squibb, MSD Pfizer, and UCB. AQ12

Dr. Proudman has received consulting fees from Actelion (more than $10,000) and has received research ~~grants~~funding ~~and consultancies for~~from Actelion, Bayer, and Glaxo Smith-Kline. AQ13

Dr. Siegel ~~is an employee of~~owns stock or stock options in ~~Genentech~~/Roche.

Dr. Furst ~~has/had consultancies with~~has received consulting fees from AbbVie, Actelion, Amgen, Bristol-Myers Squibb, Cytori, Janssen, Gilead, GlaxoSmithKline, ~~NIH,~~ Novartis, Pfizer, Genentech/-Roche, and UCB ~~and~~(less than $10,000 each) ~~does~~and honoraria for CME programs ~~with~~from AbbVie, Actelion, and UCB (less than $10,000 each) -and has ~~/had~~received research ~~grants~~funding ~~with~~from AbbVie, Actelion, Amgen, Bristol-Myers Squibb, Gilead, GlaxoSmithKline, ~~NIH,~~ Novartis, Pfizer, Genentech/ Roche, and UCB ~~and has/had consultancies with AbbVie, Actelion, Amgen, Bristol-Myers Squibb, Cytori, Janssen, Gilead, GSK, NIH, Novartis, Pfizer, Genentech/ Roche, UCB and does CME programs with AbbVie, Actelion, and UCB.~~ AQ14

~~[1]University of Michigan Scleroderma Program, Ann Arbor, MI, USA;~~
~~[2]University of Michigan, School of Public Health, Ann Arbor, MI, USA,~~
~~[3]Cincinnati Children's Hospital, Cincinnati, OH, USA,~~
~~[4]Scleroderma Research Consultants, Litchfield, CT, USA,~~
~~[5]University of Pennsylvania, Philadelphia, PA, USA,~~
~~[6]University of Texas Health Science Center Houston, TX, USA,~~
~~[7]Jewish General Hospital, McGill University, Montreal, Quebec, Canada,~~
~~[8]UCLA, Los Angeles, CA, USA,~~
~~[9]Georgetown University, Washington, DC, USA,~~ ~~[10]Boston University, Boston, MA, USA,~~
~~[11]Paris Descartes University, Cochin Hospital, Paris, France,~~

AQ12 Author: Dr. Matucci-Cerinic listed on his disclosure form the four companies (GlaxoSmithKline, Actelion, Bristol-Myers Squibb, and MSD Pfizer) from which he received fees (all under $10,000). Is it correct to list those four companies, plus the others, as having provided research funding? Please make any corrections necessary.

AQ13 Author: Is Dr. Proudmann's disclosure information correct as edited – that she received research grants from Actelion as well as consulting fees?

AQ14 Author: Regarding Dr. Furst's disclosures, please indicate if there is any company from which the fees or honoraria received were more than $10,000, and I will make those changes accordingly.

Formatted: Highlight
Formatted: Highlight
Formatted: Highlight
Formatted: Highlight
Formatted: Highlight
Formatted: Highlight
Formatted: Not Superscript/ Subscript
Formatted: Font: Univers, Bold, Not Superscript/ Subscript
Formatted: Font: Univers
Formatted: Font: Univers
Formatted: Font: Univers, Bold, Not Superscript/ Subscript
Formatted: Font: Univers
Formatted: Font: Univers
Formatted: Font: Univers, Bold, Not Superscript/ Subscript
Formatted: Font: Univers
Formatted: Font: Univers

[12]Centre for Rheumatology, Royal Free and University College London Medical School, London, UK,
[13]Division of Rheumatology, University Hospital Zurich, Zurich, Switzerland,
[14]Toronto Scleroderma Program, Toronto Western Hospital, University of Toronto, Toronto, Ontario, Canada,
[15]Department of Experimental & Clinical Medicine, Division of Rheumatology AOUC, University of Florence, Firenze, Italy,
[16]Schulich School of Medicine, Western University, London and St Joseph's Health Care, London, ON,
[17]Royal Adelaide Hospital, North Terrace and Discipline of Medicine, University of Adelaide, , Adelaide, SA, Australia,
[18]Genentech/Roche, San Francisco, CA, USA,
[19]Royal Brompton Hospital, London, UK

Corresponding author
Dinesh Khanna, MD, MSc
Professor of Medicine
Director, University of Michigan Scleroderma Program
Division of Rheumatology/Dept. of Internal Medicine Suite 7C27
300 North Ingalls Street, SPC 5422
Ann Arbor, MI 48109
Email: khannad@med.umich.edu
Phone: 734.647.8173
Fax: 734.763.5761

CONFLICTS

Dr. Khanna has/had consultancy relationship with and/or has received research funding from Bayer, Biogen Idec, Bristol-Myers Squibb, Celgene, Cytori, EMD Serono, Forward, Genentech/ Roche, Gilead, Glaxo SmithKline, Lycera, Medac, Sanofi-Aventis/Genzyme, and Seattle Genetics.
Drs. Assassi, Baron, Berrocal, Clements, Giannini, Mayes, Schiopu, Phillips, Pope, Wong and Wells have no conflicts related to this project.

Dr. Seibold has consultancies relevant to the present work with Bayer, Boehringer-Ingelheim, EMD Serono, FibroGen, Novartis, Sanofi-Aventis, Celgene, DART, InterMune and Sigma Tau.

Dr. Merkel has/had consultancy relationships with Actelion, ChemoCentryx, Glaxo-Smith-Kline, and Sanofi and has received research funding from Actelion, Bristol-Myers Squibb, Celgene, GlaxoSmithKline, and Genentech/ Roche.

Dr. Steen has/had consultancy relationship with and/or has received research funding from Actelion, Bayer, Bristol-Myers Squibb, Celgene, CSL Behring, Cytori, Genentech/ Roche, Gilead, InterMune, Sanofi-Aventis/Genzyme, and United Therapeutics.

Formatted: Not Superscript/ Subscript

Formatted: Don't add space between paragraphs of the same style

~~Dr. Simms is on the Gilead Speakers bureau, has/had consultancies with Actelion, Cytori. Has/had grant support from Actelion, Gilead, Medimmune, and InterMune.~~

~~Dr. Allanore~~ Allanore ~~has/had consultancy relationship and/or has received research funding with Actelion, Bayer, Behring, , Biogen Idec, Genentech/ Roche, Inventiva, Pfizer, Sanofi/Genzyme, Servier, and UCB.~~

~~Dr. Denton has/had consultancy relationship and/or has received research funding from Actelion, Genentech/ Roche, Pfizer, GlaxoSmithKline, BMS, CSL Behring, Novartis, Sanofi Aventis, Inventiva, and Biogen Idec.~~

~~Dr. Distler has/had consultancy relationship and/or has received research funding in the area of SSc and related conditions from Actelion, Pfizer, Ergonex, Bristol Myers Squibb, Sanofi Aventis, United BioSource Corporation, Genentech/ Roche, Medac, Biovitrium, Boehringer Ingelheim Pharma, Novartis, 4 D Science, Active Biotec, Bayer Schering, Sinoxa/ Serodapharm, EpiPharm, Biogen Idec, Inventiva, GSK and Pharmacyclics.~~

~~Dr. Johnson is supported by the Canadian Institutes of Health Research Clinician Scientist Award.~~

~~Dr. Matucci Cerinic has/had consultancy relationship and/or has received research funding with Actelion, Bayer, Behring, Bristol Myers Squibb, MSD Pfizer, and UCB.~~

~~Dr. Proudman has research grants and consultancies for Actelion, Bayer and Glaxo Smith Kline.~~

~~Dr. Siegel is an employee of Genentech.~~

~~Dr. Furst has/had research grants with AbbVie, Actelion, Amgen, Bristol Myers Squibb, Gilead, GSK, NIH, Novartis, Pfizer, Genentech/ Roche, UCB and has/had consultancies with AbbVie, Actelion, Amgen, Bristol Myers Squibb, Cytori, Janssen, Gilead, GSK, NIH, Novartis, Pfizer, Genentech/ Roche, UCB and does CME programs with AbbVie, Actelion, and UCB.~~

~~**Corresponding author**~~
Address correspondence to Dinesh Khanna, MD, MSc,
~~Professor of Medicine~~
~~Director,~~ University of Michigan, Scleroderma Program,
Division of Rheumatology~~/~~, Department~~:~~ of Internal Medicine, Suite 7C27,~~-~~
300 North Ingalls Street, SPC 5422,
Ann Arbor, MI ~~-~~48109,
E-mail: khannad@med.umich.edu.
Submitted for publication ~~xxx~~November 11, 2014; accepted in revised form ~~xxx~~October 30, 2015.<</ftnts>>
~~Phone: 734.647.8173~~
~~Fax: 734.763.5761~~

Formatted: Don't add space between paragraphs of the same style, Line spacing: single

Field Code Changed

Formatted: No underline, Font color: Auto

Formatted: Don't add space between paragraphs of the same style

~~Abstract~~<<abs>>

~~*Introduction:*~~ *Objective.* Early diffuse cutaneous systemic sclerosis (dcSSc) is characterized by rapid changes in the~~of~~ skin and internal organs. ~~Our~~ The objective of this study was to develop a composite response index in [AQ15]dcSSc (~~abbreviated~~ CRISS) for use in randomized controlled trials (RCTs).

Methods. ~~:~~ We developed 150 paper patient profiles with standardized clinical outcome elements (core set items) using patients with dcSSc. Forty scleroderma experts rated 20 patient profiles each and assessed whether each patient had improved or not improved over a period of 1 year. Using the profiles ~~where~~ for which raters had reached a consensus on whether the patients were improved ~~versus~~ not improved (79% of the profiles examined), we fit logistic regression models ~~where~~ in which the binary outcome referred to whether the patient was improved or not, and the changes in the core set items from baseline to follow-up were entered as covariates. We tested the final index in a previously completed RCT.

Results. ~~:~~ Sixteen of 31 core items were included in the patient profiles after a consensus meeting and review of test characteristics of patient-level data. The logistic regression model ~~that~~in which the ~~included the following~~ core set items were~~:~~ change over 1 year ~~s~~ in the modified Rodnan skin thickness score, the forced vital capacity ~~(FVC)% predicted~~, the patient and physician global assessments, and the Health~~y~~ Assessment Questionnaire disability index~~HAQ-DI over 1 year~~ had a sensitivity of 0.982 (95% confidence interval~~CI:~~ [AQ16]0.98~~12~~-0.983)~~,~~ and a specificity of 0.931 (95% ~~CI:~~confidence interval 0.930-0.932)~~,~~ and had the highest face validity. Subjects with a significant decline in renal or cardiopulmonary involvement were classified as not improved, regardless of improvements in other core items. With use of ~~T~~the index, ~~was able to differentiate~~ the effect of methotrexate could be differentiated ~~from~~ the effect of placebo in a 1-year RCT [AQ17](*P* ~~=<~~ 0.0~~5~~2).

Conclusion. ~~:~~ We have developed a CRISS that is appropriate for use as an outcome assessment in RCTs of early dcSSc.

<</abs>>

Formatted: Don't add space between paragraphs of the same style, Line spacing: single

Formatted: Font: Italic

Formatted ... [13]

Formatted: Indent: First line: 0.5", Don't add space between paragraphs of the same style, Line spacing: single

Formatted ... [14]

Formatted ... [15]

Formatted ... [16]

Formatted: Space After: 0 pt, Don't add space between paragraphs of the same style, Line spacing: single

---

AQ15 Author: In the ~~first~~second paragraph of the text you defined CRISS as "composite response index in SSc" whereas here it says "composite response index in dcSSc."~~re~~ Which should we be using? (I have provisionally added "dc" in the second paragraph of the text.)

Formatted ... [17]

AQ16 Author: I changed the lower end of this 95% CI from 0.981 to 0.982 (text says 0.9816).

Formatted ... [18]

AQ17 Author: I changed this P value from <0.05 to =0.02, which is what it says in the text. OK?

Formatted ... [19]

<<hd2>>Background

Systemic sclerosis (SSc; scleroderma, SSc) is one of the most life-threatening rheumatic diseases (1,-2), and is associated with substantial morbidity and many detrimental effects on health-related quality of life (3). In recent years, progress has been made in the development and validation of outcome measures and refinement of trial methodology in SSc (4--7). These advances were paralleled by an increased understanding of the pathogenesis of SSc (8) and development of potential targeted therapies (9). The mModified Rodnan Sskin thickness Sscore (10), a measure of skin thickness (6), has been used as the primary outcome measure in clinical trials of diffuse cutaneous SSc (dcSSc). However, the complexity and heterogeneity of the disease mandate a composite response measure that captures multiple organ involvement and patient-reported outcomes.

An accepted, validated, composite response index in dcSSc could substantially facilitate drug development and clinical research. Compared to individual outcome measures, a composite index has the potential to be more responsive to change (101--123), improve assessment of therapeutic interventions, and facilitate the comparison of responses across trials. Regulatory and funding agencies would then have greater confidence in proposals for interventions. We therefore undertook the present studywork

Our objective was to develop a composite response index in dcSScsystemic sclerosis (abbreviated CRISS) for use in clinical trials.

Formatted: Left, Space After: 0 pt, Don't add space between paragraphs of the same style, Line spacing: single

Formatted: Indent: First line: 0.5", Space After: 0 pt, Don't add space between paragraphs of the same style, Line spacing: single

Formatted: Indent: First line: 0.5", Don't add space between paragraphs of the same style, Line spacing: single

Formatted: Don't add space between paragraphs of the same style, Line spacing: single

Formatted: Don't add space between paragraphs of the same style

<<hd2>>Patients and mMethods

<<extp>>The index was developed using well-accepted expert consensus (134) and data-driven approaches (Figure 1),<<F1>> including the American College of Rheumatology (ACR) standards for the development of response criteria (145). Details are included in the Supplementary materialPatients and Methods, on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.?????/abstract. The basic process was as follows: i1) We conducted a consensus exercise to select domains and outcome measures (core items) for potential inclusion in the composite response index. ii2) We then tested the psychometric properties of the core items in a longitudinal cohort of patients followed up over for 1 year to assess the items' feasibility, reliability, validity, and sensitivity to change. 3iii) We developed a set of 150 patient profiles based on the data generated from the cohort study (and using the core items). Forty scleroderma experts were invited to classify each patient profile as improved or not improved. 4iv) We performed statistical reduction of the data to athe minimum number of domains and core items that, which retained the maximally responsive index and was acceptable to the experts (face validity). 5 v) We then tested the ability of the composite response index to discriminate among therapies using results from a previously published randomized controlled trial (RCT). Each of these steps isThe following paragraphs described each step in greater detail below.

(i). Structured consensus exercise to develop domains and core items. We conducted a structured, 3-round Delphi exercise to reach consensus on core items for clinical trials of SSc; the details of the excerciseexercise which have been published elsewhere (5). Briefly, an initial list of potential domains and items was composed by a steering committee and then the members of the Scleroderma Clinical Trials Consortium (SCTC). In Round 1 asked the SCTC members were asked to list items in 11 pre-defined domains domains, and in Round 2 asked respondents were asked to rate the importance of the chosen items on a 1–9 ordinal scale. This was followed by a face-to-face meeting where, withunder expert facilitators, consensus was reached using the nominal group technique (13) about the which domains and core items to test in a database (5) was reached, using the nominal group technique (14). During this exercise, the Ssteering Ccommittee discussed the feasibility, reliability, redundancy, and validity of the items.

(ii). Data collection and evaluation of psychometric properties in a longitudinal observational cohort. Due to a lack of dcSSc trials with positive findingspositive trials in dcSSc AQ18 and as a consequence of the fact that previous trials did not include some of the core items chosen in the consensus exercise (156), we launched assembled a longitudinal observational cohort (the CRISS Cohort) of patients with early dcSSc (<5 years from first 1st non-Raynaud's phenomenon sign or symptom) at 4 US sScleroderma Ccenters (the CRISS cohort) (167). The observational cohort, recruited over a 12-month period1 year, included 200 patients with dcSSc, defined as skin thickening proximal, as well as distal, to the elbows or knees, with or without involvement of the face and neck. Patients were followed up for 12 months, and

AQ18 Author: I changed "positive trials in dcSSc" to "dcSSc trials with positive findings." If this is not what was meant, please clarify.

