GENERATION OF A CORE SET OF ITEMS TO DEVELOP CLASSIFICATION CRITERIA FOR SCLERODERMA RENAL CRISIS (SRC) USING CONSENSUS METHODOLOGY

by

Emily-Ann Butler

Submitted in partial fulfilment of the requirements
for the degree of Master of Science

at

Dalhousie University
Halifax, Nova Scotia
March 2019

Dedication

I would like to thank my entire committee who were all so heavily involved in this thesis project. Thank you to my committee members, Dr. Yukiko Asada and Dr. Bradley Johnston, for their constant feedback and comments throughout this entire process. Thank you to my internal supervisor Dr. Kathleen MacPherson for her advice and continuous leadership. Lastly, I would like to thank my external supervisor Dr. Marie Hudson, whom none of this would have been made possible without. What started as a practicum placement in 2017 at the Jewish General Hospital under her supervision developed into a thesis project in the following months. Through her guidance and support this project was made possible.

Table of Contents

List of Tables

List of Figures

List of Abbreviations Used

SRC	Scleroderma renal crisis
SCTC	Scleroderma Clinical Trials Consortium
CSRG	Canadian Scleroderma Research Group
EUSTAR	European Scleroderma Trials and Research Group
ASIG	Australian Scleroderma Interest Group
SSc	Systemic sclerosis
ACE inhibitors	Angiotensin converting enzyme inhibitors
NGT	Nominal Group Technique
ACR/EULAR	American College of Rheumatology and European League against Rheumatism
LcSSc	Limited cutaneous systemic sclerosis
DcSSc	Diffuse cutaneous systemic sclerosis
SBP	Systolic blood pressure
DBP	Diastolic blood pressure
RBC	Red blood cells
HLA	Human leukocyte antigen
HPF	High Power Field
RAM	RAND Appropriateness Methods
IPR	Interpercentile range
IPRAS	Interpercentile range adjusted for symmetry
NIH	National Institute of Health
COPD	Chronic Obstructive Pulmonary Disease
IQR	Interquartile range
MAHA	Microangiopathic hemolytic anemia
MAHAT	Microangiopathic hemolytic anemia and thrombocytopenia
KDIGO	Kidney disease improving global outcomes
AKI	Acute kidney injury
LDH	Lactate dehydrogenase
ISRCS	International Scleroderma Renal Crisis Surveys
OMERACT	Outcome Measures in Rheumatology

Abstract

Background: Scleroderma Renal Crisis (SRC) is characterized by malignant hypertension and acute kidney injury. The absence of a gold standard or classification criteria for SRC has hindered research in this field. The Scleroderma Clinical Trials Consortium (SCTC) SRC Working Group was created to develop consensus and data-driven classification criteria for SRC. This project was undertaken to generate a core set of items using consensus methodology to be considered in the development of classification criteria for SRC.

Methods: A survey using items identified by a scoping review was developed (REDCap platform, Vanderbilt University, Nashville, Tennessee). An international, multidisciplinary panel of experts from the SCTC, European Scleroderma Trials and Research Group (EUSTAR), Canadian Scleroderma Research Group (CSRG), and Australian Scleroderma Interest Group (ASIG) were invited to participate in a 3-round Delphi exercise. In Round 1, participants were asked to identify omissions and clarify ambiguities regarding the items in the survey. In Round 2, participants were asked to rate the validity and feasibility of the items using Likert-type scales ranging from 1-9 (1= very invalid/unfeasible, 5 = uncertain, 9 = very valid/feasible), and to provide comments. In Round 3, participants reviewed the results and comments of Round 2, and were asked to provide final ratings. Items rated as highly valid and feasible (both median scores ≥ 7) in Round 3 were selected as the provisional core set of items. A nominal group discussion meeting followed the Delphi exercise to achieve final consensus on the core set of items.

Results: Overall, 216 experts were invited and 99 from 16 countries agreed to participate in the Delphi exercise. Of the 31 items in the survey, consensus was achieved on 13 items pertaining to hypertension, renal insufficiency, proteinuria and hemolysis. Eleven experts took part in the nominal group discussion, where consensus was achieved for 5 domains: blood pressure, kidney injury, microangiopathic hemolytic anemia, target organ dysfunction, and histopathology.

Conclusions: A core set of items defining SRC was identified using consensus methodology. Future data-driven phases of the project are planned to develop classification criteria for SRC.

CHAPTER 1. Introduction:

Scleroderma renal crisis (SRC) is a life-threatening complication of systemic sclerosis (SSc) characterized by malignant hypertension and acute kidney injury. With a high mortality rate, SRC remains a leading cause of death among patients with SSc (1). Individuals who are diagnosed promptly may have better survival, due to early initiation of treatment with angiotensin converting enzyme (ACE) inhibitors (2–4). Nevertheless, outcomes remain poor and there is an urgent need to identify novel therapeutic options (5).

One of the major hurdles in the study of SRC is the absence of a gold standard or validated classification criteria. The latter are essential to facilitate robust research, identify novel treatments and ultimately improve the outcomes of patients with SRC (6).

This research project was designed to develop a core set of items to be considered in the development of classification criteria for SRC, using consensus methodology. In Phase 1, an online modified Delphi survey was used to achieve initial consensus on a core set of items. In Phase 2, a Nominal Group Technique (NGT) meeting was held to discuss the results of the Delphi survey and to achieve final consensus on the core set.

This thesis project is a stepping stone in the development of classification criteria for SRC. Previously, a scoping review had been conducted to compile definitions of SRC proposed to date. That work was used to inform this thesis research project, consisting of a Delphi exercise followed by an NGT meeting, to generate a core set items to define SRC. These items will be moved into future data-driven phases to develop and validate classification criteria.

CHAPTER 2. Study Background and Literature Review:

2.1 Systemic Sclerosis

Systemic sclerosis (SSc), also known as scleroderma, is an autoimmune disease characterized by vasculopathy, fibrosis of the skin and internal organs, and immune abnormalities including the production of disease-specific autoantibodies (7). The 2013 American College of Rheumatology and European League against Rheumatism (ACR/EULAR) classification criteria for SSc include the following items: skin thickening of the fingers of both hands extending proximally to the metacarpophalangeal joints, Raynaud's phenomenon, SSc-related autoantibodies, fingertip lesions, telangiectasia, abnormal nailfold capillaries and pulmonary arterial hypertension and/or interstitial lung disease (8).

Across numerous studies, incidence rates for SSc appear to be relatively consistent, with approximately 20 new cases per million individuals per year (9). Prevalence has been estimated at about 240 per million adults in the United States. Systemic sclerosis is more prevalent in females and in middle-aged adults (10). Female to male ratios ranging from 4:1 to 7:1 have been documented with an increase in such ratios during child bearing years. This increase in prevalence among females of child-bearing age is thought to be associated with hormones and/or pregnancy-related events, however, research regarding this remains limited (9,11). Various genetic risk factors for the development of SSc have been documented and include HLA associated alleles (9).

Systemic sclerosis is often categorized as either limited cutaneous systemic sclerosis (lcSSc) or diffuse cutaneous systemic sclerosis (dcSSc). In lcSSc, skin involvement is restricted to the distal limbs and the face, while in dcSSc, skin involvement can extend to the proximal limbs and the trunk. LcSSc is thought to be associated with a more indolent course, while dcSSc tends to progress more rapidly and be associated with higher mortality. Approximately 40% of individuals with SSc have dcSSc (7,12).

2.2 Scleroderma Renal Crisis

2.2.1 Clinical signs and symptoms of SRC

Scleroderma renal crisis (SRC) is a life-threatening complication of SSc (13). It is usually characterized by malignant hypertension and acute kidney injury. However, the clinical spectrum

of SRC is broad, ranging from full-blown disease presenting as new onset of accelerated arterial hypertension and rapidly progressive oliguric renal failure, to more modest elevations in blood pressure and renal dysfunction, and at times, normotensive presentations. On the other hand, hypertension without uraemia, and urinary abnormalities and/or mild uraemia attributable to other factors (e.g., concomitant comorbidities such as diabetes or exposure to nephrotoxic medications) are common in SSc (14,15). These conditions should not be confused with SRC.

Existing definitions of SRC were compiled in a scoping review of the literature conducted by Hoa et al. Table 1 provides a summary of the results (see full table in Appendix 1) (16). Items used in these definitions were grouped into 11 domains: hypertension; renal insufficiency; proteinuria; hematuria; thrombocytopenia; hemolysis; encephalopathy; retinopathy; hyperreninemia; cardiac dysfunction; and abnormal kidney biopsy. Typically, in hypertensive SRC, hypertension in addition to at least one of the items listed in Table 1 are required for diagnosis. In normotensive SRC, elevated serum creatinine levels in addition to at least one other item, again listed in Table 1, are required for diagnosis (16).

Table 1. SRC domains in current literature identified by scoping review (16)

Domain	Items
Hypertension	Increased systolic blood pressure (SBP) and/or diastolic blood pressure (DPB)
Renal Insufficiency	Reduced kidney function, measured by serum creatinine levels
Proteinuria	Excess of protein in urine, measured by urine dipstick, protein:creatinine ratio, or 24-hour collection
Hematuria	Presence of blood in urine, measured by dipstick or microscopy
Thrombocytopenia	Low levels of platelets in blood
Hemolysis	Destruction of red blood cells, identified by blood smear and supported by various lab tests
Encephalopathy	Altered mental status and seizures
Retinopathy	Damage to the retina
Hyperreninemia	Elevated plasma renin levels
Cardiac Dysfunction	Flash pulmonary edema and/or pericardial effusion
Abnormal Kidney Biopsy	Abnormalities in arteries

2.2.2 Epidemiology of SRC

Scleroderma renal crisis is a rare complication of SSc, occurring in about 5% of SSc patients overall (13). SRC is more common in patients with rapidly progressing dcSSc (11%) as

compared to patients with lcSSc (4%) (17). Historically, SRC was the leading cause of death in SSc (1). However, with the advent of ACE inhibitors, one-year mortality rates have decreased significantly (2,3). Despite this decline, SRC remains a severe complication, often resulting in the need for dialysis (18). Hesselstrand et al. noted a mortality odds ratio of 4.39 (95% CI 2.10, 9.26) for SSc patients with versus without SRC (19). One-year outcomes remain poor, with over 30% mortality and 25% of patients remaining dialysis-dependent (5).