Formatted: Left, Don't add space between paragraphs of the same style, Line spacing: single
Formatted: Space After: 0 pt, Line spacing: single
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Formatted: No underline, Font color: Auto
Field Code Changed
Formatted: Indent: First line: 0.5", Don't add space between paragraphs of the same style, Line spacing: single
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Formatted: Don't add space between paragraphs of the same style, Line spacing: single
Formatted: Indent: First line: 0.5", Space After: 0 pt, Don't add space between paragraphs of the same style, Line spacing: single
Formatted: Highlight
Formatted: Highlight
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Formatted: Font: Univers, Bold, Not Superscript/ Subscript
Formatted: Font: Univers
Formatted: Font: Univers, Bold

ACR Provisional Composite Index for Scleroderma Clinical Trials

AQ19 ~~outcome~~features were ~~collected~~ recorded at baseline and 12 months. Exclusion criteria included life expectancy of ~~less than~~ ≤1 year and non-proficiency in English.

AQ20 ~~.~~All core items that emerged from the consensus meeting were included to enable an assessment of their psychometric properties (e.g., feasibility, reliability, and face, content, and construct validity [including sensitivity to change]) (1~~17~~8). Feasibility was defined as completion of the core set item by >~~-~~50% of subjects at ~~two~~ 2 time points, and AQ21 redundancy was defined as either a Spearman or Pearson correlation coefficient of at least 0.80 at baseline or during follow~~-~~up. Sensitivity to change ~~was calculated~~ over the 1-year period was calculated using appropriate patient and physician anchor and transition questions. For example, a modified Likert scale (transition health question) was ~~employed~~ used by physicians and patients at the 1-year follow~~-~~up visit to determine the change in overall condition during the prior year on a scale ~~from~~ of 1 ("much better") to 5 ("much worse"). Responses of 1 or 2 were considered an improvement in health, ratings of 4 or 5 were considered a decline in health, and a rating of 3 was considered to mean that there was no appreciable change in overall health. For this analysis, AQ22 those who answered "1" or "2" were categorized as ~~"~~improved~~"~~ on both transition questions and those who scored "3," ~~;~~ "4," or "5" were categorized as ~~"~~not improved~~"~~. Effect size ~~(ES)~~ was calculated using the transition questions as anchors and Cohen's "rule-of-thumb" for interpreting ~~ES~~effect size: values of 0.20-~~-~~0.49 represent a small change, values ~~between~~ of 0.50~~-~~-0.79 a medium change, and values of ≥0.80 a large change (1~~8~~9). Core items that were significant at predefined *P* value of ~~p~~<~~-~~0.20 (for dichotomous measures) or that had an effect size of ≥~~-~~0.20 in the "~~I~~improved" group (with respect to either patient or physician assessments) were included in the next stage.

Eight ~~Steering Committee~~steering committee members (~~see Acknowledgement section~~DK, JRS, PAM, MDM, MB, PJC, VS, and DEF) reviewed the data and scored each core item on an ordinal scale (from 1~~-~~ to 4) for feasibility, reliability, and face, content, and construct validity (~~[~~including sensitivity to change~~)]~~ using the modified content validity index matrix (~~19~~20).~~:~~ ~~a~~A score of 4 ~~(highest score)~~ was assigned when the item referred to a value or an attribute that is well ~~-~~established in the literature or through systematically obtained information~~;~~, a score of 3 indicated a value or an attribute that is somewhat known and accepted~~,~~ but ~~that~~ may need minor alteration or modification~~;~~, a score of 2 indicated that the rater was unable to assess the attribute without additional information or research~~;~~, and a score of 1 ~~(lowest score)~~ meant that the attribute should definitely not be used as a core item. Experts could also assign "not

---

AQ19, Author: I changed "outcomes" to "features" because when at the baseline time point, I don't think it would be considered an outcome. Or, we could change it to "outcome measures of interest" instead, if you prefer.

AQ20, Author: Paragraph was extremely long, and would be "dense" and difficult to read when typeset. Break here OK?

AQ21, Author: "redundancy" correct? Or should it say "reliability"?

AQ22, Author: Please clarify what is meant by "those who answered '1' or '2' on both transition questions." Was there only one transition question answered by the physician and one answered by the patient? Or did the physician and the patient each answer two different transition questions? And what if there was, for example, and answer of "1" for one transition question and an answer for "4" for the other? In those cases, the patient was not categorized as either improved or not improved?

Formatted: Highlight
Formatted: Highlight
Formatted: Highlight
Formatted: Highlight
Field Code Changed
Field Code Changed
Field Code Changed
Formatted: Highlight
Formatted: Highlight
Formatted: Highlight
Formatted: Highlight
Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed
Formatted: Font: Univers, Bold, Not Superscript/ Subscript
Formatted: Font: Univers
Formatted: Font: Univers, Bold
Formatted: Font: Univers, Bold, Not Superscript/ Subscript
Formatted: Font: Univers
Formatted: Font: Univers, Bold
Formatted: Font: Univers, Bold, Not Superscript/ Subscript
Formatted: Font: Univers
Formatted: Font: Univers, Bold
Formatted: Font: Univers, Bold, Not Superscript/ Subscript
Formatted: Font: Univers
Formatted: Underline
Formatted: No underline
Formatted: Font: Univers, Bold

applicable" if they were unfamiliar with an item or with different aspects of feasibility, reliability, and validity for the item. ~~Items s~~Scores of~~d as~~ 3 or 4 were considered supportive of an individual item.

Based on results from psychometrics analysis and expert input, a modified nominal group technique exercise was ~~conducted~~ led by one of the authors (~~by~~ EHG) via webinar, in which ~~by E. Giannini where~~ consensus was defined a priori as ≥75% agreement on each item of the matrix and overall inclusion/~~-~~exclusion of the item as a core item. During the ~~NGT~~ webinar, summary statistics were provided for each core set item, and the moderator encouraged ~~to~~ discussion of each item by each committee member and then by the group as a whole ~~then as a group~~. This process ensured that all participants had an opportunity to contribute. Subsequently, each item was rescored (if the committee member ~~felt~~ believed the score ~~that it~~ should be changed) and summary statistics were generated. Items that were found to lack feasibility, reliability, and validity (<75% of the raters assigning a score of 3 or better) were excluded from the next step.

~~(iii).~~ Development and ratings of representative patient profiles~~:~~. ~~In this step, w~~We developed 150 paper patient profiles using actual data from the CRISS ~~C~~cohort. To have sufficient data on~~for the~~ representative patients, we also obtained data ~~from~~ on patients with early dcSSc (defined as ~~the~~ a disease duration of <~~-~~5 years) in the Canadian Scleroderma Research Group ~~(CSRG)~~ database (2~~0~~1), a large observational ~~Canadian SSc~~~~scleroderma~~ cohort. Since patient interviews were not performed as part of the consensus meeting (~~S~~step 1)~~i~~), the medical literature was searched to assess the most prevalent/~~-~~bothersome issues faced by patients with SSc (2~~1~~2 ~~-~~-2~~3~~4). Based on this, pain and fatigue (assessed ~~by~~ with the Short Form ~~SF-~~36 vitality scale) (25)~~,~~ were included as part of the patient profiles.

Fifty-four international ~~scleroderma~~ experts in scleroderma clinical care and trial design were subsequently invited to participate in a web-based evaluation of 20 patient profiles each. The profiles were randomly assigned to experts based on their location (North America [~~N=~~n = 29] versus~~.~~ Europe [~~N=~~n = 21] versus~~vs.~~ Australia [~~N=~~n = 4]) and years of experience with management of SSc (>10 years [~~N=~~n = 38] versus~~.~~ ~~<~~-10 years ~~of scleroderma experience~~ [~~N=~~n = 16]), to prevent systematic bias in rating due to practice patterns. For each patient profile, the rater was asked ~~three~~ 3 questions:

1~~.~~) Do you think the patient has improved, stabilized, or worsened (or unable to tell) over 1 year? 2)

2~~.~~) If the patient was rated as improved or worsened, by how much did the patient's condition change~~?~~: considerably, somewhat, or a little? ~~.~~

3~~.~~) How would you rank the ~~three~~ 3 most important core items that influenced your decision regarding change or stability?

Consensus was considered to have been met if at least 75% ~~among~~ of those who rated the same patient profile agreed that the patient had improved, stabilized, or worsened. When there was lack of consensus, ~~the Steering Committee~~s~~S~~teering committee members were asked to rate the profiles that were not assigned to them before, followed by a web-based ~~N~~nominal g~~G~~roup ~~T~~technique exercise to discuss each profile in detail~~.~~ These patient profile ratings were then added to the previous voting, and percentage consensus was recalculated. If the proportion of agreement on a patient profile was then ~~-~~≥~~-~~75%, the case was deemed as having reached consensus. This process ~~produced~~ yielded a final list of 16 core items. Finally, we sought consensus among SSc

Formatted: Indent: First line: 0.5", Don't add space between paragraphs of the same style, Line spacing: single

Formatted: Indent: First line: 0.5", Space After: 0 pt, Don't add space between paragraphs of the same style, Line spacing: single

Field Code Changed

Field Code Changed

Field Code Changed

Field Code Changed

Formatted: Indent: First line: 0.5", Don't add space between paragraphs of the same style, Line spacing: single

Formatted: Font: Times New Roman

Formatted: Font: Not Italic

Formatted: Indent: Left: 0", First line: 0.5", Don't add space between paragraphs of the same style, Line spacing: single

Formatted: Font: Not Italic

Formatted: Font: Not Italic

Formatted: Font: Not Italic

Formatted: Font: Not Italic

Formatted: Font: Not Italic

Formatted: Highlight

Formatted: Indent: First line: 0.5", Don't add space between paragraphs of the same style, Line spacing: single

Formatted: Font: Times New Roman

experts on the level of change in internal organ involvement that ~~would~~ should be used to classify a patient as not improved.

~~(iv).~~ **Development of response definitions.**

Using only profiles ~~where~~ for which consensus was reached, we fit logistic regression models to the binary outcome measure, i.e., whether a patient had been rated by experts as being improved (~~=~~recorded as 1) versus~~.~~ not improved (recorded as ~~=~~0)~~,~~ "Not improved" included scenarios rated as either no change or worsened. ~~.~~We examined various models, increasing at each step the number of predictors (core set items) included in the logistic regression model. For each model, we calculated sensitivity, specificity, and area under the curve (AUC). Additionally, using the estimates of the logistic regression beta coefficients, we derived, for each patient profile, the predicted ~~log~~ log odds, and thus~~,~~ the predicted probability, that the patient would be rated as improved. We then compared the predicted probability to the raters' consensus opinion on the patient. Accuracy of the predictions was evaluated in several ways. Using the predicted probabilities in their continuous form, accuracy in the predictions was quantified ~~by~~ with the Brier score (2246); ~~the~~ the model with the lowest Brier ~~s~~Score is interpreted to have the best predictive performance.

We also tested whether the predicted probabilities had a different distribution for the patient profiles ~~which~~ that were rated improved by the experts and ~~for~~ those that were rated not improved. ~~We assessed~~ T~~t~~he difference in the ~~two~~ 2 distributions was assessed ~~via~~with the non-parametric Mann-Whitney test. We examined whether the predicted probabilities could be transformed into binary classifications by choosing a threshold and defining "improved" for all patients for ~~which~~ whom the predicted probability is above the chosen threshold and "not improved" for all patients for ~~which~~ whom the predicted probability is below the threshold. To identify which threshold (i.e., cut point) to use, we considered different possible cut points from 0.1 to 1.0. For each of the thresholds considered, we derived the corresponding sensitivity and specificity of the predicted binary classification of patients into improved (i.e., ~~(~~=1) or not improved (i.e., ~~=~~0). We ~~made a~~ plotted ~~of the~~ sensitivity and specificity as a function of each threshold and determined which threshold had the highest sensitivity and specificity. The data-driven definitions were discussed with the ~~Steering Committee~~Ssteering committee regarding content and face validity.

To determine whether there was a clear distinction among the 16 core items in the degree of their ability~~ir helpfulness~~ to guide raters in determining whether a patient was improved or not, we conducted a cluster analysis. To evaluate the contribution of each core component to the final CRISS, we computed the generalized coefficient of determination or pseudo R^2 for logistic regression (2257).

~~(v).~~ **Preliminary evaluation in an independent cohort.**

The composite index was tested in an RCT ~~randomized controlled trial~~ of methotrexate versus~~.~~ placebo ~~in~~ for the treatment of early dcSSc (2268). This trial was chosen ~~as~~ because individual patient data were recorded~~,~~ and all final core items were available in this database. We applied the CRISS to the ~~subjects~~ patients with complete data and~~,~~ for each ~~subject~~patient, derived the predicted probability that the~~a subject~~ individual was improved, using the predicted probability equation (see below~~Results section~~). We transformed the continuous predicted probabilities ranging from 0 to 1 into a

Formatted: Don't add space between paragraphs of the same style, Line spacing: single

Formatted: Indent: First line: 0.5", Space After: 0 pt, Don't add space between paragraphs of the same style, Line spacing: single

Field Code Changed

Field Code Changed

Field Code Changed

Field Code Changed

Field Code Changed

Field Code Changed

Field Code Changed

Field Code Changed

Field Code Changed

binary classification, by defining each ~~subject~~ patient as "~~"~~improved~~"~~ or "~~"~~not improved~~"~~ depending on whether the predicted probability was above the threshold with the highest sensitivity and specificity (identified in ~~s~~Step ~~4# iv~~). We then tested whether the probability of being improved was independent of ~~being on~~ methotrexate therapy (~~ei~~.~~ge~~., whether the probability of being improved was the same in the methotrexate-treated and the placebo-treated groups~~two groups of subjects – placebo and methotrexate~~), by ~~performing a~~ chi-square testing. We also assessed, by Mann-Whitney test, whether the distributions of the predicted probabilities differed between~~for~~ the ~~subjects~~ patients who received~~on~~ methotrexate and those who received~~subjects on~~ placebo. ~~were different using the Mann-Whitney test.~~

<<~~/~~extp>>

<<hd2>>**Results**

AQ23 ~~**(i).**~~ **Structured ~~C~~consensus ~~e~~Exercise to develop domains and core items**

. A total of 50 SCTC investigators participated in ~~Roundstep~~round 1, providing 212 unique items for the 11 domains, and rated 177 items in ~~Round step~~round 2. AQ24 ~~The~~ The ratings of the 177 items were reviewed by the ~~Steering Committees~~Steering committee, and 11 domains and 31 items were identified as the core items that met the Outcome Measures in Rheumatology (OMERACT-) filters of truth, feasibility, and discrimination. The 11 domains included~~:~~ skin, musculoskeletal, cardiac, pulmonary, gastrointestinal, renal, Raynaud's phenomenon, digital ulcers, health-related quality of life and function, global health, and biomarkers. Attendees of a 2008 OMERACT- conference (4,29)~~in 2008~~ provided input during the consensus exercise ~~(4, 27)~~.

~~(ii).~~ Data collection and evaluation of psychometric properties in the ~~CRISS cohort, a~~ longitudinal observational cohort (CRISS cohort).

~~CRISS Cohort~~

Two hundred patients with early dcSSc were recruited at baseline. For 150 of these patients, ~~and 150 had~~ both baseline and 1-year data were available. The mean ± SD age of ~~In~~ these 150 patients, ~~mean (SD) age~~ at baseline was 50.4 ~~(~~ ± 11.7~~),~~ years, and 74.7% were female~~,~~. Seventy-eight percent were white and ~~78% were Caucasian and~~ 10.7 % were Hispanic. The mean duration of disease from the time of the first ~~with mean disease duration (dated from 1st~~ non-Raynaud's phenomenon sign or symptom ~~was) of~~ 2.3 ± ~~(~~1.5~~)~~ years, the mean ~~modified Rodnan skin score (~~MRSS~~)~~ ~~of~~was 21.4 ~~(~~ ± 10.1~~)~~ ~~units~~, the mean forced vital capacity (FVC; % predicted~~)~~ ~~of~~ was 82.3 ± ~~% (~~18.5~~)~~, and the mean ~~-~~Health Assessment Questionnaire (HAQ) disability index ~~HAQ~~ (DI) (30) ~~of~~ was 1.0 ± ~~(~~0.8~~;~~ (Table 1).<<T1>>

AQ23 Author: We don't repeat the same main headings in the Methods section and the Results section. Either the heading is a description of a method or it is a description of a result. In this case, all are descriptions of a method. Therefore, please modify the headings throughout the Results section (hopefully a greater modification than just adding "Results of" to the wording that describes the method).

AQ24 Author: I changed "round 1" and "round 2" to "step 1" and "step 2," which is the terminology you used elsewhere.

Formatted: Space After: 0 pt, Don't add space between paragraphs of the same style, Line spacing: single

Formatted: Font: (Default) Times New Roman, 12 pt

Formatted: Don't add space between paragraphs of the same style

Formatted: Left, Don't add space between paragraphs of the same style, Line spacing: single

Formatted: Font: (Default) Times New Roman, 12 pt

Formatted: Highlight

Formatted: Highlight

Formatted: Indent: First line: 0.5", Don't add space between paragraphs of the same style, Line spacing: single

Formatted: Indent: First line: 0.5", Don't add space between paragraphs of the same style, Line spacing: single, Adjust space between Latin and Asian text, Adjust space between Asian text and numbers

Formatted: Highlight

Formatted: Highlight

Formatted: Don't add space between paragraphs of the same style, Line spacing: single

Formatted: Indent: First line: 0.5", Don't add space between paragraphs of the same style, Line spacing: single

Formatted: Font: Univers, Bold, Not Superscript/ Subscript

Formatted: Font: Univers

Formatted: Font: Univers, Bold

Formatted: Font: Univers, Bold, Not Superscript/ Subscript

Formatted: Font: Univers

Formatted: Font: Univers, Bold

Core items that lacked feasibility due to low completion rate (<50%) at 1 year included durometerry (a device to measure the skin hardness (28)[31]), right heart catheterization, Borg AQ25 dyspnea index (32), 6-minute walk test, and Raynaud's Condition Score (3329) (which required daily patient diary records).

When

Using the patient global assessment was used as the metric to classify patients as improved vs.versus not improved, 57% of the patients were rated as "improved" and 43% were rated as "not improved". Using physician global assessment, 58% of patients were rated as "improved" and 42% were rated as "not improved". The Spearman correlation among the definitions was 0.46, supporting use of 2 global transition questions. Using these transition questions, 55 AQ26 items were found to be not responsive to change or occurred in less than <10% of the cohort: tender joint count, presence of renal crisis, estimated glomerular filtration rateGFR, body mass index, presence of digital ulcers, and erythrocyte sedimentation rate. A modified nominal group review was performed, in which wherein consensus was achieved on 16 core items that should be used for the development of paper patients. It was decided to keep retain renal crisis and presence/absence of digital ulcers as core items due to their impact on prognosis in early dcSSc. No redundancy was noted in the core items was noted at baseline and or in the change scores, as assessed by using the correlation coefficients (Supplementary Appendix Tables 1- and 2, on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.?????/abstract).

(iii). Development and ratings of representative patient profiles.

A total of 150 patient profiles were rated by 40 of 54 invited experts (74% completion) (20 profiles rated by each expert; examples shown in the Supplementary Tables 3–5, [http://onlinelibrary.wiley.com/doi/10.1002/art.?????/abstract], upon requesthttp://onlinelibrary.wiley.com/doi/10.1002/art.?????/abstract). Appendix Tables 3-5). The median number of experts that who rated a profile was 6 , and the (range was 4-–13). In response to the instruction, "Please rank the most important core items that influenced your decision regarding change or stability,", experts ranked MRSS as the "most important" 44% of the time, followed by FVC% predicted % predicted (14.5%), patient global assessment (11.0%), physician global assessment (9.1%), and HAQ- DI (8.0%;) Table 2). All other core items were ranked as most influential in the decision making less than <2% of the time.

Initially, consensus was achieved onfor 107 (71.3%) of the patient profiles (71.3%). The Steering CommitteesSteering committee then rescored the remaining 43 profiles as improved, worsened, or stable, and final consensus was achieved in on 118

AQ25 Author: Is "Borg dyspnea index" correct, or should it say "Borg dyspnea score"? Also, I left a space in the reference list for its citation as reference 32, but didn't have the original reference (ofr whichever version was used). Please fill in the citation in the reference list.

AQ26 Author: Based on the list of items in the sentence, it looks like "5 items" should be changed to "6 items." Please advise.

Formatted: Don't add space between paragraphs of the same style, Line spacing: single

Formatted: Indent: First line: 0.5", Don't add space between paragraphs of the same style, Line spacing: single

Formatted: Highlight

Formatted: Highlight

Formatted: Don't add space between paragraphs of the same style, Line spacing: single

Formatted: Highlight

Formatted: Indent: First line: 0.5", Don't add space between paragraphs of the same style, Line spacing: single

Formatted: Highlight

Formatted: Highlight

Formatted: Don't add space between paragraphs of the same style, Line spacing: single

Formatted: Indent: First line: 0.5", Don't add space between paragraphs of the same style, Line spacing: single

Formatted: No underline, Font color: Auto

Formatted: Don't add space between paragraphs of the same style, Line spacing: single

Formatted: Indent: First line: 0.5", Don't add space between paragraphs of the same style, Line spacing: single

Formatted: Font: Univers, Bold, Not Superscript/ Subscript

Formatted: Font: Univers

Formatted: Underline

Formatted: Font: Univers, Bold

Formatted: Font: Univers, Bold, Not Superscript/ Subscript

Formatted: Font: Univers

Formatted: Font: Univers, Bold

~~(78.7%)~~ profiles (78.7%). These profiles were then used for developing the response definitions.

~~(iv).~~ **Development of response definitions.**

Logistic regression models.

The~~re were~~ 118 patient profiles ~~for~~ on which consensus was reached~~; these profiles~~ were used in the statistical models to ~~hat~~ examine~~d~~ response definitions regarding improvement based on change in the 16 core items. In 1-core item models (~~models where~~ in which only ~~one~~1 covariate was included), the AUC ranged from AQ27 0.4~~87~~ (for the model including as the single covariate the change in presence/absence of new digital ulcers) to 0.92 (for the model including as the single covariate the change in MRSS~~;~~) (~~Appendix~~ Supplementary Table 6, http://onlinelibrary.wiley.com/doi/10.1002/art.?????/abstract). In a 2-core item model, change in MRSS and change in FVC % predicted yielded the highest AUC (0.96~~;~~) (Supplementary ~~Appendix~~ Table 7, http://onlinelibrary.wiley.com/doi/10.1002/art.?????/abstract) but was deemed not to have content validity as it did not include either the patient or physician perspective. Different definitions of response and their corresponding AUCs, sensitivity, and specificity were discussed by the ~~Steering Committees~~Steering committee (data available upon request from the corresponding author).

AQ28 ~~-~~The 5-core item model including change in MRSS, FVC % predicted, physician global assessment, patient global assessment, and HAQ -DI was voted as having the greatest face validity (Table 2).<<T2>> The AQ29 clustering algorithm supported 5-core item model with the first cluster contained the following 5 items: MRSS, FVC % predicted, patient global assessment, physician global assessment, and HAQ- DI, and the second cluster included all of the remaining core items (Table 3).<<T3>> This model had a sensitivity of 0.9821 (95% confidence interval [95% CI]~~:~~ [0.9816~~,~~ 0.9827]), a specificity of 0.9310 (95% CI~~:~~ [0.9300~~,~~ 0.9321]), and an AUC of 0.9861. The Brier score was 0.038 (lower score indicates ~~a~~ better predictive performance). As the data were not normally distributed, non-parametric tests were used to assess whether the distributions of the predicted probability of improving were different ~~for~~ between~~the~~ subjects who improved and those who did not (Figure 2A).<<F2>> The distributions of predicted improvement probability were found to differ significantly (*P~~p value~~* < 0.0001~~;~~) ~~Figure 2a)~~. Using depiction of sensitivity ~~vs.~~versus specificity for identifying the improved group ~~vs.~~versus the not improved group, a threshold of 0.6 ~~had~~ was found to have the best combination of specificity and sensitivity

AQ27 Author: Due to rounding, I changed "0.47" to "0.48" (in the supplementary table, it is 0.4764). OK?

AQ28 Author: Paragraph break here OK?

AQ29 Author: "The clustering algorithm supported 5-core item model with the first cluster contained the following 5 items" is unclear. Should it be "The clustering algorithm-supported 5-core item model with the first cluster contained the following 5 items" (dash added to "clustering algorithm supported)? Or "The clustering algorithm supported the 5-core item model, with the first cluster containing the following 5 items" (addition of "the" and of a comma, and "contained" changed to "containing")? Or something else? Please clarify.

Formatted: Font: Italic
Formatted: Indent: First line: 0.5", Space After: 0 pt, Don't add space between paragraphs of the same style, Line spacing: single
Formatted: Font: Italic, No underline
Formatted: No underline
Formatted: Indent: First line: 0.5", Don't add space between paragraphs of the same style, Line spacing: single
Formatted: Highlight
Formatted: Highlight
Formatted: Highlight
Formatted: Highlight
Formatted: Highlight
Formatted: Highlight
Formatted: Font: Italic
Formatted: Font: Univers, Bold, Not Superscript/ Subscript
Formatted: Font: Univers
Formatted: Font: Univers, Bold
Formatted: Font: Univers, Bold, Not Superscript/ Subscript
Formatted: Font: Univers
Formatted: Font: Univers, Bold
Formatted: Font: Univers, Bold, Not Superscript/ Subscript
Formatted: Font: Univers
Formatted: Font: (Default) Arial
Formatted: Font: (Default) Arial, Bold
Formatted: Font: (Default) Arial, Bold
Formatted: Font: (Default) Arial, Bold
Formatted: Font: Univers, Bold

values (Figure 2B~~b~~). The 5~~-~~-core item logistic regression model can be used not only to derive predicted probabilities of improving on a 0~~-~~-1 scale, but also to derive the ~~log-~~log odds of improving for each subject. The latter can take any value: a ~~log~~ log odds of 0 means that an ~~subject~~ individual has equal odds ~~to~~ of improv~~e~~ing as ~~to~~ not improving~~e~~ (i.e., predicted probability of 0.5 or 50%) while a positive (negative) ~~log~~ log odds means that an individual ~~subject~~ has greater (lower) odds of improving.

Contribution of 5 core components to the CRISS.

We computed the pseudo R^2 for the logistic regression models that included all ~~the~~ 5 core items of the CRISS, as well as the pseudo R^2 for logistic regression models including each single predictor. Combined, the 5 core items explained 89.3% of the variability in the data. Individually, when used in a single~~-~~-core item logistic regression model, the MRSS explained 66.3% of the variation, the FVC% predicted explained 36.1%~~ of the variation~~, the physician global assessment explained 24.5%~~ of the variation~~, the patient global assessment explained 23.7%~~ of the variation~~, and the HAQ-DI explained 28.5%~~ of the variation~~.

We assessed

~~To assess~~ how changes in the core items we~~a~~re related to the predicted probabilit~~yies~~ of ~~improvment~~improvement~~ing~~ ~~on~~ for each patient profile,~~.~~ ~~Appendix T~~~~Figure 1(a)-(e) presents a scatterplot of t~~he changes (from baseline to 12 months) in the MRSS, ~~change in~~ FVC% predicted, ~~change in the~~ patient global assessment, ~~change in~~ physician global assessment, and ~~change in~~ HAQ-DI versus the predicted probabilities for the 118 patient profiles~~,~~ are depicted in Supplementary Figure 1, on the *Arthritis & Rheumatology* web site at http://onlinelibrary.wiley.com/doi/10.1002/art.?????/abstract). ~~all calculated from baseline to 12 months. A c~~Changes in the MRSS, FVC~~,% predicted~~ and HAQ-DI ~~a~~were strong indicators of whether a patient wa~~i~~s likely to be improved ~~or not~~. In each scenario, a decrease ~~of~~ in the MRSS or HAQ-DI from baseline to follow~~-~~up and an increase in the FVC% predicted correspond~~s~~ed to very high probabilities of improving. For patient ~~global~~ and physician global assessments, the association between probability of improving and change in these ~~two~~ 2 core components wa~~i~~s less evident.

Defining a patient as~~who is~~ not improved irrespective of improvement in ~~other~~ core items.

The ~~Steering Committee~~Ssteering committee considered circumstances in which a patient may improve in a particular outcome measure (such as MRSS or FVC~~% predicted~~) but have clinically significant worsening or end~~-~~-organ damage to another organ (e.g., development of renal crisis or pulmonary arterial hypertension). There was consensus that in a clinical trial, such patients should be defined as not improved ~~in a clinical trial~~. The ~~Steering Committee~~sSteering committee voted and determined that the following items met this definition: new onset of renal crisis, new-onset or worsening of lung fibrosis, new onset of pulmonary arterial hypertension, or new onset of left ventricular failure (Figure 3~~Table 4~~). ~~<<T4>>~~<<F3>> The international experts subsequently endorsed these definitions as well.

~~(v).~~ Preliminary evaluation in a randomized controlled clinical trial.

Formatted: Font: Italic
Formatted: Don't add space between paragraphs of the same style, Line spacing: single
Formatted: Font: Italic, No underline
Formatted: Indent: First line: 0.5", Don't add space between paragraphs of the same style, Line spacing: single
Formatted: Indent: First line: 0.5", Line spacing: single
Field Code Changed
Formatted: No underline, Font color: Auto
Formatted: Don't add space between paragraphs of the same style, Line spacing: single
Formatted: Font: Italic
Formatted: Font: Italic, No underline
Formatted: Indent: First line: 0.5", Don't add space between paragraphs of the same style
Formatted: Font: Italic, No underline
Formatted: Don't add space between paragraphs of the same style, Line spacing: single
Formatted: Indent: First line: 0.5", Don't add space between paragraphs of the same style, Line spacing: single

We used the individual patient data from a clinical trial ~~comparing~~ that compared treatment of dcSSc with methotrexate ~~vs.~~versus placebo (28) to assess our definition of response ~~(26)~~. Data on~~for~~ change in MRSS, FVC % predicted, patient global assessment, physician global assessment, and HAQ- DI w~~as~~ere available for 35 of 71 patients at 1 year. Using the CRISS, we derived the predicted probability of improving for each of the 35 patients with complete baseline and 1-year data and classified them ~~into~~ as improved ~~and~~ or not improved using a probability cutoff of 0.6 (determin~~eid~~ed analytically in ~~S~~step 4~~#iv~~). With this criterion, 11 of 19 ~~subjects~~ patients who received methotrexate were rated as improved, whereas 3 of 16 ~~subjects~~ patients in the placebo group were rated as improved (P~~p~~ = 0.04~~;~~) (Supplementary Figure 2, http://onlinelibrary.wiley.com/doi/10.1002/art.?????/abstract). ~~Appendix Figure 2).~~ When the data were assessed as a continuous measure, the distribution of the predicted probability for improvement was ~~statistically~~ significantly different between the placebo and ~~the~~ methotrexate groups (P~~p~~ = 0.02).

Application in a clinical trial.

The CRISS was developed with ~~a~~the goal of~~to~~ summarize~~ing~~ ~~the~~ changes in ~~the~~ clinical and patient-reported outcomes in a single composite score that conveys the likelihood (or probability) that ~~the~~ a patient with dcSSc has improved. If there is an effective agent for treatment of dcSSc, the assumption is that ~~the~~ a patient treated with the agent will have a higher probability of improvement as summarized by the CRISS ~~vs.~~versus a patient treated with placebo or an ineffective agent. The CRISS is a 2-step process for use in a clinical trial and is described in Figure 3~~Table 4~~. In ~~Step~~ step 1, ~~subjects~~ patients who develop new onset of renal crisis, new -onset or worsening of lung fibrosis, new onset of pulmonary arterial hypertension, or new onset of left ventricular failure during the trial are considered as not improved and assigned a probability of improving equal to 0.0. For the remaining ~~subjects~~ patients with complete data, s~~S~~tep 2 involves computing the predicted probability of improving for each ~~subject~~individual, using the equation shown in Figure 3~~Table 4~~. Subjects for whom the predicted probability is ~~greater or equal to~~ ≥0.60 are considered improved, while subjects for whom the predicted probability is ~~below~~ ≤0.60 are considered not improved. The 2 groups (study drug ~~vs.~~versus placebo or ~~an~~ active comparator) can then be compared in a 2 ×~~x~~ 2 table using appropriate significance tests. The predicted probabilities obtained using the CRISS can also be assessed as a continuous variable, and the distributions of the probability of improving for patients ~~on~~ receiving study drug ~~vs.~~versus placebo can be compared using non-parametric tests.