SRC commonly occurs early in the course of SSc (from one to four years after diagnosis of SSc) (14,19,20) and average age of onset is approximately 50 years (2,3,20,22,23). About 80% of SRC patients are female (2,20–23).

Other risk factors for SRC include rapidly progressing dcSSc, anti-RNAP III antibodies (22,24,25), exposure to corticosteroids (3) (26), presence of select HLA (Human leukocyte antigen) alleles (27) and presence of membrane protein CD147 (28).

2.2.3 Hypertensive SRC

Approximately 90% of patients with SRC have increased blood pressure (29,30). Previous studies have shown the importance of high blood pressure in detecting SRC, as hypertension is one of the earliest signs in many cases (21). However, definitions of increased blood pressure vary, including: systolic blood pressure (SBP) greater than 140-180 mmHg; diastolic blood pressure (DPB) greater than 90-120 mmHg; and increases in SBP of $\geq$ 30 mmHg and in DPB of $\geq$ 20 mmHg over baseline measurements (16). The varied definitions of increased blood pressure highlight the challenge in identifying hypertensive SRC.

2.2.4 Normotensive SRC

About 10% of patients with SRC have a normotensive form (29–32), characterized by acute kidney dysfunction in the absence of hypertension (29–32). Lack of hypertension may be a result of cardiac dysfunction; in particular, decreased function of the left ventricle may limit the ability to increase blood pressure (19,31,33). Some studies have shown that normotensive SRC patients may experience microangiopathic hemolytic anemia (hemolysis) and thrombocytopenia more commonly than hypertensive SRC patients (32,34)

Normotensive SRC is associated with worse outcomes than hypertensive SRC, including less recovery of renal function and higher mortality (3,13,29,33,34). It is possible that these poorer outcomes may be explained, at least in part, by delayed diagnosis (in the absence of

hypertension) or by poor cardiac function. Exposure to corticosteroids may be a greater potential risk factor for normotensive SRC than for hypertensive SRC, although evidence from the current literature is limited (13,29).

2.2.5 Treatment of SRC

Since the advent of angiotensin converting enzyme (ACE) inhibitors, SRC is no longer the leading cause of death among patients with SSc (35). Hypertension in SRC is mediated by hyper-reninemia and ACE inhibitors specifically target this pathway (32). ACE inhibitors also indirectly reduce hypertension through decreased production of angiotensin. Exposure to ACE inhibitors causes blood vessels to dilate, reducing blood pressure levels (30,36). Thus, ACE inhibitors can help control hypertension in SRC (29). However, in order to substantially improve survival rates, it is essential to administer ACE inhibitors promptly (2–4).

In addition to ACE inhibitors, dialysis is frequently required in SRC patients. If patients undergoing dialysis do not recover from renal failure, their mortality rate is high, with survival rates of less than 60% after one year, dropping to 20% after eight years (3,31).

Another possible treatment for SRC is kidney transplantation. Since renal recovery can continue up to 2 years after SRC, transplantation is delayed until then (32). Unfortunately, in addition to potential complications of the transplantation itself, patients must be aware that SRC can recur following a kidney transplant. Nevertheless, if the transplanted kidney is not rejected, survival rates are high (19).

2.3 SRC criteria proposed to date

Current research on SRC typically uses *ad hoc* criteria for defining this disease. The lack of consistent criteria hinders our understanding of SRC, due to the inability to compare and generalize research surrounding this rare disease.

To date, two key efforts have been made towards developing classification criteria for SRC. The following sections (2.2.1 and 2.2.2) discuss the relevant studies in which criteria were proposed (29) and partially validated (5).

2.3.1 ANCONA criteria proposed by Steen et al. (29)

The first efforts towards developing classification criteria for SRC were made by Steen et al. in 2003 (29). Through expert consensus, Steen and colleagues clarified the involvement of the kidney in SRC and SSc. The project went on to address the inconsistency in how renal abnormalities were used to define SRC, and how kidney signs and symptoms could result from other complications. Overall, a core set of three items for detection of renal disease in SSc patients was identified: 1) blood pressure, both systolic and diastolic, 2) serum creatinine, and 3) urinalysis, both dipstick and microscopic. In addition, 13 other items were identified for their association with SRC. These items were classified as 'other' due to their indirect involvement with alternative conditions, rather than a direct relationship to SSc and SRC. Steen et al. concluded that SRC should be classified primarily by specific SRC abnormalities, and proposed the following criteria known as the ANCONA criteria, for hypertensive and normotensive SRC:

A) Hypertensive scleroderma renal crisis

New onset hypertension; defined as any of the following:

a) Systolic blood pressure ≥ 140 mg Hg

b) Diastolic blood pressure ≥ 90 mg Hg

c) Rise in systolic blood pressure ≥ 30 mm Hg

d) Rise in diastolic blood pressure ≥ 20 mm Hg

And 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 plts/mm3

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 scleroderma renal crisis

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

And 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 /mm3

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)

There was no attempt to validate the proposed criteria.

2.3.2 Criteria proposed by Hudson et al. (5)

More recently, a study by Hudson et al. (2014) produced the first set of SRC criteria to be partially validated. In a prospective cohort study of incident SRC patients, the main objective of this study was to determine if ACE inhibitors administered prior to SRC onset would result in worse health outcomes (mortality rates and dialysis during the first year following SRC onset). Among the 75 incident SRC cases included, 21% were previously exposed to ACE inhibitors. The overall one year mortality rate was 36%, and 25% of patients remained on dialysis. Exposure to ACE inhibitors prior to SRC diagnosis was associated with a greater then 2-fold risk of death, compared to patients not exposed to ACE inhibitors prior to SRC diagnosis (5).

Nested within the study, the investigators attempted to validate criteria for SRC. They proposed the following criteria, which are different from the ANCONA criteria in 2 key respects, namely inclusion of an item for hypertensive encephalopathy and non-inclusion of renal biopsy findings for normotensive SRC:

A) Hypertensive SRC

Any one of the following:

a) Systolic blood pressure > 140 mmHg, or

b) Diastolic blood pressure > 90 mmHg, or

c) Rise in systolic blood pressure > 30 mmHg compared to baseline, or

d) Rise in diastolic blood pressure > 20 mmHg compared to baseline

And one of the following features:

a) Increase in serum creatinine >50% over baseline OR serum creatinine >120% of upper limit of normal for local laboratory

b) Proteinuria: >2+ by dipstick and confirmed by protein:creatinine ratio > upper limit of normal

c) Hematuria: >2+ by dipstick or >10 RBCs/HPF (without menstruation)

d) Thrombocytopenia: <100,000 platelets/mm3

e) Hemolysis: by blood smear or increased reticulocyte count

f) Hypertensive encephalopathy

B) Normotensive SRC

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

And one of the following features:

a) Proteinuria: 42þ by dipstick and confirmed by protein:creatinine ratio 4 upper limits of normal

b) Hematuria: >2+ by dipstick or >10 RBCs/HPF (without menstruation)

c) Thrombocytopenia: <100,000 platelets/mm3

d) Hemolysis: by blood smear or increased reticulocyte count

e) Hypertensive encephalopathy

In the absence of a true gold standard, Hudson et al. used the physician diagnosis of SRC as the reference standard. They found that 70/70 hypertensive SRC patients met the proposed criteria for hypertensive SRC whereas only two of the five normotensive patients met the criteria for normotensive SRC. The results were the same when they used the ANCONA criteria. However, kidney biopsies were not available for any of the patients with normotensive SRC. Whether this would have resulted in better performance of the ANCONA criteria for normotensive SRC remains unknown.

The main limitations of the criteria in this study therefore include a set of criteria generated in an *ad hoc* manner by the study investigators, an imperfect gold standard, and inability to correctly classify normotensive SRC with the proposed criteria. These limitations

highlight the need to develop and validate classification criteria for SRC using robust methodology.

2.3.3 Scoping review by Hoa et al. (16)

The Scleroderma Clinical Trials Consortium identified the need to develop and validate classification criteria for SRC as a priority in 2016. Thus, the SRC Working Group was created. The first effort of the working group was to perform a scoping review to identify definitions of SRC that have been used in the published literature to date. This review included an extensive search in three online databases, Medline, Embase and non-Ovid Pubmed. A 'snowball technique' and search of reference lists contributed to the search of relevant material to be included. Articles written in English and specifically addressing SRC were considered. Articles were excluded if they did not use human data (16).

The search identified 4,158 articles, of which 415 met inclusion criteria. Forty original definitions of SRC, with significant heterogeneity, were identified from 36 studies, nine reviews and two editorials. All noted items were included as candidate items for defining SRC. The final list consisted of 11 domains and 48 items (16).

The 11 domains were: hypertension, renal insufficiency, proteinuria, hematuria, thrombocytopenia, hemolysis, hypertensive encephalopathy, hypertensive retinopathy, hyperreninemia, abnormal kidney biopsy, and flash pulmonary edema. Each domain, except flash pulmonary edema, had a variable number of items: renal insufficiency had 14, hypertension had 12, hemolysis had five, proteinuria and hematuria each had four, abnormal kidney biopsy had three, hypertensive retinopathy had two, and thrombocytopenia, hypertensive encephalopathy and hyperreninemia each had one (16). For the complete list of domains and items, see Appendix 1. The scoping review laid key ground work for this thesis project, by providing a starting list of possible items to include in the Delphi exercise.