AQ30 ~~For trials that incorporate components of CRISS at multiple time points, t~~The CRISS was developed using data ~~at~~ from 12 month~~s~~ of treatment. Therefore, with regard to trials that incorporate components of the CRISS at multiple time points, there is a lack of data to support its performance at earlier time periods. We recommend using 12-month ~~data~~ findings as primary/-secondary outcome measure~~s~~ and using data from other~~s~~ time points, such as baseline to 3, 6, and/or 9 months, as exploratory outcomes. -We recommend capturing the data ~~at~~ during each patient visit, using specific case report forms for organ involvement. We also encourage ~~developing~~ inclusion of an adjudication

AQ30 Author: Paragraph break here OK?

Formatted: Font: Italic

Formatted: Font: Italic

Formatted: Don't add space between paragraphs of the same style, Line spacing: single

Formatted: Indent: First line: 0.5", Space After: 0 pt, Don't add space between paragraphs of the same style, Line spacing: single

Formatted: Indent: First line: 0.5", Don't add space between paragraphs of the same style, Line spacing: single, Widow/Orphan control, Adjust space between Latin and Asian text, Adjust space between Asian text and numbers

Formatted: Highlight

Formatted: Highlight

Formatted: Font: Univers, Bold, Not Superscript/ Subscript

Formatted: Font: Univers

Formatted: Font: Univers, Bold

committee that can help with validating ~~that~~ the occurrence of cardio-pulmonary or ~~—~~ renal involvement ~~occurred~~. AQ31 If case report forms are not developed and included in the trial, ~~then~~ this information~~se~~ should be captured as part of the accounting of adverse events ~~[~~(all of the~~m~~se occurrences should be classified as serious adverse events~~]~~). ~~Specifically, n~~Non-availability of thi~~e~~se data on specific case report forms ~~[~~(i.e., if such ~~no specific case report forms are~~forms were not developed prospectively for use in the trial)~~upfront]~~ should not be taken as missing data as, again, these occurrences should be captured as ~~adverse events/~~ serious adverse events. If there ~~is~~ are missing data for the components of ~~S~~step 2, we recommend considering the reason for missingness and using appropriate statistical methods. Missing data for the 5 components in ~~S~~step 2 should be imputed t~~ll~~hrough m~~M~~onth 12 before calculating the score.

AQ31, Author: The sentence beginning "If case report forms" and the sentence following it had to be edited quite a bit, for clarity. Please check that all remains accurate.

Formatted: Highlight

Formatted: Highlight

Formatted: Don't add space between paragraphs of the same style, Line spacing: single

Formatted: Don't add space between paragraphs of the same style

Formatted: Don't add space between paragraphs of the same style, Line spacing: single

Formatted: Don't add space between paragraphs of the same style

Formatted: Font: (Default) Times New Roman, 12 pt, Not Bold

Formatted: Font: Univers, Bold, Not Superscript/ Subscript

Formatted: Font: Univers

Formatted: Font: Univers, Bold

<<hd2>>Discussion

We have developed a composite response index for trials ~~(CRISS) in~~of early dcSSc ~~(~~the CRISS~~)~~ using well-established consensus and data-driven approaches. The CRISS includes core items that assess change in ~~two~~ 2 common and prominent manifestations of early dcSSc (skin and interstitial lung disease), functional disability (as assessed by the HAQ~~-~~ DI), and patient and physician global assessments. In addition, the CRISS captures clinically meaningful declines in internal organ involvement requiring treatment, that classify the patient as having not improved (regardless of changes in other parameters) during the clinical trial. We subsequently tested the CRISS using data from a clinical trial and, using this ~~showed that the CRISS~~ index, identified~~identified~~ different probabilities of improvement ~~for~~ among methotrexate-treated versus placebo-treated patients with early dcSSc. The findings of this analysis~~, subjects in the placebo and methotrexate groups,~~ suggest~~inged~~ that methotrexate has the potential to improve the overall health ~~health~~ of patients with~~condition in the~~ dcSSc ~~subjects~~ after 1 year of treatment.

Traditionally, trials in early dcSSc have focused on skin or lung involvement (3~~0~~4,-3~~1~~5). The MRSS has been used as the primary outcome measure ~~for~~ in the trials of skin fibrosis (6). ~~MRSS~~ It meets the OMERACT criteria as a fully validated measure of outcome (3~~2~~6), but is also a surrogate ~~of~~for internal organ involvement and mortality in early dcSSc (3~~3~~7,-3~~4~~8). However, clinical trials in dcSSc to date have largely ~~been~~ yielded ~~"~~negative~~"~~ results, and ~~and~~ the MRSS has been questioned as a primary outcome measure when~~re~~ post~~-~~ hoc analysis of "negative" trials has shown stability/improvement in the MRSS over time (1~~3~~5,-3~~6~~9). The CRISS incorporates multisystem involvement in dcSSc and includes the patient perspective and the impact of the disease on functional disability. AQ32 ~~CRISS~~ It was developed with ~~a~~the goal ~~to~~of summarize~~ing~~ ~~the~~ changes in ~~the~~ clinical and patient-reported outcomes in a single composite score that conveys the likelihood (or probability) that the patient has improved. If a ~~treatment~~therapy is~~For an~~ effective ~~treatment~~ for dcSSc, the assumption is that patients treated with the agent will have a higher probability of improvement, as summarized by the CRISS, than patients treated with~~vs.versus~~ placebo or an ineffective agent.

The CRISS is calculated as a 2-step process (Figure 3~~Table 4~~). The first step evaluates clinically significant decline in renal or cardiopulmonary involvement that requires treatment; if this is present, the patient is classified as not improved. The definitions chosen for internal organ involvement were based on published data and expert opinion regarding involvement that is~~that was felt to be~~ clinically significant and would trigger pharmacologic management. ~~–~~The second step assesses remaining patients and calculates the predicted probability of improvement. Here, the ~~Steering Committee~~Ssteering committee discussed different response definitions and decided on the use of~~using~~ a data-driven definition as suggested by the ACR Criteria ~~s~~Subcommittee (14~~37~~). In addition, data-driven definitions of disease activity have been successfully used for regulatory approval in other rheumatic diseases (40~~38~~,~~-39~~41).

AQ32 Author: The sentence that starts "It was developed with the goal" and the sentence that follows it are too redundant with the first two sentences in the section of Results that (currently) has the heading "Application in a clinical trial." Please delete or modify in one of the two places.

Formatted: Left, Don't add space between paragraphs of the same style, Line spacing: single

Formatted: Indent: First line: 0.5", Space After: 0 pt, Don't add space between paragraphs of the same style, Line spacing: single

Formatted: Highlight

Formatted: Highlight

Formatted: Indent: First line: 0.5", Space After: 0 pt, Don't add space between paragraphs of the same style, Line spacing: single, Tab stops: 4.06", Left

Formatted: Font: Univers, Bold, Not Superscript/ Subscript

Formatted: Font: Univers

Formatted: Font: Univers, Bold

The ~~goal~~ purpose of the CRISS is to assess ~~if~~ whether new pharmacologic agents have an impact on overall disease activity/severity. Our hope is that its~~that the~~ use ~~of~~ CRISS in clinical trials of~~n~~ dcSSc will greatly facilitate the interpretation of results and form the basis for drug approvals. Rather than using numerous outcome measures that vary from trial to trial, the core set of items used in the CRISS will produce a single efficacy measure. This process will lessen the ambiguity associated with ~~the~~ presentation of multiple test statistics, some of which may be significant and others not, and facilitate meta-analyses. It will likely also allow a ~~decrease~~ reduction in the number of patients ~~necessary~~ needed for appropriately powered clinical trials, as has been the case ~~for~~ with other composite indices in rheumatoid arthritis. It should ~~also~~ be noted that ~~the~~ use of the CRISS does not preclude the addition of other items in a trial; it simply provides one standardized outcome that can be easily compared and understood across trials. The individual components of the CRISS would each likely be important secondary outcomes to assess in any trial. If the goal of a trial is to focus on a particular organ (e.g., use of vasodilators for underlying digital ulcers), then the CRISS can be used as a secondary measure.

The initial panel of domains (~~N=~~n = 11) and items (~~N=~~n = 31) offered a comprehensive view of the marked heterogeneity of SSc and at first was modeled on the comprehensive structure of the ~~BILAG~~ British Isles Lupus Assessment Group and ~~SLEDAI~~ Systemic Lupus Erythematosus Disease Activity Index measures used in trials of systemic lupus erythematous (~~44~~02,43). However, many items were discarded based on lack of sensitivity to change in our actual data-~~-~~gathering exercise, and others were shown to lack feasibility. As an example, the CRISS does not include items for worsening gastrointestinal disease or digital ulcers, but it is anticipated that patient and physician global assessments will capture these. The data-driven approach used in the development of the CRISS strongly supports the relatively simple and accessible panel of items that was selected.

~~There are o~~Other indices for SSc ~~that~~ have been de~~velop~~scribed ~~in SSc~~. The European Scleroderma Study Group (~~41~~44) has proposed a composite index to assess SSc-related disease activity in routine clinical care, but it has not been validated as an outcome measure in clinical trials. A severity index (4~~2~~5), a measure that encompasses disease activity and damage, has been proposed and can be used in trials to complement the CRISS.

This study has several strengths. It is the first concerted effort by the scleroderma research community to address the lack of a robust composite index for this multisystem disease. We used well-accepted expert consensus and data-driven methodologies and successfully derived the index for use in patients with early dcSSc. The index addresses several domains of illness by capturing single-organ involvement in early dcSSc, patient assessment of overall disease, functional disability, and physician global assessment. We were ~~only~~ able to test the index ~~only~~ in only a single, small RCT in which a substantial number of patients were lost~~that had loss~~ to follow-up; therefore, ~~CRISS therefore requires~~ further validation of the CRISS in a prospective RCT of adequate size is needed.

~~Our~~ The study is also not without limitations. The CRISS was developed for early dcSSc and may not be valid for late dcSSc or limited cutaneous SSc (lcSSc). A similar exercise in late lcSSc might focus on vascular complications such as digital ulcers, calcinosis, or pulmonary arterial hypertension but might not include the MRSS. The

Formatted: Indent: First line: 0.5", Space After: 0 pt, Don't add space between paragraphs of the same style, Line spacing: single

Formatted: Indent: First line: 0.5", Line spacing: single

majority of past and ongoing ~~therapeutic~~ clinical trials are focused on early dcSSc due to dynamic changes in skin and internal organ involvement that may be responsive to pharmacologic intervention. We did not obtain patient input during the development of the index. We acknowledge this limitation and searched the literature for patient input regarding scleroderma (~~21,~~ 22,23); this led to inclusion of fatigue and pain during the development of patient profiles, but neither measure remained in the final core set of items following the nominal group exercises. Nonetheless, ~~two~~ 2 of the constituent core items of the CRISS include patient global assessment and patient-reported functional assessment.

AQ33 We also note that the CRISS should be considered as a preliminary index. Although i~~the index~~ was tested in an RCT, missing data in the~~e~~at trial (>50%) precludes definitive conclusions, and the CRISS may need to be revised as more data ~~becomes available~~ from future trials become available. We had 118 paper patient profiles ~~where~~ for which there was expert consensus, and these profiles were used to develop different response definitions. Although this is standard methodology, ~~this~~ it may be suboptimal for testing 16 core set items. This may also explain the high AUC of AQ34 0.9~~68~~86 for the index.

Last~~ly~~, as our goal was to develop a response index for change, baseline scores are not included in the algorithm. Other indices such as ACR 20% improvement criteria ~~for~~ for rheumatoid arthritis (13) or the ACR 30% improvement criteria for juvenile idiopathic arthritis (46) also ~~employ~~ address only changes in core items, and not baseline values. Although ~~the~~ baseline scores can influence the change~~d~~ scores, randomization should provide a balanced cohort.

In conclusion, we have developed a novel composite index for use in clinical trials in early dcSSc. The index should be considered provisional, and needs to be validated in RCTs of dcSSc.

Formatted: Highlight
Formatted: Highlight
Formatted: Highlight
Formatted: Highlight
Formatted: Indent: First line: 0.5", Space After: 0 pt, Don't add space between paragraphs of the same style, Line spacing: single
Formatted: Highlight
Formatted: Indent: First line: 0.5", Don't add space between paragraphs of the same style, Line spacing: single
Formatted: Don't add space between paragraphs of the same style

AQ33 Author: Paragraph break here OK?
AQ34 Author: Per text in Results and Table 3, I changed "0.968" to "0.986." OK?

Formatted: Font: Univers, Bold, Not Superscript/ Subscript
Formatted: Font: Univers
Formatted: Font: Univers, Bold
Formatted: Font: Univers, Bold, Not Superscript/ Subscript
Formatted: Font: Univers
Formatted: Underline
Formatted: Underline
Formatted: Font: Univers, Bold

~~**ACKNOWLEDGEMENT**~~

~~Research reported in this publication was supported by the National Institute of Arthritis and Musculoskeletal and Skin Diseases of the National Institutes of Health under Award Number UO1 AR055057. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.~~

~~Drs. Khanna and Berrocal were also supported by NIH/NIAMS K24 AR063120. Dr. Johnson is supported by the Canadian Institutes of Health Research Clinician Scientist Award.~~

~~The Steering committee included: Murray Baron, MD, Philip J. Clements, MD, MPH, Daniel E. Furst, MD, Dinesh Khanna, MD, MS, Maureen D. Mayes, MD, MPH, Peter A. Merkel, MD, MPH, James R. Seibold, MD, and Virginia Steen, MD~~

<<hd1>>**ACKNOWLEDGEMENTS**

<<ack>>We thank ~~AQ35 following colleagues for participating in rating of patient profiles~~Drs. AQ36 Jerome Avouac, ~~Jerome;~~Patricia Carreira, ~~Patricia;~~Lorinda Chung, ~~Lorinda;~~Mary Ellen Csuka, ~~Mary Ellen;~~Laszlo Czirjak, ~~Laszlo;~~ Tracy Frech, ~~Tracy;~~Ariane Herrick, ~~Ariane;~~Monique Hinchcliff, ~~Monique;~~Vivian Hsu, ~~Vivian;~~Murat Inanc, ~~Murat;~~Sergio Jimenez, ~~Sergio;~~Bashar Kahaleh, Otylia~~Bashar;~~ Kowal-Bielecka, ~~Otylia;~~Thomas A. Medsger Jr., ~~Thomas A;~~Ulf Müller-Ladner, ~~Ulf;~~Mandana Nikpour, ~~Mandana;~~ Ami Shah, ~~Ami;~~Wendy Stevens, ~~Wendy;~~ Gabriele Valentini, ~~Gabriele ;~~ Jacob M. van Laar, ~~Jacob M;~~John Varga, ~~John;~~ Madelon Vonk, ~~Madelon;~~and Ulrich A. Walker for participating in rating of patient profiles., ~~Ulrich A~~ <<ack>>

<9-point TRB, All Caps, Centered>>**AUTHOR CONTRIBUTIONS**

<<8/9-point TR; indent first sentence only>>

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Dr. Khanna had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Khanna, Berrocal, Giannini, Seibold, Merkel, Clements, Phillips, Simms, Denton, Johnson, Matucci-Cerinic, Pope, Siegel, Wong,

Acquisition of data. Khanna, Seibold, Merkel, Mayes, Baron, Clements, Steen, Assassi, Schiopu, Phillips, Simms, Denton, Johnson, Matucci-Cerinic, Pope, Siegel,

Analysis and interpretation of data. Khanna, Berrocal, Giannini, Seibold, Merkel, Baron, Clements, Steen, Assassi, Phillips, Simms, Allanore, Denton, Distler, Johnson, Matucci-Cerinic, Pope, Proudman, Siegel, Wong, Wells, Furst,

<<9-point TRB, All Caps, Centered>>**ADDITIONAL DISCLOSURES**

AQ35 Author: Names of steering committee members were deleted from Acknowledgments per style, because they are all authors. The are identified (by initials, per style) in the Patients and Methods section of the text.

AQ36 Author: Do all of the individuals listed have the title "Dr." as edited?

Formatted: Left, Don't add space between paragraphs of the same style, Line spacing: single

Formatted: Space After: 0 pt, Don't add space between paragraphs of the same style

Formatted: Don't add space between paragraphs of the same style

Formatted: Not Superscript/ Subscript

Formatted: Don't add space between paragraphs of the same style

Formatted: Indent: First line: 0.5", Don't add space between paragraphs of the same style, Line spacing: single

Formatted: Highlight

Formatted: Highlight

Formatted: Highlight

Formatted: Highlight

Formatted: Indent: First line: 0.5", Don't add space between paragraphs of the same style

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Indent: First line: 0.5", Don't add space between paragraphs of the same style

Formatted: Font: (Default) Times New Roman, 12 pt

Formatted: Font: Univers, Bold, Not Superscript/ Subscript

Formatted: Font: Univers

Formatted: Font: Univers, Bold

Formatted: Font: Univers, Bold, Not Superscript/ Subscript

Formatted: Font: Univers

Formatted: Font: Univers, Bold

——<<8/9-point TR; indent>>Author Siegel is an employee of Genentech/Roche.