2.3.4 Definitions of Acute Kidney Injury as well as Microangiopathic Hemolytic Anemia and Thrombocytopenia

Various items identified in the scoping review have been defined and validated in settings outside of SRC. The Kidney Disease Improving Global Outcomes (KDIGO) is a global organization that works towards developing guidelines for Kidney Disease. These guidelines are

developed for preventing and managing various kidney diseases, including Acute Kidney Injury (AKI) (37). In this setting, they have defined Acute Kidney Injury (AKI) as follows:

(1) Increase in serum creatinine by ≥0.3 mg/dl (≥26.5 lmol/l) within 48 hours; or
(2) Increase in serum creatinine to ≥1.5 times baseline, which is known or presumed to have occurred within the prior 7 days; or
(3) Urine volume <0.5 ml/kg/h for 6 hours.

Microangiopathic hemolytic anemia and thrombocytopenia (MAHAT) has been defined by the International Working Group on thrombotic thrombocytopenia purpura (TTP) and associated thrombotic microangiopathies (TMAs) and the American Society of Hematology. It includes MAHAT defined as new or worsening anemia not due to other causes, schistocytes or other RBC fragments on blood smear, laboratory evidence of hemolysis that includes elevated lactate dehydrogenase (LDH) and reticulocytes and/or low/absent haptoglobin and a negative direct anti-globulin test. Additionally, it includes a platelet count of ≤ 100,000 confirmed by blood smear for thrombocytopenia (38,39).

These definitions were used in final stages of this project to inform the definitions of items retained in the core set.

2.4 Assessing validity

As this study focuses on the concept of working towards developing possible classification criteria, validation of such criteria should be discussed. Validity in this context refers to the ability for such criteria to actually classify the disease of interest.

Measurement validity reflects the extent to which an instrument truly measures what it was intended to measure (40). Face validity looks at whether, on the surface, the instrument being tested appears to measure the construct of interest (41). Content validity, also termed logic and rational validity, examines the extent to which an instrument or set of criteria incorporates the relevant construct or domain (40,42,43). Construct validity looks at how well an instrument actually measures what it claims to measure (44). This can be assessed in a number of ways. When a previous instrument exists, then the comparison of old and new instruments through the administration of both can help demonstrate validity. Further comparisons of specificity and sensitivity can also be used to test the new instrument against existing criteria. If no other instrument exists, observing the relationships that arise from administration of the instrument

compared to general anticipated outcomes can also help establish validity. The use of different populations can also be used to assess an anticipated hypothesis; if the expected relationship is found, the instrument can be inferred to be measuring what it is intended to measure (41). Finally, criterion validity examines how well a certain instrument measures the construct compared to the gold standard (45). However, since there is no gold standard for SRC, this level of validation is challenging in this setting.

This thesis project will explore face and content validity on a core set of items. The judgement of experts can support the face or content validity of criteria. Having experts discuss and develop criteria can further contribute towards validation, given that they have the greatest working knowledge on the subject at hand (43). Achieving content validity can be done in a two-stage process, where 1) a construct is identified and domains pertaining to this construct are produced and organized into an instrument, and then 2) a panel of experts discuss, modify and agree on domains and items within each domain to develop an overall instrument for measuring the construct of interest (40,42). Consensus agreement on domains and items is necessary for content validation (43).

2.5 Consensus building methods

This thesis project was conducted using consensus methodology. In this section, various types of consensus methods that exist will be discussed. Through the exploration of the various methods including their respective strengths and limitations, this section will provide rationale for the methodology used for this project.

Different techniques can be used for decision making and reaching consensus in healthcare research. These techniques include brainstorming, focus groups, nominal group technique (NGT) meetings, and the Delphi method. Brainstorming groups feature several individuals discussing a topic at hand. Although brainstorming is a good practice for preliminary steps in research, this method does not work towards achieving consensus but rather bringing about ideas (46–48). A focus group, typically consisting of several participants, works with a moderator to discuss a common area of interest or topic at hand. Analyzing results, drawing conclusions and assessing consensus can be challenging, due to the qualitative, free discussion nature of the process. Further issues can arise when there is pressure to conform due to limited numbers of participants and hierarchies that may exist within the groups.

A similar approach is an NGT meeting which comprises an in-person gathering of a small group of individuals, specifically, experts in the field of interest. A moderator typically leads the discussion. In addition to the qualitative data that arises from discussion, NGT members vote on topics discussed. This approach can suffer from pressure to conform, similar to focus groups. In contrast, the Delphi method provides a more quantitative approach to decision making. This technique consists of multiple rounds of almost identical surveys distributed to a variety of individuals. Participants are asked to rank items within the survey and are provided with results from prior rounds to inform ranking, as they work towards consensus on items.

Other methods have been used in previous research for consensus building but are more specific to their respective applications. For example, the National Institute of Health (NIH) works at providing consensus on safety and appropriateness of medical practices, devices and drugs. Consensus development conferences follow a structured format, beginning with a literature review, followed by presentations by investigators, an open discussion to allow for questions and comments, then ending with closed deliberations by a smaller group of individuals. Everyone must agree on the final decision for consensus to occur. Another example is Glaser's approach to describing current knowledge around Chronic Obstructive Pulmonary Disease (COPD) and developing COPD guidelines (42,45). Glaser himself reached out to a small group of individuals, who were in turn encouraged to contact others. A report was developed and then circulated to all participants for opinions/comments and approvals before the completion of a final draft.

Overall, the combination of a Delphi exercise and a NGT meeting is what we believe to be the most appropriate and rigorous approach to achieve the objective of this study. These approaches complement each another; some limitations of one are strengths of the other. The additional combination of both quantitative and qualitative data collection will contribute to the strength of the combination of these methodologies. Within this thesis research project, the results of the Delphi provide the starting point for the NGT meeting. The in-person NGT meeting seeks to further elaborate on the Delphi results and further strengthen agreement. Thus, by conducting both the Delphi and NGT, these methods aim to achieve overall, well-acknowledged agreement.

2.6 RAND/UCLA Appropriateness Methods (RAM)

2.6.1 Delphi Exercise

The Delphi exercise was designed and developed by RAND Corporation in the 1950s and is best outlined in the RAND/UCLA Appropriateness Methods (RAM) manual (49). The Delphi exercise can be used for a variety of purposes and has been adopted by medical and health professionals in numerous settings. The process can act as an initial stage of research to identify key items/opinions around a topic of interest (50,51). Prior to conducting the Delphi surveys, a literature review should be carried out to compile available working knowledge.

The Delphi exercise consists of a series of surveys that are administered in rounds, with each round presenting a new survey that has been adapted based on results from the previous round. Typically, results are anonymous, and summarized after each Delphi round is completed. The summarized results are sent back to participants and used to facilitate the next round, which consists of a similar survey from the prior round with slight modifications based on feedback and comments provided. The process is repeated, to slowly reach agreement on answers to survey questions. The entire process is iterative and continues until opinions begin to align and/or agreement is achieved, such as a high (or low) median score on a Likert scale rating system (49). When only minimal changes in answers are noted between rounds, consensus has been reached and the process stops; alternatively, a pre-determined criterion, such as a set number of rounds can be achieved to end the Delphi process. The incorporation of summary statistics provides each participant with a perspective on other participants' opinions. The analysis of group statistics encourages participants to not only consider their own perspective, but also the views of others, in order to then re-evaluate an answer to the same question, building towards consensus.

The use of a Delphi exercise has many advantages. It allows multiple people to participate without physical barriers. As everything is conducted online, participants can be in different countries or even continents, and can complete the survey rounds when convenient for them. The procedure is cost- and time-effective. Additionally, the online, anonymous format provides participants with equal opportunity to voice opinions and helps prevent a single, particularly compelling or powerful voice from determining the end results.

Furthermore, outcome measures in rheumatology (OMERACT) is an organization of rheumatologists, epidemiologists and biostatisticians whose work focuses on improving measurements for rheumatic diseases (52). OMERACT has published a suggested checklist for

Delphi processes used in determining core sets for rheumatic diseases as well as the OMERACT filter for developing core outcome measurement sets. The checklist proposed includes the use of clinicians and patients, asking open questions in initial phases and minimising attrition throughout the process (53). The incorporation of this checklist when conducting a Delphi can be used to strengthen this method for achieving consensus. The filter further suggests that truth (validity), discrimination and feasibility all be considered in such development stages and should be followed when developing outcome measurement sets for research in rheumatology (54,55).

2.6.2 Measuring disagreement

As a part of the RAM process, methods for calculating disagreement have been developed and tested (49). A typical measurement of disagreement uses the Interpercentile range (IPR). Through observation of limitations presented when using IPR to measure the spread of votes in panel-like scale rating exercises, the IPRAS method was developed by RAM. This method uses the interpercentile range adjusted for symmetry (IPRAS), as opposed to the IPR alone, and further allows for increased sensitivity to symmetric rating systems. The IPRAS smooths the rating scale for values between 6-7 and 3-4, creating a better measure of dispersion. Additionally, the IPR is centered at 5 on a 1-9 scale, whereas the IPRAS is centered more proportionally to the ratings obtained, creating a better measure of dispersion for each case used in relation to the ratings presented. The formula for IPRAS is as follows:

$$\text{IPRAS} = 2.35 + [\text{Asymmetry Index} \times 1.5]$$

where the asymmetry index is the difference between the central point on the rating scale used (such as 5 on a 1-9 Likert scale) and the central point of the IPR. The interpercentile range required for disagreement when perfect symmetry exists is a set value of 2.35. The correction factor for asymmetry is also a pre-determined set value of 1.5. All of these factors make up the IPRAS equation for calculating disagreement.

When the IPRAS for an individual rating is smaller than the IPR, disagreement exists. In testing this method, IPRAS received a sensitivity rating of one and noted ‘good’ specificity although no numerical value is provided. Testing of the method was noted by RAND to have occurred in six data sets with well over 5,000 variables rated. It proved to have several advantages, including providing a better measure of dispersion, and is thus beneficial when ratings on a scale are scattered. Since development, the IPRAS method for disagreement is

documented to have been tested in over 16,000 theoretical cases and over 6,500 real cases with high success rates and is thus a well-developed method for calculating consensus (49).