Formatted: Font: Bold

Formatted: Don't add space between paragraphs of the same style, Line spacing: single

Formatted: Don't add space between paragraphs of the same style

<<hd1>>**REFERENCES**

<<ref21>>

AQ37 1. Ioannidis JP, Vlachoyiannopoulos PG, Haidich AB, Medsger TA, Jr., Lucas M, Michet CJ, et al. Mortality in systemic sclerosis: an international meta-analysis of individual patient data. Am J Med 2005;118(1):2--10.

2. Elhai M, Meune C, Avouac J, Kahan A, Allanore Y. Trends in mortality in patients with systemic sclerosis over 40 years: a systematic review and meta-analysis of cohort studies. Rheumatology (Oxford) 2012;51(6):1017--26.

3. Khanna D, Kowal-Bielecka O, Khanna PP, Lapinska A, Asch SM, Wenger N, et al. Quality indicator set for systemic sclerosis. Clin Exp Rheumatol 2011;29(2 Suppl 65):33--9.

4. Khanna D, Distler O, Avouac J, Behrens F, Clements PJ, Denton C, et al, for the Investigators in CRISS and EPOSS. Measures of response in clinical trials of systemic sclerosis: the Ceombined Rresponse Iindex for Ssystemic Ssclerosis (CRISS) and Outcome Measures in Pulmonary Arterial Hypertension related to Systemic Sclerosis (EPOSS). J Rheumatol 2009;36(10):2356--61.

5. Khanna D, Lovell DJ, Giannini E, Clements PJ, Merkel PA, Seibold JR, et al. Development of a provisional core set of response measures for clinical trials of systemic sclerosis. Ann Rheum Dis 2008;67(5):703--9.

6. Khanna D, Merkel PA. Outcome measures in systemic sclerosis: an update on instruments and current research. Curr Rheumatol Rep 2007;9(2):151--7.

7. Chung L, Denton CP, Distler O, Furst DE, Khanna D, Merkel PA. Clinical trial design in scleroderma: where are we and where do we go next? Clin Exp Rheumatol 2012;30(2 Suppl 71):97--102.

8. Abraham DJ, Varga J. Scleroderma: from cell and molecular mechanisms to disease models. Trends Immunol 2005;26(11):587--95.

9. Nagaraja V, Denton CP, Khanna D. Old medications and new targeted therapies in systemic sclerosis. Rheumatology (Oxford) 2015;54:1944--534.

10. Clements P, Lachenbruch P, Seibold J, White B, Weiner S, Martin R, et al. Inter and intraobserver variability of total skin thickness score (modified Rodnan TSS) in systemic sclerosis. J Rheumatol 1995;22:1281--5.

110. v Van der Heijde DM, van 't Hof MA, van Riel PL, Theunisse LA, Lubberts EW, van Leeuwen MA, et al. Judging disease activity in clinical practice in rheumatoid arthritis: first step in the development of a disease activity score. Ann Rheum Dis 1990;49(11):916--20.

121. Paulus HE, Egger MJ, Ward JR, Williams HJ, and the Cooperative Systematic Studies of Rheumatic Diseases Group. Analysis of improvement in individual rheumatoid

AQ37 Author: Starting with reference 10, the references have been renumbered: Reference citations were inserted for the MRSS (ref. 10), the SF-36 (ref. 25), the HAQ (ref. 30), and the ACR Pedi 30 (ref. 46) (and a space was left for the citation of the Borg index as noted in the query above). Please confirm that all the references I inserted are the correct ones, especially for the HAQ as I cited the original HAQ but I believe there is also a specific one for scleroderma. The reference for the BILAG and SLEDAI was replaced with the original citations for those (refs. 42 and 43). Also, reference 35 was a duplicate of reference 15 and reference 37 was a duplicate of reference 14.

Formatted: Left, Indent: Left: 0", Don't add space between paragraphs of the same style, Line spacing: single

Formatted: Indent: Left: 0", Don't add space between paragraphs of the same style

Formatted: Highlight

Formatted: Indent: Left: 0", First line: 0", Don't add space between paragraphs of the same style

Formatted: Highlight

Formatted: Indent: Left: 0", First line: 0", Don't add space between paragraphs of the same style, Pattern: Clear (Custom Color(RGB(255,253,233)))

Formatted: Font: (Default) Times New Roman, 12 pt, Not Bold, Font color: Auto

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: (Default) Times New Roman, , Check spelling and grammar

Formatted: Indent: Left: 0", First line: 0", Don't add space between paragraphs of the same style

Formatted: Font: Times New Roman

Formatted: Indent: Left: 0", First line: 0", Don't add space between paragraphs of the same style

Formatted: Font: Univers, Bold, Not Superscript/ Subscript

Formatted: Font: Univers

Formatted: Font: Univers, Bold

arthritis patients treated with disease-modifying antirheumatic drugs, based on the findings in patients treated with placebo. ~~The Cooperative Systematic Studies of Rheumatic Diseases Group.~~ Arthritis Rheum 1990;33~~(4)~~:477-~~-~~84.
1~~3~~2.~~—~~ Felson DT, Anderson JJ, Boers M, Bombardier C, Furst D, Goldsmith C, et al. American College of Rheumatology~~.~~ p~~P~~reliminary definition of improvement in rheumatoid arthritis. Arthritis Rheum 1995;38~~(6)~~:727-~~-~~35.
1~~3~~4.~~—~~ Nair R, Aggarwal R, Khanna D. Methods of formal consensus in classification/diagnostic criteria and guideline development. Semin Arthritis Rheum 2011;41~~(2)~~:95-~~-~~105.
1~~4~~5.~~—~~ Classification and Response Criteria Subcommittee of the American College of Rheumatology Committee on Quality Measures~~Singh JA, Solomon DH, Dougados M, Felson D, Hawker G, Katz P, et al~~. Development of classification and response criteria for rheumatic diseases. Arthritis Rheum 2006;55~~(3)~~:348-~~-~~52.
1~~5~~6.~~—~~ Merkel PA, Silliman NP, Clements PJ, Denton CP, Furst DE, Mayes MD, et al, for the Scleroderma Clinical Trials Consortium. Patterns and predictors of change in outcome measures in clinical trials in scleroderma: an individual patient meta-analysis of 629 subjects with diffuse cutaneous systemic sclerosis. Arthritis Rheum 2012;64:3420--9. ~~Merkel PA, Silliman NP, Clements P, Denton CP, Furst DE, Mayes M, et al. Patterns and pPredictors of cChange in oOutcome mMeasures in cClinical tTrials in sScleroderma. Arthritis Rheum 2005;52:282--3.~~ ~~AQ38~~
1~~6~~7.~~—~~ Wiese AB, Berrocal VJ, Furst DE, Seibold JR, Merkel PA, Mayes MD, et al. Correlates and responsiveness to change of measures of skin and musculoskeletal disease in early diffuse systemic sclerosis. Arthritis Care Res (Hoboken) 2014;66:1731-~~-~~9.
1~~7~~8.~~—~~ Hays RD, Hadorn D. Responsiveness to change: an aspect of validity, not a separate dimension. Qual Life Res 1992;1~~(1)~~:73-~~-~~5.
1~~8~~9.~~—~~ Cohen J. The analysis of variance and covariance. In: Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale~~,~~ (NJ): Lawrence Erlbaum Associates; 1988. p. 273-~~-~~406.
20~~19~~.~~—~~ Davies EH, Surtees R, DeVile C, Schoon I, Vellodi A. A severity scoring tool to assess the neurological features of neuronopathic Gaucher disease. J Inherit Metab Dis 2007;30~~(5)~~:768-~~-~~82.
2~~0~~1.~~—~~ Fan X, Pope J, the Canadian Scleroderma Research Group, Baron M. What is the relationship between disease activity, severity and damage in a large Canadian systemic sclerosis cohort? Results from the Canadian Scleroderma Research Group (CSRG). Rheumatol Int 2009;30:1205-~~-~~10.
2~~1~~2.~~—~~ Bassel M, Hudson M, Taillefer SS, Schieir O, Baron M, Thombs BD. Frequency and impact of symptoms experienced by patients with systemic sclerosis: results from a Canadian National Survey. Rheumatology (Oxford) 2011;50~~(4)~~:762-~~-~~67.
2~~3~~2.~~—~~ Suarez-Almazor ME, Kallen MA, Roundtree AK, Mayes M. Disease and s~~S~~ymptom b~~B~~urden in s~~S~~ystemic s~~S~~clerosis: a~~A~~ p~~P~~atient p~~P~~erspective. J Rheumatol 2007;34~~(8)~~:1718-~~-~~26.

~~AQ~~AQ38 ~~Author: Reference 15: There is no article in the 2005 issue of A&R that begins on page 282. The authors and title listed here are the same as in reference 35 (which has the title and subtitle). Please check this reference and advise.~~

Formatted: Font: (Default) Times New Roman, 12 pt
Formatted: Font: (Default) Times New Roman, 12 pt
Formatted: Indent: Left: 0", First line: 0", Don't add space between paragraphs of the same style, Pattern: Clear (White)
Formatted: Font: Check spelling and grammar
Formatted: Indent: Left: 0", First line: 0", Don't add space between paragraphs of the same style
Formatted: Font: (Default) Times New Roman, Highlight
Formatted: Font: (Default) Times New Roman, Highlight
Formatted: Font: , Check spelling and grammar
Formatted: Font: Bold, Not Superscript/ Subscript
Formatted: Not Superscript/ Subscript
Formatted: Font: Not Bold
Formatted: Font: Univers, Not Superscript/ Subscript
Formatted: Font: Univers
Formatted: Underline
Formatted: Font: Univers

243. Stamm TA, Mattsson M, Mihai C, Stocker J, Binder A, Bauernfeind B, et al. Concepts of functioning and health important to people with systemic sclerosis: a qualitative study in four European countries. Ann Rheum Dis 2011;70(6):1074-9.

25. Ware JE Jr, Snow KK, Kosinski M, Gandek B. SF-36 health survey: manual and interpretation guide. Boston: The Health Institute, New England Medical Center; 1993.

246. Gneiting T, Raftery A. Strictly proper scoring rules. J Am Stat Assoc 2007;102:359-78.

275. Nagelkerke NGD. A note on a general definition of the coefficient of dDetermination. Biometrika 1991;78(3):691-2.

2628. Pope JE, Bellamy N, Seibold JR, Baron M, Ellman M, Carette S, et al. A randomized, controlled trial of methotrexate versus placebo in early diffuse scleroderma. Arthritis Rheum 2001;44(6):1351-8.

297. Furst D, Khanna D, Matucci-Cerinic M, Clements P, Steen V, Pope J, et al. Systemic sclerosis: continuing progress in developing clinical measures of response. J Rheumatol 2007;34(5):1194-1200.

30. Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. Arthritis Rheum 1980;23:137-45.

2831. Merkel PA, Silliman NP, Denton CP, Furst DE, Khanna D, Emery P, et al, for the CAT-192 Research Group and the Scleroderma Clinical Trials Consortium. Validity, reliability, and feasibility of durometer measurements of scleroderma skin disease in a multicenter treatment trial. Arthritis Rheum 2008;59(5):699-705.

32. AQ37 org Dyspnea ref. TK

2933. Merkel PA, Herlyn K, Martin RW, Anderson JJ, Mayes MD, Bell P, et al, for the Scleroderma Clinical Trials Consortium. Measuring disease activity and functional status in patients with scleroderma and Raynaud's phenomenon. Arthritis Rheum 2002;46(9):2410-20.

304. Khanna D, Clements PJ, Furst DE, Korn JH, Ellman M, Rothfield N, et al, for the Relaxin Investigators and the Scleroderma Clinical Trials Consortium. Recombinant human relaxin in the treatment of systemic sclerosis with diffuse cutaneous involvement: aA randomized, double-blind, placebo-controlled trial. Arthritis Rheum 2009;60(4):1102-11.

315. Tashkin DP, Elashoff R, Clements PJ, Goldin J, Roth MD, Furst DE, et al. Cyclophosphamide versus placebo in scleroderma lung disease. N Engl J Med 2006;354(25):2655-66.

362. Merkel PA, Clements PJ, Reveille JD, Suarez-Almazor ME, Valentini G, Furst DE. Current status of outcome measure development for clinical trials in systemic sclerosis: rReport from OMERACT 6. J Rheumatol 2003;30(7):1630-47.

337. Clements PJ, Hurwitz EL, Wong WK, Seibold JR, Mayes M, White B, et al. Skin thickness score as a predictor and correlate of outcome in systemic sclerosis: high-dose versus low-dose penicillamine trial. Arthritis Rheum 2000;43(11):2445-54.

348. Steen VD, Medsger TA, Jr. Severe organ involvement in systemic sclerosis with diffuse scleroderma. Arthritis Rheum 2000;43(11):2437-44.

Formatted: Font: Times New Roman

Formatted: Indent: Left: 0", First line: 0", Don't add space between paragraphs of the same style

Formatted: Font: Times New Roman

Formatted: Indent: Left: 0", First line: 0", Don't add space between paragraphs of the same style

Formatted: Font: Times New Roman, Highlight

Formatted: Font: Times New Roman

Formatted: Indent: Left: 0", First line: 0", Don't add space between paragraphs of the same style

Formatted: Font: (Default) Times New Roman, 12 pt

Formatted: Font: (Default) Times New Roman, 12 pt

35. Merkel PA, Silliman NP, Clements PJ, Denton CP, Furst DE, Mayes MD, et al, for the Scleroderma Clinical Trials Consortium. Patterns and predictors of change in outcome measures in clinical trials in scleroderma: an individual patient meta-analysis of 629 subjects with diffuse cutaneous systemic sclerosis. Arthritis Rheum 2012;64(10):3420–9.

369. Amjadi S, Maranian P, Furst DE, Clements PJ, Wong WK, Postlethwaite AE, et al, for the Investigators of the D-Penicillamine, Human Recombinant Relaxin, and Oral Bovine Type I Collagen Clinical Trials. Course of the modified Rodnan skin thickness score in systemic sclerosis clinical trials: aAnalysis of three large multicenter, double-blind, randomized controlled trials. Arthritis Rheum 2009;60(8):2490--8.

37. DUPLICATE OF REF 14. Classification and Response Criteria Subcommittee of the American College of Rheumatology Committee on Quality MeasuresSingh JA, Solomon DH, Dougados M, Felson D, Hawker G, Katz P, et al. Development of classification and response criteria for rheumatic diseases. Arthritis Rheum 2006;55(3):348–52.

3840. v Van der Heijde DM, van 't Hof MA, van Riel PL, van Leeuwen MA, van Rijswijk MH, van de Putte LB. Validity of single variables and composite indices for measuring disease activity in rheumatoid arthritis. Ann Rheum Dis 1992;51(2):177--81.

4139. Luijten KM, Tekstra J, Bijlsma JW, Bijl M. The Systemic Lupus Erythematosus Responder Index (SRI); a new SLE disease activity assessment. Autoimmun Rev 2012;11(5):326--9.

42. Symmons DP, Coppock JS, Bacon PA, Bresnihan B, Isenberg DA, Maddison P, et al, and Members of the British Isles Lupus Assessment Group (BILAG). Development and assessment of a computerized index of clinical disease activity in systemic lupus erythematosus. Q J Med 1988;69:927--37.

43. Bombardier C, Gladman DD, Urowitz MB, Caron D, Chang DH, and the Committee on Prognosis Studies in SLE. Derivation of the SLEDAI: a disease activity index for lupus patients. Arthritis Rheum 1992;35:630--40.

40. Yee CS, Isenberg DA, Prabu A, Sokoll K, Teh LS, Rahman A, et al. BILAG-2004 index captures systemic lupus erythematosus disease activity better than SLEDAI-2000. Ann Rheum Dis 2008;67(6):873–6.

414. Valentini G, D'Angelo S, Della RA, Bencivelli W, Bombardieri S. European Scleroderma Study Group to define disease activity criteria for systemic sclerosis. IV. Assessment of skin thickening by modified Rodnan skin score. Ann Rheum Dis 2003;62(9):904--5.

452. Medsger TA, Jr., Silman AJ, Steen VD, Black CM, Akesson A, Bacon PA, et al. A disease severity scale for systemic sclerosis: development and testing. J Rheumatol 1999;26(10):2159--67.

46. Giannini EH, Ruperto N, Ravelli A, Lovell DJ, Felson DT, Martini A. Preliminary definition of improvement in juvenile arthritis. Arthritis Rheum 1997;40:1202--9.

437. Steen VD, Mayes MD, Merkel PA. Assessment of kidney involvement. Clin Exp Rheumatol 2003;21(3 Suppl 29):29--31.

Formatted: Font: (Default) Times New Roman, 12 pt

Formatted: Indent: Left: 0", First line: 0", Don't add space between paragraphs of the same style, Pattern: Clear (White)

Formatted: Font: (Default) Times New Roman, 12 pt

Formatted: Indent: Left: 0", First line: 0", Don't add space between paragraphs of the same style

Formatted: Font color: Blue

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Indent: Left: 0", First line: 0", Don't add space between paragraphs of the same style

Formatted: Font: Times New Roman

Formatted: Indent: Left: 0", First line: 0", Don't add space between paragraphs of the same style

448. Hoeper MM, Bogaard HJ, Condliffe R, Frantz R, Khanna D, Kurzyna M, et al. Definitions and diagnosis of pulmonary hypertension. J Am Coll Cardiol 2013;62(25 Suppl):42-50.

<</ref1>>

<<label>>

Figure 1. Expert consensus and data-driven approaches used to develop the composite response index in systemic sclerosis (CRISS). dcSSc = diffuse cutaneous systemic sclerosis; OMERACT = Outcome Measures in Rheumatology.

Figure 2. A, Distribution of the predicted probability of improving among patients rated by the experts as improved (red curve) and patients rated by the experts as not improved (blue curve). **B,** Sensitivity and specificity of the predicted classification of patients as improved or not improved as a function of the predicted probability cutoff. The cutoffs considered were 0.1, 0.2, 0.3, … 0.9, and the predicted classifications were derived as follows: if the predicted probability for a patient is greater than the probability cutoff, the patient is rated as improved; otherwise, the patient is rated as not improved.

~~Figure 2. (a) Distribution of the predicted probability of improving for patients rated improved by the experts (red curve) and patients rated not improved by experts (blue curve). (b) Sensitivity (red line) and specificity (blue line) of the predicted classification of patients into "improved" and "not improved" as a function of the predicted probability cutoff. The cutoffs considered are 0.1, 0.2, 0.3, … 0.9 and the predicted classifications are derived as follow: if the predicted probability for a subject is greater than the probability cutoff, the subject is rated as "improved", otherwise subject is not.~~

<</label>>

Formatted: Font: Not Bold

Formatted: Font: Not Bold

Formatted: Font: (Default) Times New Roman, 12 pt

Formatted: Font: Not Bold

Formatted: Indent: Left: 0", First line: 0", Don't add space between paragraphs of the same style

Formatted: Indent: Left: 0", First line: 0", Space After: 0 pt, Don't add space between paragraphs of the same style

ACR Provisional Composite Index for Scleroderma Clinical Trials

Table 1: Baseline demographics of patients who participated in the CRISS Cohort with baseline and 1 year data

	Baseline N	
Age, mean (SD)	150	50.4 (11.7)
Female, N (%)		112 (75%)
Race, N (%) Caucasian African American Asian Other or not provided	150	 117 (78%) 13 (9%) 11 (7%) 9 (6%)
Ethnicity, N (%) Hispanic Non-Hispanic	150	 16 (11%) 134 (89%)
Disease duration from first non-Raynaud symptom (yrs), mean (SD)	144	1.59 (1.34)
Years since first Raynaud symptom, mean (SD)	128	2.87 (2.49)
Years since first non-Raynaud symptom, mean (SD)	129	2.32 (1.5)
Body mass index, mean (SD)	96	26.02 (7.1)
Modified Rodnan skin score, mean (SD)	150	21.4 (10.1)
Durometer, mean (SD)	113	272.4 (64.5)
Forced vital capacity % predicted, mean (SD)	140	82.32 (18.5)
Total lung capacity % predicted, mean (SD)	109	87.83 (20.4)
Diffusion capacity of carbon monoxide % predicted, mean (SD)	140	65.05 (20.9)
High resolution computer tomography consistent with interstitial lung disease, N (%)	99	79 (80)
6-minute walking distance, mean (SD)	50	421.6 (139.2)
Borg dyspnea (0-10 scale), mean (SD)	46	1.92 (1.51)
Tendon friction rubs, N (%)	140	40 (29)
Small joint contractures, N (%)	133	78 (59)
Large joint contractures, N (%)	133	39 (29)
Digital ulcers, N (%)	150	15 (10)
Health assessment questionnaire disability index, mean (SD)	150	1.0 (0.8)
Digital ulcers VAS (0-150), mean (SD)	134	20.9 (40.9)
Raynaud's VAS (0-150), mean (SD)	135	32.7 (40.8)
Breathing VAS (0-150), mean (SD)	138	23.1 (36.7)
GI VAS (0-150), mean (SD)	136	22.6 (34.4)
Disease severity VAS (0-150), mean (SD)	138	56.4 (42.9)
Pain VAS (0-10), mean (SD)	140	4.0 (2.8)
SF-36 PCS, mean (SD)	138	37.6 (12.9)
SF-36 MCS, mean (SD)	138	44.2 (6.0)

Formatted: Don't add space between paragraphs of the same style

Formatted ... [20]

Formatted ... [21]

Formatted ... [22]

Formatted ... [23]

Formatted ... [24]

Formatted ... [25]

Formatted ... [26]

Formatted ... [27]

Formatted ... [28]

Formatted ... [29]

Formatted ... [30]

Formatted ... [31]

Formatted ... [32]

ACR Provisional Composite Index for Scleroderma Clinical Trials

~~Physician global assessment VAS (0-10 cm), mean (SD)~~	~~143~~	~~4.4 (2.2)~~
~~Patient global assessment VAS (0-10 cm), mean (SD)~~	~~140~~	~~4.1 (4.0)~~
~~Antinuclear antibody, N (%)~~	~~116~~	~~94 (81)~~
~~Anti-SCL-70 antibody, N (%)~~	~~115~~	~~34 (30)~~
~~Serum creatine phosphokinase (IU/L) , mean (SD)~~	~~127~~	~~143.9 (184.5)~~
~~Serum platelets (k/uL), mean (SD)~~	~~143~~	~~315.2 (102.5)~~
~~Serum brain natriuretic peptide (pg/ml), mean (SD)~~	~~105~~	~~161.3 (824.0)~~
~~Serum erythrocyte sedimentation rate (mm/hr), mean (SD)~~	~~121~~	~~23.4 (22.6)~~
~~Serum C-reactive protein (mg/dL), mean (SD)~~	~~116~~	~~2.1 (4.9)~~

~~VAS=visual analog scale; PCS=Physical component scale; MCS=Mental component scale~~

Formatted: Don't add space between paragraphs of the same style

~~**Table 2. Final CRISS model consisting of 5 core items with highest face validity**~~

~~Core items (calculated as changed from baseline to 1 year)~~	~~Area under the curve (AUC)~~	~~Sensitivity (95% CI)~~	~~Specificity (95% CI)~~	~~Unadjusted Beta coefficients~~	~~Standard errors~~
~~MRSS~~ ~~FVC predicted~~ ~~HAQ-DI~~ ~~Patient global assessment~~ ~~Physician global assessment~~	~~0.9861~~	~~0.9821 (0.9816, 0.9827)~~	~~0.9310 (0.9300, 0.9321)~~	~~-0.81~~ ~~0.21~~ ~~-0.40~~ ~~-0.44~~ ~~-3.41~~	~~0.21~~ ~~0.08~~ ~~0.24~~ ~~0.26~~ ~~1.75~~

Formatted: Don't add space between paragraphs of the same style

Formatted: Don't add space between paragraphs of the same style

~~MRSS= modified Rodnan skin score, FVC= Forced vital capacity, HAQ-DI= health assessment questionnaire-disability index, MRSS= modified Rodnan skin score~~

ACR Provisional Composite Index for Scleroderma Clinical Trials

Table 3. The table describes ranking of the 16 core items by scleroderma experts and results of the cluster analysis

Core item	Rank 1 (%)	Rank 2 (%)	Rank 3 (%)	Cluster
MRSS	374 (44.1%)	131 (15.5%)	75 (8.9%)	1
FVC_% predicted	123 (14.5%)	148 (17.5%)	72 (8.5%)	1
Physician global assessment	77 (9.1%)	116 (13.7%)	88 (10.4%)	1
Patient global assessment	93 (11%)	69 (8.2%)	115 (13.6%)	1
HAQ-DI	68 (8%)	112 (13.2%)	99 (11.7%)	1
Vitality SF-36	12 (1.4%)	37 (4.4%)	101 (11.9%)	2
GI VAS	25 (2.9%)	44 (5.2%)	43 (5.1%)	2
Pain	11 (1.3%)	38 (4.5%)	82 (9.7%)	2
Tendon friction rubs	11 (1.3%)	33 (3.9%)	23 (2.7%)	2
Breathing VAS	13 (1.5%)	25 (3%)	32 (3.8%)	2
Digital ulcers VAS	7 (0.8%)	38 (4.5%)	17 (2%)	2
Raynaud's VAS	11 (1.3%)	18 (2.1%)	43 (5.1%)	2
Patient skin interference last month	2 (0.2%)	21 (2.5%)	22 (2.6%)	2
Number of digital ulcers	9 (1.1%)	11 (1.3%)	17 (2%)	2
Presence of renal crisis	11 (1.3%)	3 (0.4%)	2 (0.2%)	2
Body mass index	1 (0.1%)	3 (0.4%)	15 (1.8%)	2

MRSS= modified Rodnan skin score, FVC= Forced vital capacity, HAQ-DI= health assessment questionnaire-disability index, GI= gastrointestinal, VAS= visual analog scale, MRSS= modified Rodnan skin score

Formatted: Don't add space between paragraphs of the same style

Table 4. Application of CRISS in a clinical trial

CRISS is a 2-step process.

Step 1: Subjects who develop new or worsening of cardiopulmonary and/or renal involvement due to systemic sclerosis are considered as not improved (irrespective of improvement in other core items) and assigned a probability of improving equal to 0.0. Specifically if a subject develops any of the following

- **New scleroderma renal crisis (43)**
- **Decline in forced vital capacity (FVC)% predicted ≥15% (relative), confirmed by another FVC% within a month, high resolution computer tomography (HRCT) to confirm interstitial lung disease (ILD; if previous high resolution computer tomography of chest did not show ILD) and FVC% predicted below 80% predicted***
- **New onset of left ventricular failure (defined as left ventricular ejection fraction ≤45%) requiring treatment***
- **New onset of pulmonary arterial hypertension (PAH) on right heart catheterization (44) requiring treatment*.** PAH is defined as mean pulmonary artery pressure ≥ 25 mm Hg at rest and an end-expiratory pulmonary artery wedge pressure ≤ 15 mm Hg and a pulmonary vascular resistance >3 Wood units
-

***= Attributable to systemic sclerosis**

Step 2: For the remaining subjects, Step 2 involves computing the predicted probability of improving for each subject using the following equation (equation to derive predicted probabilities from a logistic regression model):

$$\frac{exp[-5.54 - 0.81 * \Delta_{MRSS} + 0.21 * \Delta_{FVC\%} - 0.40 * \Delta_{Pt-glob} - 0.44 * \Delta_{MD-glob} - 3.41 * \Delta_{HAQ-DI}]}{1 + exp[-5.54 - 0.81 * \Delta_{MRSS} + 0.21 * \Delta_{FVC\%} - 0.40 * \Delta_{Pt-glob} - 0.44 * \Delta_{MD-glob} - 3.41 * \Delta_{HAQ-DI}]}$$

where Δ_{MRSS} indicates the change in MRSS from baseline to follow up, Δ_{FVC} denotes the change in FVC% predicted from baseline to follow up, $\Delta_{Pt\text{-}glob}$ indicates the change in patient global assessment, $\Delta_{MD\text{-}glob}$ denotes the change in physician global assessment, and $\Delta_{HAQ\text{-}DI}$ is the change in HAQ-DI. All changes are absolute change ($Time_2 - Time_{baseline}$).

Definition of scleroderma renal crisis [adapted from **(43)**]

Formatted: Don't add space between paragraphs of the same style, Line spacing: single

Formatted: Don't add space between paragraphs of the same style

Formatted: Don't add space between paragraphs of the same style, No bullets or numbering

Formatted: Indent: Left: 0", Don't add space between paragraphs of the same style

Formatted: Don't add space between paragraphs of the same style

Formatted: Don't add space between paragraphs of the same style, Line spacing: single

Formatted: Don't add space between paragraphs of the same style

ACR Provisional Composite Index for Scleroderma Clinical Trials

A. Hypertensive SRC (fulfills both A1 and A2)

1. New onset hypertension, defined as any of the following:

a) Systolic blood pressure ≥ 140 mgHg
b) Diastolic blood pressure ≥ 90 mgHg
c) Rise in systolic blood pressure ≥ 30 mmHg
d) Rise in diastolic blood pressure ≥ 20 mmHg

AND

2. One (1) of the following five (5) features:

a) Increase in serum creatinine by 50+% over baseline OR serum creatinine ≥120% of upper limit of normal for local laboratory
b) Proteinuria ≥2+ by dipstick
c) Hematuria ≥2+ by dipstick or ≥10 RBCs/HPF
d) Thrombocytopenia: <100,000 platelets/mm^3
e) Hemolysis defined as anemia not due to other causes and either of the following:
(1) Schistocytes or other RBC fragments seen on blood smear
(2) increased reticulocyte count

B. Normotensive SRC (fulfills both B1 and B2)

1. Increase in serum creatinine >50% over baseline OR serum creatinine ≥120% of upper limit of normal for local laboratory

AND

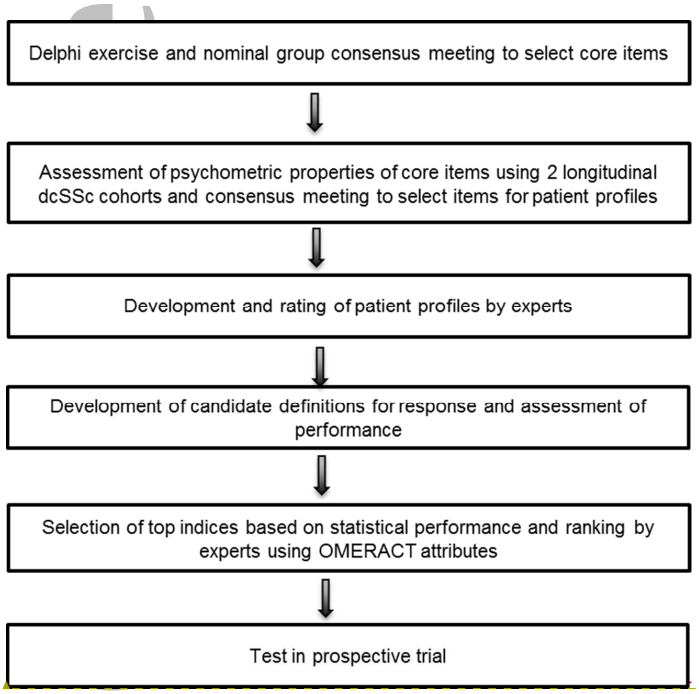
2. One (1) of the following five (5) features:

a) Proteinuria ≥2+ by dipstick
b) Hematuria ≥2+ by dipstick or ≥10 RBCs/HPF
c) Thrombocytopenia: <100,000/mm^3
d) Hemolysis defined as anemia not due to other causes and either of the following:
(1) Schistocytes or other RBC fragments seen on blood smear
(2) Increased reticulocyte count
e) Renal biopsy findings consistent with scleroderma renal crisis (microangiopathy)