2.6.3 Nominal Group Technique (NGT) meeting

Similar to the Delphi exercise, a nominal group technique (NGT) meeting is intended to build consensus on a specific issue/topic. Although the aims of the Delphi exercise and NGT meeting are similar, the approaches to achieving consensus are different. An NGT meeting consists of an in-person gathering where experts share opinions and thoughts. Through structured discussion led by a moderator, agreement is achieved in a relatively short timeframe, normally within 1-2 hours. Typically, the moderator will pose a specific question or topic and each participant is directed to write down or discuss his/her opinion. The moderator will ensure inclusiveness of each participant throughout the process. Discussion occurs for each question/topic to clarify any ambiguities or provide additional feedback. Finally, after all questions are posed, answered and discussed, voting on each idea takes place (47,56). To finalize results, a majority is required. To help avoid a split vote there may be ranking or rating of items from most important to least important (47,48,56); alternatively, rewording and revisiting of items can occur until the vote is no longer split (57).

The structured NGT meeting has numerous strengths. It enables discussion and feedback to occur simultaneously, leading to better decision-making (48). Time constraints are minimal and it is both efficient and cost-effective (56). There are also disadvantages to this process (47). Although discussion is encouraged, due to the procedure and time constraints, opinions may not be shared to the fullest extent, and some participants' opinions may be overshadowed by the opinions of the more senior, more powerful, or more articulate members of the group.

2.7 Summary

The absence of classification criteria is an important challenge for research on SRC. The overall goal of the SCTC Scleroderma Renal Crisis Working Group is to develop classification criteria for SRC. To date, the Working Group has completed a scoping review of the literature. The detailed results of the scoping review can be viewed in Appendix 1 and in the supporting literature by Hoa et al. 2017 (16). This thesis project builds on the scoping review, and conducts a three-round Delphi exercise, followed by a NGT meeting.

CHAPTER 3. Objectives:

For this thesis project, 'domain' is an umbrella term referring to a distinct area of focus of an individual's health. Within each domain, 'items' are the specific indicators that measure and assess these health areas. The ultimate aim is to generate a concise list of domains and items that will be used to develop classification criteria for SRC in future research. The classification criteria will be used to facilitate research, including both clinical trials and observational studies. Such research will, in turn, provide further insight into SRC diagnosis, treatment and prevention.

We aim to create items to be considered for classification criteria specifically. Classification criteria are developed for clinical research purposes. They are developed to encompass a variety of individuals with a specific disease, however, they are not broad enough to include everyone. Diagnostic criteria differ in that they tend to be much broader, typically consist of associated signs and symptoms and are mainly used for patient care (6). The items to be considered for SRC criteria will be used for future research purposes and are thus classification criteria.

CHAPTER 4. Method Overview:

4.1 Overview

Although various definitions of SRC have been proposed, none has been developed and validated using robust consensus and evidence-based methodology. The aim of this project was to create a core set of items to develop classification criteria for scleroderma renal crisis (SRC) using consensus methodology. This was done in two phases, in which consensus was 1) initially achieved on a preliminary list of items using a Delphi exercise and then 2) further achieved through refinement of the list of items in a structured NGT meeting. The Delphi exercise consisted of three rounds. The surveys used in Round 1 and Round 2 can be found in Appendix 2 and Appendix 3, respectively. This project represents part of a larger program of research. Previously, a scoping review was conducted to inform the development of this project. Future phases of work will occur following this project to further develop the core set of items to produce classification criteria for SRC. The overall process involved in the development of SRC classification criteria can be seen in Figure 1.

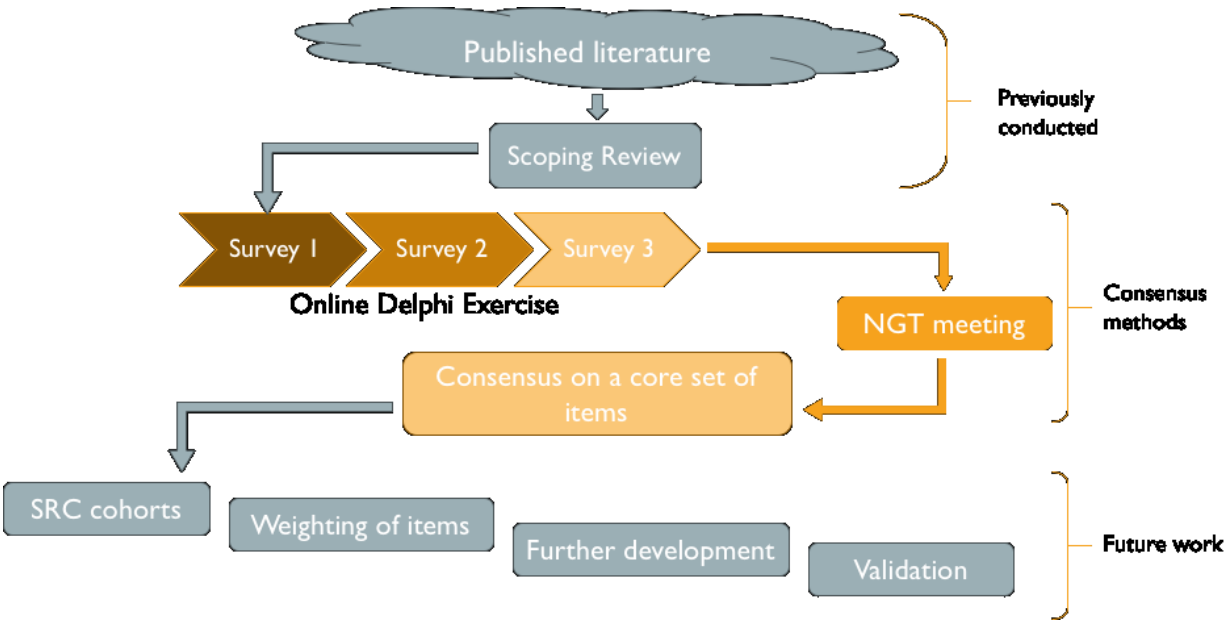

Figure 1. Flow chart of the overall process for the development of classification criteria of SRC

4.2 Role of the thesis author

As part of my personal contribution to this research project, I was responsible for the entire Delphi process outlined below. Within the process, I developed the surveys, contacted the participants, analyzed and drafted the results. These results were then distributed to participants of the NGT meeting. I was not present at the meeting and therefore my supervisor, Dr. Hudson, was responsible for delivering the Delphi results and writing up the results from the NGT. Nevertheless, the NGT meeting was audio recorded and I was able to listen and become completely familiar with that part of the project. Additionally, I was a part of all communication through email correspondence with NGT participants, both prior to and following the NGT. I was then responsible for analyzing participant characteristics and for further documenting and summarizing all NGT results.

4.3 Ethics

Ethics approval for this project was obtained from the Jewish General Hospital Research Ethics Board, Montreal, Quebec, Canada (Ethics Protocol # CODIM-MBM-17-104). Prior to the start of Round 1 of the Delphi exercise, all participants provided informed consent.

CHAPTER 5. Manuscript:

GENERATION OF A CORE SET OF ITEMS TO DEVELOP CLASSIFICATION CRITERIA FOR SCLERODERMA RENAL CRISIS USING CONSENSUS METHODOLOGY*

Emily-Ann Butler, BSc (Hons)[1]; Murray Baron, MD[2,3]; Agnes B. Fogo, MD[4]; Tracy Frech, MD MS[5]; Cybele Ghossein, MD[6]; Eric Hachulla, MD[7]; Sabrina Hoa, MD[3,8]; Sindhu R. Johnson, MD PhD[9]; Dinesh Khanna, MD MS[10]; Luc Mouthon, MD PhD[11]; Mandana Nikpour, MD PhD[12]; Susanna Proudman, MB BS (Hons)[13]; Virginia Steen, MD[14]; Edward Stern, MD[15]; John Varga, MD[6]; Christopher Denton, PhD FRCP[15]; Marie Hudson, MD MPH[2,3,8]

Institutional affiliations: [1]Dalhousie University, Halifax, Canada; [2]Jewish General Hospital, Montreal, Canada; [3]McGill University, Montreal, Canada; [4]Vanderbilt University Medical Center, Nashville, USA; [5]University of Utah, Salt Lake City, USA; [6]Northwestern University, Chicago, USA; [7]University of Lille - Hospital Claude Huriez, Lille, France; [8]Lady Davis Institute, Montreal, Canada; [9]University of Toronto, Toronto, Canada; [10]University of Michigan, Ann Arbor, USA; [11]Department of Internal Medicine, Cochin Hospital, Paris-Descartes University, Paris, France; [12]University of Melbourne at St. Vincent's Hospital, Victoria, Australia; [13]Royal Adelaide Hospital and University of Adelaide, Adelaide, Australia; [14]Georgetown University, Washington DC, USA; [15]Royal Free Hospital, London, UK .

Funding: This study was funded by a grant from the Scleroderma Clinical Trials Consortium. Dr Hudson is funded as a Clinician Scientist by the Fonds de Recherche du Quebec – Sante. Dr Stern was supported by MRC grant MR/K015230/1. The funding sources had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

---

* Butler, E. , Baron, M. , Fogo, A. B., Frech, T. , Ghossein, C. , Hachulla, E. , Hoa, S. , Johnson, S. R., Khanna, D. , Mouthon, L. , Nikpour, M. , Proudman, S. , Steen, V. , Stern, E. , Varga, J. , Denton, C. , Hudson, M., Scleroderma Clinical Trials Consortium Scleroderma Renal Crisis Working Group , (2019), *Generation of a core set of items to develop classification criteria for Scleroderma renal crisis using consensus methodology*. Arthritis Rheumatol. Accepted Author Manuscript. doi:10.1002/art.40809.

Abstract

Objective: This project was undertaken to generate a core set of items to develop classification criteria for scleroderma renal crisis (SRC) using consensus methodology.