Figure 1: Expert consensus and data-driven approaches used to develop CRISS

Formatted: Don't add space between paragraphs of the same style, Widow/Orphan control, Adjust space between Latin and Asian text, Adjust space between Asian text and numbers
Formatted: Font: (Default) Times New Roman, 12 pt, Font color: Auto
Formatted: Don't add space between paragraphs of the same style, Widow/Orphan control, Adjust space between Latin and Asian text, Adjust space between Asian text and numbers
Formatted: Font: (Default) Times New Roman, 12 pt, Font color: Auto
Formatted: Don't add space between paragraphs of the same style, Widow/Orphan control, Adjust space between Latin and Asian text, Adjust space between Asian text and numbers
Formatted: Font: (Default) Times New Roman, 12 pt, Font color: Auto
Formatted: Don't add space between paragraphs of the same style, Widow/Orphan control, Adjust space between Latin and Asian text, Adjust space between Asian text and numbers
Formatted: Font: (Default) Times New Roman, 12 pt, Font color: Auto
Formatted: Don't add space between paragraphs of the same style, Widow/Orphan control, Adjust space between Latin and Asian text, Adjust space between Asian text and numbers
Formatted: Font: (Default) Times New Roman, 12 pt, Font color: Auto
Formatted ... [33]
Formatted ... [34]
Formatted ... [35]
Formatted ... [36]
Formatted ... [37]
Formatted ... [38]
Formatted ... [39]
Formatted ... [40]
Formatted ... [41]
Formatted ... [42]
Formatted ... [43]
Formatted ... [44]
Formatted ... [45]
Formatted ... [46]
Formatted ... [47]

ACR Provisional Composite Index for Scleroderma Clinical Trials



Formatted: Font: (Default) Times New Roman

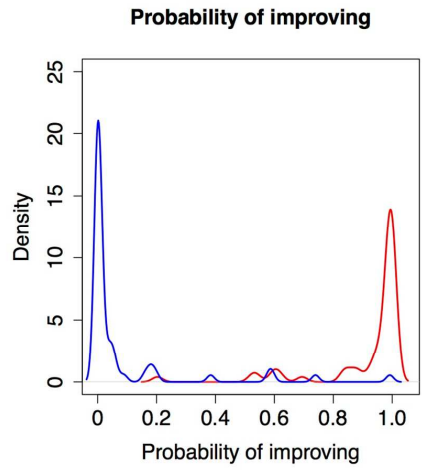
Formatted: Don't add space between paragraphs of the same style

Formatted: Left, Don't add space between paragraphs of the same style

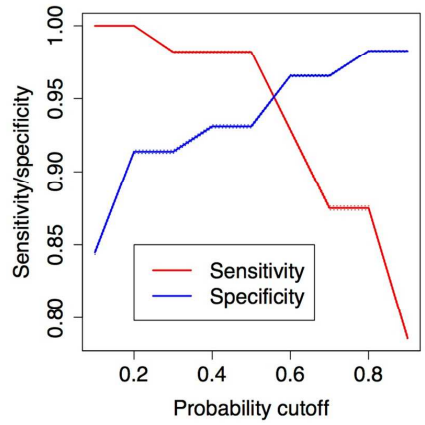
Formatted: Don't add space between paragraphs of the same style

Formatted: Left, Don't add space between paragraphs of the same style

Formatted: Don't add space between paragraphs of the same style



(a)



(b)

Figure 2. (a) Distribution of the predicted probability of improving for patients rated improved by the experts (red curve) and patients rated not improved by experts (blue curve). (b) Sensitivity (red line) and specificity (blue line) of the predicted classification of patients into "improved" and "not improved" as a function of the predicted probability cutoff. The cutoffs considered are 0.1, 0.2, 0.3, … 0.9 and the predicted classifications are derived as follow: if the predicted probability for a subject is greater than the probability cutoff, the subject is rated as "improved", otherwise subject is not.

Formatted: Font: (Default) Times New Roman

Formatted: Left, Don't add space between paragraphs of the same style

Formatted: Don't add space between paragraphs of the same style

Formatted: Font: (Default) Times New Roman

Formatted: Left, Don't add space between paragraphs of the same style

Formatted: Don't add space between paragraphs of the same style

ACR Provisional Composite Index for Scleroderma Clinical Trials

Appendix Table 1. Correlation between the continuous core items among the 14 core items at baseline.*

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14
V1	1.0	-0.26	0.43	0.60	0.33	0.49	0.31	0.04	0.16	0.09	0.09	0.04	0.03	0.17
V2		1.0	-0.22	-0.33	-0.23	-0.20	-0.18	0.02	-0.03	-0.17	-0.003	-0.11	-0.27	-0.16
V3			1.0	0.46	0.57	0.66	0.56	0.23	0.26	0.17	0.02	-0.06	0.28	0.25
V4				1.0	0.45	0.54	0.33	0.17	0.18	0.11	0.04	0.08	0.13	0.10
V5					1.0	0.55	0.57	0.35	0.35	0.19	-0.02	0.01	0.41	0.30
V6						1.0	0.60	0.19	0.44	0.26	0.11	0.06	0.30	0.22
V7							1.0	0.17	0.47	0.41	0.11	0.09	0.34	0.33
V8								1.0	0.15	0.06	-0.05	0.06	0.26	0.07
V9									1.0	0.35	0.20	0.15	0.39	0.45
V10										1.0	0.16	0.11	0.20	0.23
V11											1.0	-0.04	-0.02	0.02
V12												1.0	0.19	0.07
V13													1.0	0.36
V14														1.0

V1=MRSS, V2=FVC% predicted, V3=HAQ-DI, V4=Physician global, V5=Patient global, V6=Patient skin interference, V7=Pain, V8=Vitality, V9=Raynaud VAS, V10=Digital Ulcers VAS, V11=Number of digital ulcers, V12=BMI, V13=Breathing VAS, V14=GI VAS
*renal crisis and tendon friction rubs not included

Formatted: Don't add space between paragraphs of the same style

Appendix Table 2. Correlation between the change scores in the 14 core continuous core items.*

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14
V1	1.0	-0.30	0.22	0.26	0.16	0.32	0.21	0.12	0.17	0.17	-0.10	0.07	0.08	0.17
V2		1.0	-0.39	-0.31	-0.27	-0.29	-0.33	0.03	-0.06	-0.17	0.10	0.002	-0.30	-0.10
V3			1.0	0.17	0.27	0.31	0.23	-0.005	0.08	-0.05	-0.009	-0.18	0.30	0.05
V4				1.0	0.25	0.46	0.19	-0.09	0.18	0.03	-0.08	0.04	0.33	0.26
V5					1.0	0.13	0.25	-0.007	0.002	0.05	-0.14	-0.10	0.16	0.25
V6						1.0	0.28	-0.08	0.15	-0.07	-0.02	0.22	0.30	0.02
V7							1.0	0.07	0.27	0.10	0.22	0.11	0.33	0.23
V8								1.0	0.001	-0.12	-0.03	0.01	-0.12	-0.14
V9									1.0	0.20	0.35	0.20	0.23	0.47
V10										1.0	-0.13	0.11	0.05	0.36
V11											1.0	0.008	0.06	0.05
V12												1.0	0.16	0.07
V13													1.0	0.28
V14														1.0

V1=MRSS, V2=FVC% predicted, V3=HAQ-DI, V4=Physician global, V5=Patient global, V6=Patient skin interference, V7=Pain, V8=Vitality, V9=Raynaud VAS, V10=Finger Ulcers VAS, V11=Number of digital ulcers, V12=BMI, V13=Breathing VAS, V14=GI VAS

*renal crisis and tendon friction rubs not included

Formatted: Don't add space between paragraphs of the same style

ACR Provisional Composite Index for Scleroderma Clinical Trials

Appendix Table 3. Example of a patient rated "improved" by the experts. Predicted probability of improving is 0.99 according to CRISS.

	Baseline	Follow-up	Absolute change
Age	51.6 years		
Disease duration (months)	12.98		
Global assessments			
*Patient global assessment (0-10)**	3	1	-2
*Physician global assessment (0-10)**	3	3	0
Musculoskeletal			
*HAQ-DI (0-3)**	0.625	0	-0.625
Tendon friction rubs*	No	No	No change
Skin			
*MRSS (0-51)**	13	3	-10
Patient skin interference last month	2	0	-2
Lung			
*FVC% predicted**	62	75	13
Breathing VAS (0-10)	2	0	-2
Renal			
*Renal crisis***	No	No	No change
Gastrointestinal			
GI VAS (0-10)	3	3	0
Body Mass Index (BMI)	25.40	26.58	1.18
Raynaud's			
Raynaud's VAS (0-10)	2	1	-1
Digital ulcers			
Digital ulcers VAS (0-10)	0	0	0
Number of digital ulcers	0	0	0
HRQOL			
Pain VAS (0-10)	3	1	-2
Fatigue (SF-36 Vitality scale) (0-100)	42.31	35.12	-7.19

*included in Step 2; ** included in Step 1
HAQ-DI= health assessment questionnaire-disability index, MRSS= modified Rodnan skin score, FVC= Forced vital capacity, GI= gastrointestinal, VAS= visual analog scale

Formatted: Don't add space between paragraphs of the same style

ACR Provisional Composite Index for Scleroderma Clinical Trials

Appendix Table 4. Example of a patient rated "improved" by the experts. Predicted probability of improving is 0.60 according to CRISS.

	Baseline	Follow-up	Absolute change
Age	64.65 years		
Disease duration (months)	30.74		
Global assessments			
*Patient global assessment (0-10)**	1	0	-1
*Physician global assessment (0-10)**	7	4	-3
Musculoskeletal			
*HAQ-DI (0-3)**	0.375	0.250	-0.125
Tendon friction rubs*	No	No	No change
Skin			
*MRSS (0-51)**	21	15	-6
Patient skin interference last month	8	5	-3
Lung			
*FVC% predicted**	86	81	-5
Breathing VAS (0-10)	0	0	0
Renal			
*Renal crisis***	Yes	Yes	No change
Gastrointestinal			
GI VAS (0-10)	0	0	0
Body Mass Index (BMI)	25.12	24.82	-0.3
Raynaud's			
Raynaud's VAS (0-10)	3	4	1
Digital ulcers			
Digital ulcers VAS (0-10)	0	8	8
Number of digital ulcers	0	0	0
HRQOL			
Pain VAS (0-10)	0	2	2
Fatigue (SF-36 Vitality scale) (0-100)	35.12	35.12	0.0

*included in Step 2; ** included in Step 1

HAQ-DI= health assessment questionnaire disability index, MRSS= modified Rodnan skin score, FVC= Forced vital capacity, GI= gastrointestinal, VAS= visual analog scale

Formatted: Don't add space between paragraphs of the same style

Formatted [56]

Formatted [57]

Formatted [58]

Formatted [59]

Formatted [60]

Formatted [61]

Formatted [62]

Formatted [63]

Appendix Table 5. Example of a patient rated "worsened" by the experts. Predicted probability of improving is 0.002 according to the CRISS.

	Baseline	Follow-up	Absolute Change
Age	53.6 years		
Disease duration (months)	43.3		
Global assessments			
*Patient global assessment (0-10)**	1	2	1
*Physician global assessment (0-10)**	1	2	1
Musculoskeletal			
*HAQ-DI (0-3)**	0	0	0
Tendon friction rubs*	No	Yes	Change to worsen
Skin			
*MRSS (0-51)**	7	5	-2
Patient skin interference last month	3	2	-1
Lung			
*FVC% predicted**	87	80	-7
Breathing VAS (0-10)	0	1	1
Renal			
*Renal crisis***	No	No	No change
Gastrointestinal			
GI VAS (0-10)	0	1	1
Body Mass Index (BMI)	24.68	24.68	0
Raynaud's			
Raynaud's VAS (0-10)	0	3	3
Digital ulcers			
Digital ulcers VAS (0-10)	0	0	0
Number of digital ulcers	0	0	0
HRQOL			
Pain VAS (0-10)	1	1	0
Fatigue (SF-36 Vitality scale) (0-100)	37.52	35.10	-2.42

*included in Step 2; ** included in Step 1

HAQ-DI= health assessment questionnaire disability index, MRSS= modified Rodnan skin score, FVC= Forced vital capacity, GI= gastrointestinal, VAS= visual analog scale

Formatted: Don't add space between paragraphs of the same style

Formatted: Don't add space between paragraphs of the same style

Formatted: Don't add space between paragraphs of the same style

Formatted: Don't add space between paragraphs of the same style

Formatted: Don't add space between paragraphs of the same style

Formatted: Don't add space between paragraphs of the same style

Formatted: Don't add space between paragraphs of the same style

Formatted: Don't add space between paragraphs of the same style

Formatted: Don't add space between paragraphs of the same style

Formatted: Don't add space between paragraphs of the same style

Formatted: Don't add space between paragraphs of the same style

Formatted: Don't add space between paragraphs of the same style

Formatted: Don't add space between paragraphs of the same style

Formatted: Don't add space between paragraphs of the same style

Formatted: Don't add space between paragraphs of the same style

Formatted: Don't add space between paragraphs of the same style

Formatted: Don't add space between paragraphs of the same style

Formatted: Don't add space between paragraphs of the same style

Formatted: Don't add space between paragraphs of the same style

Formatted ... [64]

Formatted ... [65]

Formatted ... [66]

Formatted ... [67]

Formatted ... [68]

Formatted ... [69]

Formatted ... [70]

Formatted ... [71]

Appendix Table 6. One core item logistic model using expert consensus definition of improved vs.versus not

Core item	Area under the curve (AUC)	Sensitivity	Specificity	Brier Score
MRSS	0.9231	0.8392	0.8793	0.108
FVC% predicted	0.7906	0.6429	0.7586	0.184
Physician global	0.7743	0.7143	0.7241	0.197
Patient global	0.7448	0.7143	0.6207	0.204
HAQ DI	0.7107	0.6429	0.6897	0.200
Pain	0.6857	0.6071	0.7586	0.218
Vitality	0.6856	0.4643	0.7414	0.225
VAS Breathing	0.6670	0.375	0.8103	0.219
GI VAS	0.6667	0.7857	0.4483	0.220
Patient skin interference last month	0.6601	0.5179	0.7586	0.226
Raynaud's VAS	0.6190	0.4286	0.7241	0.238
Tendon friction rubs	0.5640	0.2321	0.8966	0.245
Digital ulcers VAS	0.5503	0.2857	0.7931	0.247
Body mass index	0.4946	0.1786	0.8276	0.250
Number of digital ulcers	0.4764	0.0179	0.931	0.249

HAQ DI= health assessment questionnaire disability index, MRSS= modified Rodnan skin score, FVC= Forced vital capacity, GI= gastrointestinal, VAS= visual analog scale

Formatted: Left, Don't add space between paragraphs of the same style

Formatted: Left, Don't add space between paragraphs of the same style

Formatted: Left, Don't add space between paragraphs of the same style

Formatted: Left, Don't add space between paragraphs of the same style

Formatted: Left, Don't add space between paragraphs of the same style

Formatted: Left, Don't add space between paragraphs of the same style

Formatted: Left, Don't add space between paragraphs of the same style

Formatted: Left, Don't add space between paragraphs of the same style

Formatted: Left, Don't add space between paragraphs of the same style

Formatted: Left, Don't add space between paragraphs of the same style

Formatted: Left, Don't add space between paragraphs of the same style

Formatted: Left, Don't add space between paragraphs of the same style

Formatted: Left, Don't add space between paragraphs of the same style

Formatted: Left, Don't add space between paragraphs of the same style

Formatted: Left, Don't add space between paragraphs of the same style

Formatted: Don't add space between paragraphs of the same style

~~Appendix Table 7. Two core item logistic model using expert consensus definition of improved vs. versus not~~

~~Core item~~	~~Area under the curve (AUC)~~	~~Sensitivity~~	~~Specificity~~	~~Brier Score~~
~~MRSS, FVC% predicted~~	~~0.9632~~	~~0.8929~~	~~0.9138~~	~~0.068~~
~~MRSS, HAQ-DI~~	~~0.9615~~	~~0.9107~~	~~0.8793~~	~~0.076~~
~~MRSS, Patient global~~	~~0.9560~~	~~0.875~~	~~0.8966~~	~~0.081~~
~~MRSS, physician global~~	~~0.9450~~	~~0.875~~	~~0.9310~~	~~0.094~~
~~FVC% predicted, HAQ-DI~~	~~0.8519~~	~~0.7679~~	~~0.8448~~	~~0.158~~
~~FVC% predicted, Patient global~~	~~0.8548~~	~~0.7679~~	~~0.8448~~	~~0.152~~
~~FVC% predicted, physician global~~	~~0.8544~~	~~0.750~~	~~0.8103~~	~~0.158~~
~~HAQ-DI, patient global~~	~~0.7982~~	~~0.7143~~	~~0.7241~~	~~0.184~~
~~HAQ-DI, physician global~~	~~0.8094~~	~~0.6607~~	~~0.7931~~	~~0.181~~
~~Patient global, physician global~~	~~0.8265~~	~~0.7321~~	~~0.7759~~	~~0.170~~

~~HAQ-DI= health assessment questionnaire disability index, MRSS= modified Rodnan skin score, FVC= Forced vital capacity~~

Formatted: Left, Don't add space between paragraphs of the same style

Formatted: Left, Don't add space between paragraphs of the same style

Formatted: Left, Don't add space between paragraphs of the same style

Formatted: Left, Don't add space between paragraphs of the same style

Formatted: Left, Don't add space between paragraphs of the same style

Formatted: Left, Don't add space between paragraphs of the same style

Formatted: Left, Don't add space between paragraphs of the same style

Formatted: Left, Don't add space between paragraphs of the same style

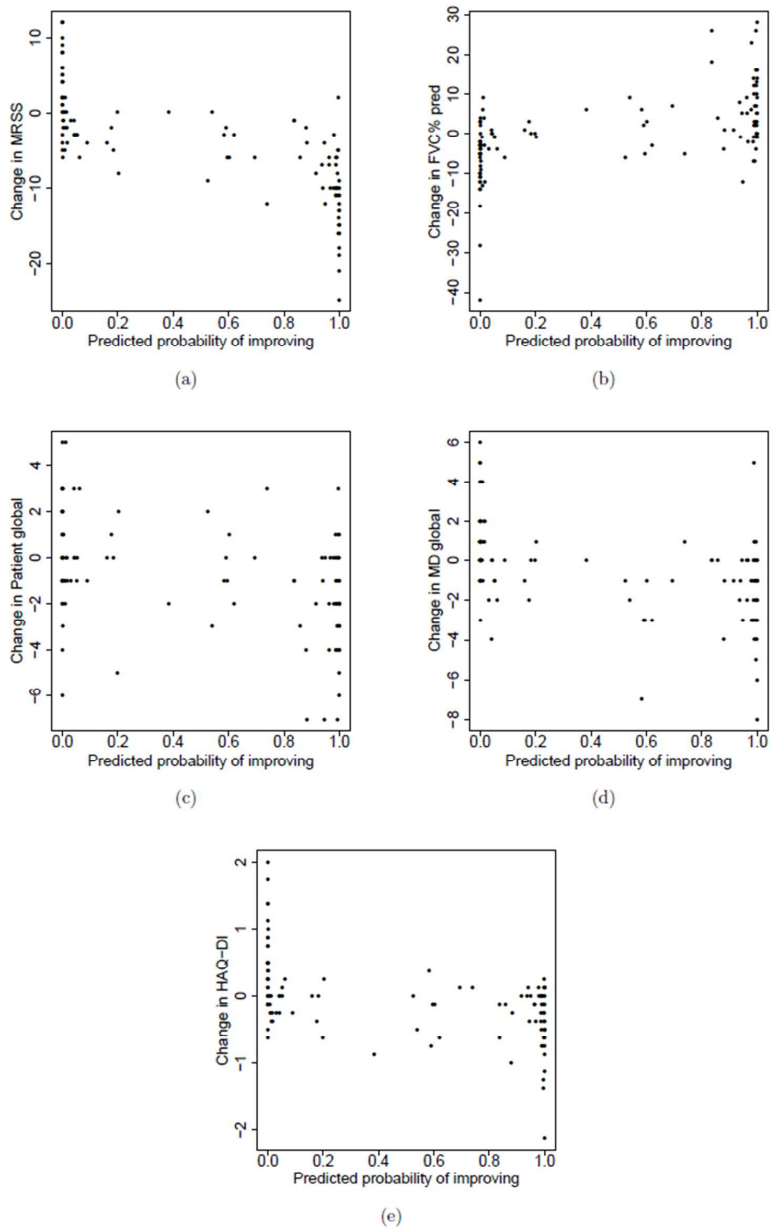
Formatted: Left, Don't add space between paragraphs of the same style

Formatted: Left, Don't add space between paragraphs of the same style

Formatted: Don't add space between paragraphs of the same style

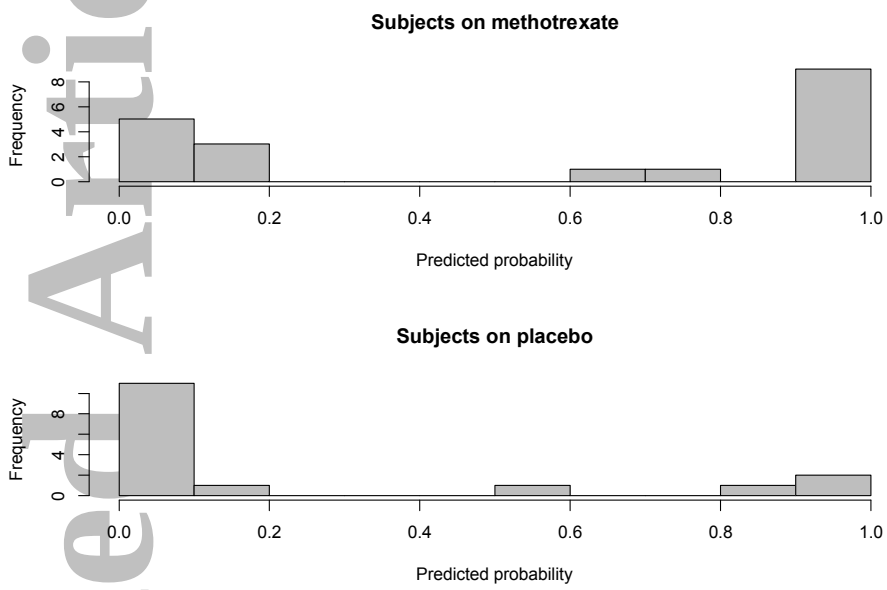
ACR Provisional Composite Index for Scleroderma Clinical Trials

~~Appendix Figure 1. (a) Change in MRSS, (b) Change in FVC% predicted, (c) Change in patient global assessment, (d) Change in physician global assessment, and (e) Change in HAQ-DI versus the predicted probability of improving yielded by CRISS.~~



Formatted: Font: (Default) Times New Roman

~~Appendix Figure 2. Histogram of the predicted probabilities of improving in subjects in the RCT study of methotrexate vs.~~versus ~~placebo.~~



Formatted: Font: (Default) Times New Roman

Page 1: [1] Formatted **Author**

Font: (Default) Times New Roman, 12 pt, Not Bold

Page 1: [2] Formatted **Author**

Font: (Default) Times New Roman, 12 pt, Not Bold

Page 1: [3] Formatted **Author**

Font: (Default) Times New Roman, 12 pt, Not Bold

Page 1: [4] Formatted **Author**

Don't add space between paragraphs of the same style

Page 1: [5] Formatted **Author**

Don't add space between paragraphs of the same style, Line spacing: single

Page 1: [6] Formatted **Author**

Font: Univers, Bold, Superscript, Highlight

Page 1: [7] Formatted **Author**

Font: (Default) Times New Roman, 12 pt, Font color: Blue

Page 1: [8] Formatted **Author**

Font: (Default) Times New Roman, 12 pt, Font color: Blue

Page 1: [9] Formatted **Author**

Don't add space between paragraphs of the same style, Line spacing: single

Page 1: [10] Formatted **Author**

Font: Univers, Bold, Not Superscript/ Subscript

Page 1: [11] Formatted **Author**

Indent: Left: 0", Hanging: 0.5"

Page 1: [12] Formatted **Author**

Font: Univers, Bold, Not Superscript/ Subscript

Page 9: [13] Formatted **Author**

Font: Bold

Page 9: [13] Formatted **Author**

Font: Bold

Page 9: [13] Formatted **Author**

Font: Bold

Page 9: [13] Formatted **Author**

Font: Bold

Page 9: [13] Formatted **Author**

Font: Bold

Page 9: [13] Formatted **Author**

Font: Bold

Page 9: [13] Formatted **Author**

Page 9: [13] Formatted **Author**
Font: Bold
Page 9: [13] Formatted **Author**
Font: Bold
Page 9: [14] Formatted **Author**
Font: Italic
Page 9: [14] Formatted **Author**
Font: Italic
Page 9: [14] Formatted **Author**
Font: Italic
Page 9: [14] Formatted **Author**
Font: Italic
Page 9: [14] Formatted **Author**
Font: Italic
Page 9: [14] Formatted **Author**
Font: Italic
Page 9: [14] Formatted **Author**
Font: Italic
Page 9: [14] Formatted **Author**
Font: Italic
Page 9: [14] Formatted **Author**
Font: Italic
Page 9: [14] Formatted **Author**
Font: Italic
Page 9: [14] Formatted **Author**
Font: Italic
Page 9: [14] Formatted **Author**
Font: Italic
Page 9: [15] Formatted **Author**
Font: Italic
Page 9: [15] Formatted **Author**
Font: Italic
Page 9: [15] Formatted **Author**
Font: Italic
Page 9: [15] Formatted **Author**
Font: Italic
Page 9: [15] Formatted **Author**

Font: Italic

Page 9: [15] Formatted Author

Font: Italic

Page 9: [15] Formatted Author

Font: Italic

Page 9: [15] Formatted Author

Font: Italic

Page 9: [15] Formatted Author

Font: Italic

Page 9: [15] Formatted Author

Font: Italic

Page 9: [15] Formatted Author

Font: Italic

Page 9: [15] Formatted Author

Font: Italic

Page 9: [15] Formatted Author

Font: Italic

Page 9: [15] Formatted Author

Font: Italic

Page 9: [15] Formatted Author

Font: Italic

Page 9: [15] Formatted Author

Font: Italic

Page 9: [15] Formatted Author

Font: Italic

Page 9: [15] Formatted Author

Font: Italic

Page 9: [15] Formatted Author

Font: Italic

Page 9: [15] Formatted Author

Font: Italic

Page 9: [15] Formatted Author

Font: Italic

Page 9: [15] Formatted Author

Font: Italic

Page 9: [15] Formatted Author

Font: Italic

Font: Italic

Page 9: [15] Formatted	Author

Font: Italic

Page 9: [15] Formatted	Author

Font: Italic

Page 9: [15] Formatted	Author

Font: Italic

Page 9: [15] Formatted	Author

Font: Italic

Page 9: [15] Formatted	Author

Font: Italic

Page 9: [15] Formatted	Author

Font: Italic

Page 9: [16] Formatted	Author

Font: Italic

Page 9: [16] Formatted	Author

Font: Italic

Page 9: [16] Formatted	Author

Font: Italic

Page 9: [16] Formatted	Author

Font: Italic

Page 9: [16] Formatted	Author

Font: Italic

Page 9: [17] Formatted	Author

Font: Univers, Bold, Not Superscript/ Subscript

Page 9: [17] Formatted	Author

Font: Univers, Bold, Not Superscript/ Subscript

Page 9: [17] Formatted	Author

Font: Univers, Bold, Not Superscript/ Subscript

Page 9: [17] Formatted	Author

Font: Univers, Bold, Not Superscript/ Subscript

Page 9: [17] Formatted	Author

Font: Univers, Bold, Not Superscript/ Subscript

Page 9: [17] Formatted	Author

Font: Univers, Bold, Not Superscript/ Subscript

Page 9: [17] Formatted	Author

Font: Univers, Bold, Not Superscript/ Subscript

Font: Univers, Bold, Not Superscript/ Subscript

Page 9: [18] Formatted Author

Font: Univers, Bold, Not Superscript/ Subscript

Page 9: [18] Formatted Author

Font: Univers, Bold, Not Superscript/ Subscript

Page 9: [19] Formatted Author

Font: Univers, Bold, Not Superscript/ Subscript

Page 9: [19] Formatted Author

Font: Univers, Bold, Not Superscript/ Subscript

Page 9: [19] Formatted Author

Font: Univers, Bold, Not Superscript/ Subscript

Page 31: [20] Formatted Author

Don't add space between paragraphs of the same style

Page 31: [21] Formatted Author

Don't add space between paragraphs of the same style

Page 31: [22] Formatted Author

Don't add space between paragraphs of the same style

Page 31: [23] Formatted Author

Don't add space between paragraphs of the same style

Page 31: [24] Formatted Author

Don't add space between paragraphs of the same style

Page 31: [25] Formatted Author

Don't add space between paragraphs of the same style

Page 31: [26] Formatted Author

Don't add space between paragraphs of the same style

Page 31: [27] Formatted Author

Don't add space between paragraphs of the same style

Page 31: [28] Formatted Author

Don't add space between paragraphs of the same style

Page 31: [29] Formatted Author

Don't add space between paragraphs of the same style

Page 31: [30] Formatted Author

Don't add space between paragraphs of the same style

Page 31: [31] Formatted Author

Don't add space between paragraphs of the same style

Page 31: [32] Formatted Author

Don't add space between paragraphs of the same style

Don't add space between paragraphs of the same style, Widow/Orphan control, Adjust space between Latin and Asian text, Adjust space between Asian text and numbers

Page 36: [34] Formatted **Author**

Font: (Default) Times New Roman, 12 pt, Font color: Auto

Page 36: [35] Formatted **Author**

Don't add space between paragraphs of the same style, Widow/Orphan control, Adjust space between Latin and Asian text, Adjust space between Asian text and numbers

Page 36: [36] Formatted **Author**

Font: (Default) Times New Roman, 12 pt, Font color: Auto

Page 36: [37] Formatted **Author**

Don't add space between paragraphs of the same style, Widow/Orphan control, Adjust space between Latin and Asian text, Adjust space between Asian text and numbers

Page 36: [38] Formatted **Author**

Font: (Default) Times New Roman, 12 pt, Font color: Auto

Page 36: [39] Formatted **Author**

Don't add space between paragraphs of the same style, Widow/Orphan control, Adjust space between Latin and Asian text, Adjust space between Asian text and numbers

Page 36: [40] Formatted **Author**

Font: (Default) Times New Roman, 12 pt, Font color: Auto

Page 36: [41] Formatted **Author**

Don't add space between paragraphs of the same style, Widow/Orphan control, Adjust space between Latin and Asian text, Adjust space between Asian text and numbers

Page 36: [42] Formatted **Author**

Font: (Default) Times New Roman, 12 pt, Font color: Auto

Page 36: [43] Formatted **Author**

Don't add space between paragraphs of the same style, Widow/Orphan control, Adjust space between Latin and Asian text, Adjust space between Asian text and numbers

Page 36: [44] Formatted **Author**

Font: (Default) Times New Roman, 12 pt, Font color: Auto

Page 36: [45] Formatted **Author**

Don't add space between paragraphs of the same style, Widow/Orphan control, Adjust space between Latin and Asian text, Adjust space between Asian text and numbers

Page 36: [46] Formatted **Author**

Don't add space between paragraphs of the same style

Page 36: [47] Formatted **Author**

Left, Don't add space between paragraphs of the same style

Page 41: [48] Formatted **Author**

Don't add space between paragraphs of the same style

Don't add space between paragraphs of the same style

Page 41: [50] Formatted Author

Don't add space between paragraphs of the same style

Page 41: [51] Formatted Author

Don't add space between paragraphs of the same style

Page 41: [52] Formatted Author

Don't add space between paragraphs of the same style

Page 41: [53] Formatted Author

Don't add space between paragraphs of the same style

Page 41: [54] Formatted Author

Don't add space between paragraphs of the same style

Page 41: [55] Formatted Author

Don't add space between paragraphs of the same style

Page 42: [56] Formatted Author

Don't add space between paragraphs of the same style

Page 42: [57] Formatted Author

Don't add space between paragraphs of the same style

Page 42: [58] Formatted Author

Don't add space between paragraphs of the same style

Page 42: [59] Formatted Author

Don't add space between paragraphs of the same style

Page 42: [60] Formatted Author

Don't add space between paragraphs of the same style

Page 42: [61] Formatted Author

Don't add space between paragraphs of the same style

Page 42: [62] Formatted Author

Don't add space between paragraphs of the same style

Page 42: [63] Formatted Author

Don't add space between paragraphs of the same style

Page 43: [64] Formatted Author

Don't add space between paragraphs of the same style

Page 43: [65] Formatted Author

Don't add space between paragraphs of the same style

Page 43: [66] Formatted Author

Don't add space between paragraphs of the same style

Page 43: [67] Formatted Author

Don't add space between paragraphs of the same style

Don't add space between paragraphs of the same style

Page 43: [69] Formatted Author

Don't add space between paragraphs of the same style

Page 43: [70] Formatted Author

Don't add space between paragraphs of the same style

Page 43: [71] Formatted Author

Don't add space between paragraphs of the same style