Methods: An international, multidisciplinary panel of experts was invited to participate in a 3-round Delphi exercise developed based on items identified by a scoping review. In Round 1, participants were asked to identify omissions and clarify ambiguities regarding the items in the survey. In Round 2, participants were asked to rate the validity and feasibility of the items using Likert-type scales ranging from 1-9 (1= very invalid/unfeasible, 5 = uncertain, 9 = very valid/feasible). In Round 3, participants reviewed the results and comments of Round 2, and were asked to provide final ratings. Items rated as highly valid and feasible (both median scores ≥7) in Round 3 were selected as the provisional core set of items. A consensus meeting using nominal group technique (NGT) followed to further reduce the core set of items.

Results: Ninety-nine experts from 16 countries participated in the Delphi exercise. Of the 31 items in the survey, consensus was achieved on 13, including hypertension, renal insufficiency, proteinuria and hemolysis. Eleven experts took part in the NGT discussion, where consensus was achieved in 5 domains: blood pressure, acute kidney injury, microangiopathic hemolytic anemia, target organ dysfunction, and renal histopathology.

Conclusions: A core set of items that characterize SRC was identified using consensus methodology. This core set will be used in future data-driven phases of this project to develop classification criteria for SRC.

Scleroderma renal crisis (SRC) is a life-threatening complication of systemic sclerosis (SSc) (13,14,33,34). It is usually characterized by malignant hypertension and acute kidney injury (13). However, the clinical spectrum of SRC is broad, ranging from full-blown disease presenting as new onset accelerated arterial hypertension and rapidly progressive oliguric renal failure, to more modest elevations in blood pressure and renal dysfunction, and, more rarely, normotensive presentations. On the other hand, hypertension without uraemia, urinary abnormalities and/or mild uraemia attributable to other factors (e.g., concomitant comorbidities such as diabetes or exposure to nephrotoxic medications) are common in SSc (14,15). These conditions should not be confused with SRC.

Scleroderma renal crisis is relatively rare, occurring in about 5% of all SSc patients (13). It is more common in patients with rapidly progressing diffuse cutaneous SSc (dcSSc) (11%) as compared to patients with limited cutaneous SSc (lcSSc) (4%) (17). SRC can be further sub-categorized into hypertensive or normotensive forms, representing approximately 90% and 10% of SRC cases, respectively (29,30). Historically, SRC was the leading cause of death in SSc (1). However, with the advent of angiotensin converting enzyme (ACE) inhibitors, mortality rates have decreased significantly (2,3). Nevertheless, one-year outcomes remain poor, with over 30% mortality and 25% of patients remaining dialysis-dependent (5). There is an urgent need to undertake research to identify novel treatments and to improve SRC outcomes.

In addition to heterogeneity and rarity, the absence of a consensus classification criteria is an important challenge for research on SRC. To date, most studies of SRC have used *ad hoc* criteria that have varied considerably from study to study. In a scoping review of the literature, 40 original definitions of SRC, with significant heterogeneity among them, were identified (16). Only one study to date has partially validated criteria for SRC (5).

The Scleroderma Clinical Trials Consortium (SCTC) SRC Working Group was created to develop consensus classification criteria for SRC. The objective of this phase of the study was to generate a core set of domains with corresponding items to classify SRC using consensus methodology. Future studies will be required to develop and validate classification criteria for SRC.

Methods

A scoping review of the literature to identify domains and corresponding items used to classify SRC has been published (16). The results of this review were used to inform this project, which consisted of two phases: 1) a modified online Delphi survey to develop provisional consensus on a core set of domains with corresponding items to classify SRC and 2) a consensus meeting using nominal group technique (NGT) to further reduce the core set. Ethics approval for this project was obtained from the Jewish General Hospital Research Ethics Board, Montreal, Quebec, Canada (Protocol # CODIM-MBM-17-104).

Phase 1: Delphi

To develop initial consensus, a modified, online, three-round Delphi survey was conducted (58,59). We identified two hundred and sixteen experts identified through the SCTC, European Scleroderma Trials and Research Group (EUSTAR), Canadian Scleroderma Research Group (CSRG) and Australian Scleroderma Interest Group (ASIG) and we sent a letter of invitation via email to participate. In addition, pathologists and nephrologists known through these organizations with interest in SRC were invited to participate to provide additional perspective on key items pertaining to SRC and are included in the 216 expert count provided.

All individuals interested in participating in the online Delphi survey were asked to explicitly accept the invitation by return email. All individuals who accepted were then considered study participants, and thereby constituted the denominator for the participation rates.

The online Delphi survey was developed and managed through the REDCap platform (Vanderbilt University, Nashville, Tennessee). The survey consisted of 31 items identified by the scoping review, grouped in 11 domains: hypertension; renal insufficiency; proteinuria; hematuria; thrombocytopenia; hemolysis; encephalopathy; retinopathy; hyper-reninemia; cardiac dysfunction; and abnormal kidney biopsy.

The Delphi survey consisted of three rounds. At the start of Round 1, consent to participate was obtained and contact, demographic and personal information was collected for all participants. Subsequently, Round 1 asked participants to consider the domains and corresponding items identified in the scoping review and requested them to clarify ambiguities, identify omissions and to provide comments. Items were modified, re-worded and re-organized according to the feedback from Round 1.

In Round 2, participants were asked to rate the validity and feasibility of the items using Likert-type scales ranging from 1-9 (1= very invalid/unfeasible, 5 = uncertain, 9 = very valid/feasible) and to provide comments. Participants were provided links to full-text copies of the scoping review and all of the papers included therein addressing studies providing SRC definitions or classification criteria, totaling 24 total papers. Scientific validity was defined as items supported by published literature on SRC and empirical validity was defined by personal experience and knowledge of SRC content. Feasibility was defined in terms of whether the item could be performed/tested in an easy or convenient matter. In addition, specific questions to identify cut-offs or clinical values were included, using multiple-choice question format. These questions pertained to blood pressure, serum creatinine, proteinuria, hematuria and thrombocytopenia.

In Round 3, the results of Round 2 were presented using summary statistics, including medians and interquartile ranges, and bar graphs. Participants were also shown their answers and anonymized comments from other participants in Round 2. After reviewing the results of Round 2, participants were then asked to provide their final rating on scientific validity, empirical validity and feasibility of the items.

Participants were informed of the timeline for the Delphi survey and given 2 weeks to complete the first round. Upon completion of Round 1, participants were prompted with a reminder of the upcoming rounds. After closing Round 1, results were analyzed and the survey modified accordingly during a 2-week period. If an individual had agreed to participate, but did not complete Round 1 in the allotted time, they were still allowed to participate in Rounds 2 and 3, as the first round primarily gathered input and comments for a more structured second and third round. However, given the links between Rounds 2 and 3, only those who participated in Round 2 were presented with their answers. If an individual did not complete Round 2 in the allotted time, they were only provided with group summary statistics and comments in Round 3.

Consensus was defined as items rated highly scientifically valid and feasible (both median scores $\geq$7) in Round 3, and for which there was no disagreement, calculated using the RAND/UCLA Appropriateness Method formula. Disagreement existed when the inter-percentile range (IPR: difference between the 30^{th} and 70^{th} percentiles) was larger than the IPR adjusted for symmetry (IPRAS), calculated as follows:

$$\text{IPRAS} = 2.35 + [\text{Asymmetry Index} \times 1.5]$$

Derivation of the formula is shown in the RAND/UCLA Appropriateness Method handbook (49).

Phase 2: NGT meeting

The second phase of this study was conducted to further reduce the number of items and achieve consensus using NGT (60). International experts, including rheumatologists, internists and nephrologists, were invited to participate in a 2-hour face-to-face meeting held in November 2017 in San Diego (California, USA). Dr. Dinesh Khanna moderated the discussion based on expertise and previous experience in the fields of SRC and NGT techniques (60,61).

For the purposes of the NGT meeting, the 11 domains from the Delphi survey were re-organized and collapsed into five core domains (hypertension, renal dysfunction, microangiopathic hemolytic anemia with thrombocytopenia, target organ dysfunction [encephalopathy, retinopathy and cardiac dysfunction] and renal histopathology). Each domain was discussed in turn with each panelist invited to provide comments. At the end of the discussion, the panelists were asked to vote by a show of hands if the items corresponding to the core domains should be included. A simple majority was required to include the item.

During the NGT meeting, it became clear that some items required content expertise beyond rheumatology, internal medicine and nephrology. Thus, some items were conditionally included, pending further review with content experts. Experts in hematology, neurology, ophthalmology, and cardiology were then contacted and asked to provide input and published evidence to define items in those domains.

A final list of core domains and corresponding items (and their definitions and/or descriptions) was compiled and circulated among the participants of the NGT meeting for final approval.

Secondary objectives of the NGT were to define a list of diseases with similar clinical presentations to SRC (to improve the specificity of the criteria) and to discuss how the classification criteria for hypertensive and normotensive SRC should be different. Although the former was achieved, the panel decided that distinction between hypertensive and normotensive SRC should be based on data collected in future phases of this project.

Results

Phase 1: Delphi

We contacted 216 professionals with an interest in SRC, of whom 99 agreed to participate in the modified online Delphi survey. Of those, 77 (78%), 60 (61%) and 69 (70%) participated in Rounds 1, 2 and 3, respectively, and 49 (49%) completed all three rounds of the survey.

Participant characteristics are shown in Table 2 and the geographical distribution of those participants in Table 3. Participants were mainly rheumatologists (86%) with some internists, nephrologists and pathologists. Most participants worked as clinicians for >11 years, with only a few having less than 10 years of experience (13%). The majority of participants were from the United States (35%) followed by Canada (11%); 16 other countries were also represented.

The Delphi survey consisted of three rounds in which Round 1 allowed participants to provide feedback on the content of the survey, Round 2 allowed participants to rate items for validity and feasibility, in addition to providing optional comments, and Round 3 allowed participants to review their own and the group's ratings from Round 2 and to provide final ratings for validity and feasibility. A total of 31 items in 11 domains were included in the Delphi survey. The 11 domains included: hypertension; renal insufficiency; proteinuria; hematuria; thrombocytopenia; hemolysis; encephalopathy; retinopathy; hyper-reninemia; cardiac dysfunction; and abnormal kidney biopsy. Of these, 13 items in four domains (five items in hypertension, two in renal insufficiency, one in proteinuria and five items in hemolysis) achieved consensus in Round 3 (median ratings $\geq$ 7 on validity and feasibility with no disagreement). Disagreement on feasibility, calculated with the IPRAS formula, was only present for hyper-reninemia. In any case, that item had not achieved consensus on feasibility either. Of note, all items that reached consensus in Round 2, also reached consensus in Round 3 with no additional items reaching consensus in Round 3. However, the IQR for the majority of items became smaller in Round 3, demonstrating growing consensus. The median ratings and IQR for each item for Rounds 2 and 3 are presented in Table 4.

After completion of the Delphi survey, only scientific validity and feasibility (not empirical validity) were used in calculating consensus. This slight modification allowed for the inclusion of one additional item; reticulocyte count above normal range for local laboratory under the category of hemolysis. This approach was used in an effort to be as inclusive as

possible and to enhance content validity, by producing ratings based on literature and research to date, while minimizing personal opinion and bias.

In addition to the rating of items, questions pertaining to cut-offs for blood pressure, creatinine, proteinuria, hematuria and thrombocytopenia were included in Rounds 2 and 3 (Table 5). Under hypertension, six questions pertaining to blood pressure cut-offs, increases in SBP and DBP as well as the frequency and timing of blood pressure measurements were addressed. Two questions addressing serum creatinine level increases for renal insufficiency were posed. Four questions under proteinuria addressed dipstick measurements and urine: protein ratios. Similarly, four questions for hematuria addressing dipstick levels and RBC counts were also addressed. Finally, one question regarding platelet count for thrombocytopenia was included in the Delphi survey. All questions were duplicated in Rounds 2 and 3. The results showed considerable variability, emphasizing the need to identify uniform cut-offs supported by evidence.

Phase 2: NGT meeting

Seventeen international experts, including rheumatologists, internists and nephrologists, were invited to participate in the face-to-face NGT meeting. Six were not available. Thus, the panel consisted of 11 participants, 10 rheumatologists and one nephrologist, from the USA, Canada, United Kingdom, France, Netherlands and Australia. All but one of the NGT participants were also participants in the prior Delphi survey. Prior to the NGT meeting, the 11 domains from the Delphi survey were re-organized into five domains (hypertension, renal dysfunction [renal insufficiency, proteinuria, hematuria and hyper-reninemia], microangiopathic hemolytic anemia with thrombocytopenia, target organ dysfunction [encephalopathy, retinopathy and cardiac dysfunction] and renal histopathology). Prior to and at the meeting, it was agreed that items should be defined as much as possible according to evidence and/or international guidelines.

After discussion, the participants at the NGT agreed that hypertension should be re-worded as *Rise in blood pressure* and defined according to international guidelines using cut-offs of 140 mmHg for systolic blood pressure and 90 mmHg for diastolic blood pressure (62–64). Since “rise in blood pressure” is a concept that is intrinsic to SRC and is meant to include patients with blood pressure within normal ranges but with clinically significant rise over baseline and for which there are no established guidelines, cut-offs of 30 mmHg above normal for rise in systolic blood pressure and 20 mmHg above normal for rise in diastolic blood pressure were retained based on the consensus in the Delphi exercise (Table 5).

Similarly, the participants at the NGT agreed that renal dysfunction should be re-worded as *Acute Kidney Injury* and defined according to the Kidney Disease Improving Global Outcomes (KDIGO) guidelines (37). These guidelines define acute kidney injury as follows: increase in serum creatinine by > 26.5 umol/L (> 0.3 mg/dl) within 48 hours; increase in serum creatinine to >1.5 times baseline, which is known or presumed to have occurred within the prior seven days; and urine volume < 0.5 ml/kg/h for six hours.

The panel discussed *Microangiopathic hemolytic anemia and thrombocytopenia* and *Target organ dysfunction (encephalopathy, retinopathy, cardiomyopathy)*. It was agreed that these domains could be retained in the core set but that specific item definitions should be finalized after consulting with content experts in hematology, neurology, ophthalmology, and cardiology. Following these consultations, the items were defined as follows:

Microangiopathic hemolytic anemia and thrombocytopenia (MAHAT) was defined as new or worsening anemia not due to other causes, schistocytes or other RBC fragments on blood smear, laboratory evidence of hemolysis that includes elevated lactate dehydrogenase (LDH) and reticulocytes and/or low/absent haptoglobin and a negative direct anti-globulin test. Thrombocytopenia was defined as a platelet count of $\leq$ 100,000 confirmed by blood smear (38,39). There was discussion about including a specific cut-off for schistocytes, such as >1% (2,65) or > 2 per high powered field (66). However, this was not retained because automated quantification is not widely available, manual quantification is subjective and neither of these cut-offs have been validated.

Encephalopathy was defined as headache, altered mental status, seizures, visual disturbances and/or other focal or diffuse neurologic signs not attributable to other cause. In the absence of an evidence-based definition of hypertensive encephalopathy, the definition proposed by Lamy and Mas (67) was felt to describe the syndrome best and was retained.

Retinopathy was defined as hemorrhages, hard and soft (cotton wool) exudates, and/or disc edema, not attributable to other causes and confirmed by an ophthalmologist. This definition was based on key items in the Keith-Wagener-Baker and Modified Scheie classification criteria (68,69), and required confirmation by an ophthalmologists because it has been shown that the reliability of these criteria is low when ophthalmoscopic exam is performed by other physicians (69).

Cardiomyopathy was divided into *Acute heart failure* and *Acute pericarditis. Acute heart failure* is a syndrome and its definition was based on the US and Canadian guidelines for the management

of heart failure (70–72). It is characterized by typical symptoms including breathlessness, ankle swelling and fatigue that may be accompanied by signs such as elevated jugular venous pressure, pulmonary crackles and peripheral edema. *Acute pericarditis* was defined according to the 2015 European Society of Cardiology Guidelines for the diagnosis and management of pericardial diseases. It is diagnosed with at least two of the four following criteria: 1) chest pain due to pericarditis; 2) pericardial rub; 3) new widespread ST-elevation or PR depression on electrocardiogram; 4) pericardial effusion (new or worsening) on cardiac echocardiography (72).

A detailed description of the renal histopathological changes in SRC was prepared by an experienced pathologist and can be found in Table 6 (73).

The final core set of items (and definitions) to develop classification criteria for SRC is presented in Table 6. It was approved by the participants at the NGT. After the NGT and consultation with content experts, some items that reached consensus in the Delphi exercise were not retained in the core set. The domain of renal insufficiency was discussed and agreed to be replaced with kidney injury to meet KDIGO guidelines and definition for AKI (37). This resulted in the removal of the corresponding item of serum creatinine $\geq$ 120% (or 1.2 times) the upper limit of normal for local laboratory as this is not part of KDIGO guidelines. Proteinuria was discarded after NGT discussion as low-level proteinuria was believed to be too common, dipstick urine protein to creatinine ratio was not reliable. Additionally, when turning to KDIGO guidelines, proteinuria is not included as part of AKI definitions. Other items that did not achieve consensus in the Delphi exercise (e.g. thrombocytopenia < 100,000 platelets/mm3 and elevated serum lactate dehydrogenase, as part of the definition for microangiopathic hemolytic anemia) were included in the final core set. In an effort to be as inclusive as possible in the core set of domains with corresponding items, domains and items that did not reach consensus during the Delphi were retained after NGT discussion. Although hemolysis as a domain had consensus on all but one item during the Delphi, all items were retained, specifically serum LDH and/or indirect bilirubin above normal ranges, as it was agreed that MAHAT guidelines were agreed to be followed – thus, modifications to meet these guidelines resulted in item retention (38,39). Additionally, thrombocytopenia was retained to meet MAHAT guidelines. The domains of retinopathy, encephalopathy and cardiac dysfunction with all respective items did not reach consensus during the Delphi but were retained during the NGT meeting in an effort to defer to neurologists and cardiologists to provide supportive evidence. Finally, abnormal kidney biopsy

was also retained and modified to include histopathology to meet the expert definition proposed by Agnes Fogo (Vanderbilt) for inclusion in order to once again be as inclusive as possible. The final core set was then distributed to all participants of the NGT meeting and the Delphi exercise for final approval.

Finally, as a secondary objective of the NGT, a list of SRC mimickers was compiled and approved by the panel (Table 7). Indeed, kidney injury in SSc is not always due to SRC and mimickers can also occur in SSc. In addition, mimickers of SRC may also share other clinical features with SRC, such as hypertension and MAHAT, and renal histopathology may overlap (16,74,75). Excluding patients with these conditions will improve the specificity of the future classification criteria (76).

Discussion

In this study, we generated a core set of items to classify SRC using consensus methodology. This core set includes five domains and 13 items. The definitions for each item were evidence-based or, in the absence of evidence, determined in consultation with content experts.

The progress made to date to develop classification criteria for SRC demonstrates the importance of using the best evidence available. A scoping review of the literature identified 40 heterogeneous definitions of SRC using more than 40 items with variable definitions (16). The Delphi exercise led to consensus on 13 of these items. However, the need to go beyond consensus in the rheumatology community and to get the input of content experts emerged as a critical factor at the NGT meeting. Thus, the input from content experts was sought to finalize the core set. Proteinuria is a perfect example of how this approach allowed the core set to evolve. Indeed, low-level proteinuria is common in SSc (14), dipstick and urine protein-to-creatinine ratio are not reliable in AKI, proteinuria is not part the KDIGO definition of AKI, and proteinuria would compromise specificity of SRC criteria. Thus, despite the fact that there was consensus to include proteinuria in the core set after the Delphi exercise, this item was excluded after the NGT meeting and discussion with nephrologists.

A core set of variables to define SRC was proposed by experts in 2003 (29). It included items for systolic and diastolic blood pressure, serum creatinine, proteinuria, hematuria, microangiopathic hemolytic anemia and renal histopathology. These are known as the ANCONA criteria for SRC. Our core set has similarities to the ANCONA criteria in particular with respect

to blood pressure. However, there are also notable differences in defining acute kidney injury (including the exclusion of proteinuria and hematuria). In addition, our core set includes target organ dysfunction and a detailed histopathological description of SRC.

In 2016, the UK Scleroderma Study Group proposed criteria for the diagnosis of SRC (77). The criteria were divided into categories: diagnostic criteria (essential) and supportive evidence (desirable) with blood pressure and AKI as the former, MAHAT, hypertensive retinopathy, hematuria, oliguria or anuria, renal biopsy consistent with SRC features and flash pulmonary edema as the latter. Discrepancies with our proposed criteria are found in the slightly modified cut-off values for blood pressure (150/85 mmHg versus 140/90 mmHg) and additionally, there is no noted rise in diastolic blood pressure, only $\geq$20 mmHg for systolic blood pressure which is lower than $\geq$30 mmHg proposed in this study. Further, the UK criteria included hematuria. Additionally, oliguria and flash pulmonary edema were proposed as stand-alone items whereas in our list, these items are grouped into the AKI and acute heart failure definitions, respectively. Our core set provides a more in-depth detailed definition for each item, specifically for AKI, MAHAT and renal histopathology.

Only one study to date has attempted to validate the ANCONA criteria and another slightly different set of criteria for SRC that included encephalopathy (5). In that study, a diagnosis of SRC confirmed by a study physician was used as the gold standard for SRC. Compared to the gold standard, the two sets of criteria identified 70/70 subjects with hypertensive, but only 2/5 subjects with normotensive SRC. We believe that our core set, which was developed using robust consensus methodology and evidence-based content, represents a significant advancement over these definitions. In addition, it defines target organ involvement and provides a detailed histopathological description to define the term “findings consistent with SRC”.

This study has some limitations. First, only 99/216 experts invited to participate accepted and 77 (78%), 60 (61%) and 69 (70%) of these participated in Rounds 1-3 of the Delphi, respectively. We cannot exclude some response bias. Part of the reason for the low response rates may have been that the Delphi exercise was conducted during the summer and early fall in the Northern hemisphere. Numerous out of office replies were returned. On the other hand, to mitigate this source of bias, reminder emails were sent to optimize participation rates and the final sample was still substantial and representative. Second, there are large gaps in knowledge on SRC. Hence, participants in the Delphi may have rated validity based more on empirical, rather than on scientific

evidence. Nevertheless, we provided the Delphi participants with the scoping review and all of the original papers included therein in every Round for easy access to the available literature. Third, recruitment of participants with a broad range of expertise is critical to the success of a consensus-building exercise. Although there were a few specialists other than rheumatologists who participated in the Delphi, it became clear at the NGT meeting that content expertise in hematology, neurology, ophthalmology, and cardiology was lacking. We therefore recruited experts in all of these fields to help finalize the relevant items.

This study has substantial strengths. The emphasis on evidence and input from content experts ensured that the final core set had face and content validity (78). The geographic range of participants contributed to the generalizability of the results. There was important complementarity in the use of both a Delphi exercise and a semi-structured NGT consensus meeting. The Delphi provided a cost-effective approach to survey a larger sample of international experts working anonymously. The NGT meeting allowed for a time-efficient, face-to-face discussion of a smaller sample of experts led by an experienced moderator.

Conclusion and future steps

In conclusion, using consensus methodology, we generated a core set of items, and the definition of those items, to be used in the development of classification criteria for SRC. To determine if and how these items should be incorporated into classification criteria for SRC, two future phases of this research project are now in planning. The first, modeled on the *International Scleroderma Renal Crisis Survey* (5), will be to recruit an inception SRC cohort and collect the items in the core set. A comparison cohort consisting of subjects with conditions that mimic SRC (Table 7) will also be assembled. These data will be used to develop and validate classification criteria for SRC. The second will be a forced choice study using multi-criteria decision analysis methods (79) to assign weights to the items in the criteria and to set probability values for definite, probable and possible SRC. The resulting classification criteria will facilitate rigorous research in SRC.

Acknowledgements

We would like to thank all the participants of the Delphi survey, the participants of the Nominal Group Meeting (Murray Baron, Mary Csuka, Jeska de vries Bouwstra, Christopher

Denton, Tracy Frech, Cybele Ghossein, Sindhu Johnson, Luc Mouthon, Mandana Nikpour, Richard Silver, Virginia Steen) and the content experts (hematology: Mark Blotstein, Stephen Caplan, Nathalie Johnson; cardiology: Vartan Mardigyan, Richard Sheppard; neurology: Robert Altman, Colin Chalk, Rami Massie; ophthalmology: Julius Gomolin).

Table 2. Characteristics of participants in the Delphi exercise

		N (%)
Specialty	Rheumatologist	61 (85.9)
	Nephrologist	2 (2.8)
	Pathologist	1 (1.4)
	Internist	5 (7.0)
	Other	2 (2.8)
Years as a clinician	1-10 years	9 (12.7)
	11-20 years	22 (31.0)
	21-30 years	24 (33.8)
	>30 years	16 (22.5)
Unique systemic sclerosis patients seen each year	1-30 patients	10 (14.1)
	31-60 patients	8 (11.3)
	61-100 patients	12 (16.9)
	>100 patients	41 (57.7)
New scleroderma renal crisis patients seen each year	0 patients	4 (5.6)
	1-2 patients	45 (63.4)
	3-5 patients	16 (22.5)
	>5 patients	6 (8.5)
Returning scleroderma renal crisis patients seen each year	0 patients	5 (7.0)
	1-5 patients	26 (36.6)
	6-10 patients	23 (32.4)
	11-15 patients	14 (19.7)
	>15 patients	3 (4.2)

Table 3. Geographical distribution of participants in the Delphi exercise

	N (%)
Argentina	1 (1.4)
Australia	6 (8.5)
Belgium	2 (2.8)
Canada	8 (11.3)
Denmark	1 (1.4)
France	3 (4.2)
Germany	2 (2.8)
Israel	1 (1.4)
Italy	5 (7.0)
Japan	3 (4.2)
Mexico	1 (1.4)
Netherlands	2 (2.8)
Spain	2 (2.8)
Switzerland	2 (2.8)
United Kingdom	6 (8.5)
United States of America	25 (35.2)

Table 4. Results from Rounds 2 and 3 of the Delphi exercise and consensus achieved after Round 3

Criteria Category	Question		Round 2		Round 3		Consensus
			Validity	Feasibility	Validity	Feasibility	
Hypertension	New onset or deterioration of pre-existing hypertension, defined as any of the following:	Systolic blood pressure ≥ 140 mmHg	7(2)*	8(2)	7(1)	8(1)	yes
		Diastolic blood pressure ≥ 90 mmHg	7(2)	8(1)	7(0.5)	8(1)	yes
		Rise in systolic blood pressure ≥ 30 mmHg	7(2)	8(1)	7(1)	8(1)	yes
		Rise in diastolic blood pressure ≥ 20 mmHg	7(2)	8(2)	7(1)	8(0)	yes
	Increase in both systolic and diastolic blood pressure should be present.		6(3)	8(2)	6(2)	8(0.5)	no
	In the absence of signs and symptoms, blood pressure measurements should be measured on at least 2 occasions.		7(3)	8(1)	7(1)	8(1)	yes
Renal insufficiency	Increase in serum creatinine ≥50% over baseline or, if no baseline available, serum creatinine ≥120% (or 1.2 times) the upper limit of normal for local laboratory (with measurement repeated if necessary to rule out lab error).		7(2)	8(2)	7(1)	8(1)	yes
Proteinuria	New proteinuria defined as ≥ 1+ (30-100 mg/dL range) by urine dipstick or worsening proteinuria defined as a ≥ 1 point increase in protein on urine (1+ to ≥ 2+, 2+ to ≥ 3+, etc).		5(2)	7(2)	5(1)	7(1)	no
	New proteinuria defined as ≥ 2+ (100-300 mg/dL range) by urine dipstick or worsening proteinuria defined as a ≥ 1 point increase in protein on urine (2+ to ≥ 3+, 3+ to ≥ 4+, etc).		7(2)	8(1)	7(1)	8(1)	yes
	Proteinuria should be confirmed by urine protein:creatinine ratio.		7(2)	8(2)	7(1)	8(0)	yes
	Proteinuria should be confirmed by 24-hour urine collection.		6(4)	6(3)	6(2)	6(2)	no
Hematuria	New hematuria defined as ≥ 1+ by urine dipstick or worsening hematuria defined as a a ≥ 1 point increase on urine dipstick (1+ to ≥ 2+, 2+ to ≥ 3+, etc).		6(3)	8(1)	6(1)	8(1)	no
	New hematuria defined as ≥ 2+ by urine dipstick or worsening hematuria defined as a ≥ 1 point increase on urine dipstick (2+ to ≥ 3+, 3+ to ≥ 4+, etc).		6(3)	8(1)	6(1)	8(1)	no
	New hematuria defined as ≥ 10 red blood cells per high powered field on urine microscopy or worsening hematuria defined as a doubling of baseline hematuria on urine microscopy.		6(2)	7(2)	6(2)	7(1)	no
Thrombocytopenia	≤ 100,000 platelets/mm3		6(3)	8(1)	6(1)	8(1)	no
	Thrombocytopenia should be confirmed by manual blood smear.		6(2)	6(2)	6(2)	6(1)	no
Hemolysis	Microangiopathic hemolytic anemia defined as new or worsening anemia not due to other causes and supported by the presence of one of the following:	Schistocytes or other red blood cell fragments on blood smear.	8(1)	8(1)	8(0)	8(0)	yes
		Reticulocyte count above normal range for local laboratory.	7(3)	7(1)	7(1)	7(1)	yes
		Serum lactate dehydrogenase and/or indirect bilirubin above normal ranges for local laboratory.	6(2)	8(2)	6(1)	8(1)	no
		Serum haptoglobin below normal range for local laboratory.	7(2)	8(2)	7(1)	8(1)	yes
	Microangiopathic hemolytic anemia defined as new or worsening anemia not due to other causes and supported by the presence of at least two lab abnormalities (red blood cell fragments, elevated reticulocyte count, elevated serum lactate dehydrogenase/indirect bilirubin, low haptoglobin).		8(1)	8(1)	8(0)	8(0)	yes
	A direct anti-globulin test should be documented to rule out autoimmune hemolytic anemia.		7(3)	7(2)	7(0)	7(1)	yes

* Median values (inter-quartile range)

Table 4. Results from Rounds 2 and 3 of the Delphi exercise and consensus achieved after Round 3 – Continued

Criteria Category	Question	Round 2		Round 3		Consensus
		Validity	Feasibility	Validity	Feasibility	
Encephalopathy	Encephalopathy defined by the American Academy of Neurology as follows: 'Any diffuse disease of the brain that alters brain function or structure. The hallmark of encephalopathy is an altered mental state. Depending on the type and severity of encephalopathy, common neurological symptoms are progressive loss of memory and cognitive ability, subtle personality changes, inability to concentrate, lethargy, and progressive loss of consciousness. Other neurological symptoms may include myoclonus (involuntary twitching of a muscle or group of muscles), nystagmus (rapid, involuntary eye movement), tremor, muscle atrophy and weakness, dementia, seizures, and loss of ability to swallow or speak'.	6(3)*	7(2)	6(1)	7(1)	no
Retinopathy	Retinopathy typical of malignant hypertension	7(2)	6(3)	7(1)	6(1)	no
	Grade III (flame-shaped hemorrhages and/or "cotton-wool" exudates) or IV (papilledema) retinopathy, according to Keith-Wagener classification	7(3)	6(3)	7(1)	6(2)	no
Hyperreninemia	Elevation of plasma renin activity ≥ 2 times the upper limit of normal	7(3)	4(4)	7(1)	5(2)	no
Cardiac dysfunction	Presence of flash pulmonary edema based on all available information and clinical judgement.	6(2)	7(2)	6(1)	7(0)	no
	Presence of symptomatic pericardial effusion based on all available information and clinical judgement.	6(2)	6(2)	6(1)	6(1)	no
Abnormal kidney biopsy	Findings consistent with scleroderma renal crisis (microangiopathy)	8(2)	6(4)	8(0)	6(2)	no
	Accumulation of mucoid (myxoid) in interlobular arteries (indistinguishable from accelerated hypertension) and/or fibrinoid necrosis of arteries	7(2)	6(4)	7(1)	6(2)	no
	Histopathological findings on kidney biopsy consistent with SRC may include the following: small vessel (arcuate and interlobular arteries) changes predominate over glomerular alterations. Early vascular abnormalities include intimal accumulation of myxoid material, thrombosis, fibrinoid necrosis, sometimes resulting in cortical necrosis. Narrowing and obliteration of the vascular lumen lead to glomerular ischemia. Juxtaglomerular apparatus hyperplasia, while relatively rare (10%), can be observed. Late changes are manifested by intimal thickening and proliferation (which lead to characteristic vascular "onion-skin" lesions), glomerulosclerosis and interstitial fibrosis. Since none of these findings are specific for scleroderma renal crisis, the pathological diagnosis must be supported by appropriate clinical and serological data.	8(2)	6(3)	8(0)	6(2)	no

* Median values (inter-quartile range)

Table 5. Results from the Delphi exercise for questions pertaining to cut-offs

Domain	Questions		Round 2	Round 3
Hypertension	What are the most appropriate cutoffs for high blood pressure? - Absolute SBP	140 mmHg	16*	13
		150 mmHg	**16**	**40**
		160 mmHg	9	7
		170 mmHg	1	0
		180 mmHg	1	0
		Other	2	0
	What are the most appropriate cutoffs for high blood pressure? - Absolute DBP	**90 mmHg**	**24**	**38**
		100 mmHg	18	21
		110 mmHg	1	1
		120 mmHg	0	0
		130 mmHg	0	0
		Other	2	0
	What are the most appropriate cutoffs for high blood pressure? - Increase in SBP	10 mmHg	0	0
		20 mmHg	11	5
		30 mmHg	**33**	**55**
		40 mmHg	1	0
		Other	0	0
	What are the most appropriate cutoffs for high blood pressure? - Increase in DBP	10 mmHg	6	3
		20 mmHg	**35**	**57**
		30 mmHg	4	0
		40 mmHg	0	0
		50 mmHg	0	0
		Other	0	0
	What are the most appropriate frequency and intervals for repeated measurements?	Only once is enough	1	1
		2 times	**30**	**51**
		3 times	13	8
		4 times	0	0
		Other	1	0
	What are the most appropriate frequency and intervals for repeated measurements?	**12 hours apart**	**29**	**45**
		24 hours apart	7	3
		48 hours apart	2	0
		72 hours apart	2	0
		1 week apart	0	0
		Other	5	12
Renal Insufficiency	What are the most appropriate cutoffs for increase in serum creatinine? - Increase above baseline	20%	2	0
		30%	7	7
		40%	7	6
		50%	**25**	**43**
		60%	1	1
		70%	0	1
		80%	0	0
		90%	0	0
		100% (doubling)	2	0
		Other	0	1
	What are the most appropriate cutoffs for increase in serum creatinine? - Increase above upper limit of local laboratory	**120%**	**21**	**41**
		130%	7	7
		140%	3	3
		150%	10	6
		175%	0	0
		200%	2	0
		Other	1	2

* Count of number of responses

Table 5. Results from the Delphi exercise for questions pertaining to cut-offs - Continued

Domain	Questions		Round 2	Round 3
Proteinuria	What are the most appropriate cutoffs for new proteinuria? - Dipstick	1+	3*	2
		2+	**40**	**56**
		3+	0	0
		4+	0	0
		Other	0	1
	What are the most appropriate cutoffs for new proteinuria? - urine protein:creatinine ratio	≥ 0.15 g/day †	3	2
		≥ 0.5 g/day	**28**	**57**
		≥ 1.0 g/day	10	0
		≥ 2.0 g/day	1	0
		Other	1	0
	What are the most appropriate cutoffs for worsening proteinuria? - Dipstick	a ≥ 1 point increase	18	6
		a ≥ 2 point increase	**25**	**51**
		Other	0	2
	What are the most appropriate cutoffs for worsening proteinuria? - urine protein:creatinine ratio	**Doubling**	**37**	**51**
		Tripling	4	1
		Quadrupling	0	0
		Other	2	6
Hematuria	What are the most appropriate cutoffs for new hematuria? - Dipstick	1+	4	3
		2+	**37**	**55**
		3+	2	0
		4+	0	0
		Other	0	1
	What are the most appropriate cutoffs for new hematuria? - Microscopy	**≥ 10 RBCs/HPF** §	**28**	**50**
		≥ 20 RBCs/HPF	9	6
		≥ 30 RBCs/HPF	4	0
		≥ 50 RBCs/HPF	1	1
		Other	1	2
	What are the most appropriate cutoffs for worsening hematuria? - Dipstick	a ≥ 1 point increase	20	8
		a ≥ 2 point increase	**22**	**48**
		Other	1	3
	What are the most appropriate cutoffs for worsening hematuria? - Microscopy	**doubling**	**34**	**50**
		tripling	7	2
		quadrupling	1	0
		Other	1	7
Thrombocytopenia	What is the most appropriate cutoff for thrombocytopenia? - Range from 50,000 to 140,000 platelets/mm3	50 000 platelets/mm^3	1	1
		60 000 platelets/mm^3	2	0
		70 000 platelets/mm^3	2	0
		80 000 platelets/mm^3	0	1
		90 000 platelets/mm^3	1	3
		100 000 platelets/mm^3	**29**	**47**
		110 000 platelets/mm^3	0	2
		120 000 platelets/mm^3	7	3
		130 000 platelets/mm^3	1	0
		140 000 platelets/mm^3	0	0
		Other	0	0

* Count of number of responses
† Grams per day
§ Red blood cell per high power field

Table 6. Final core set of items to develop classification criteria for SRC

Domain	Item	
Blood pressure	Acute rise in blood pressure defined as any of the following:	Systolic blood pressure ≥ 140 mmHg Diastolic blood pressure ≥ 90mmHg A rise in systolic blood pressure ≥ 30 mmHg above normal A rise in diastolic blood pressure ≥ 20 mmHg above normal
	Blood pressure measurement should be taken twice separated by at least 5 minutes. If blood pressure readings are discordant, repeat readings should be obtained until 2 consistent readings are obtained.	
Kidney injury	Acute kidney injury defined as any of the following:	Increase in serum creatinine by ≥ 26.5 umol/L (≥ 0.3 mg/dl) within 48 hours Increase in serum creatinine to ≥1.5 times baseline, which is known or presumed to have occurred within the prior 7 days Urine volume < 0.5 ml/kg/h for 6 hours
Microangiopathic hemolytic anemia and thrombocytopenia	New or worsening anemia not due to other causes. Schistocytes or other red blood cell fragments on blood smear. Thrombocytopenia ≤ 100,000, confirmed by manual smear. Laboratory evidence of hemolysis, including elevated lactate dehydrogenase, reticulocytosis and/or low/absent haptoglobin A negative direct anti-globulin test.	
Target organ dysfunction	*Hypertensive retinopathy* (hemorrhages, hard and soft (cotton wool) exudates, and/or disc edema, not attributable to other causes), confirmed by an ophthalmologist. *Hypertensive encephalopathy*, characterized by headache, altered mental status, seizures, visual disturbances and/or other focal or diffuse neurologic signs not attributable to other causes. *Acute heart failure*, characterized by typical symptoms (e.g. breathlessness, ankle swelling and fatigue) that may be accompanied by signs (e.g. elevated jugular venous pressure, pulmonary crackles and peripheral edema). *Acute pericarditis*, diagnosed with at least 2 of the 4 following criteria: 1) pericarditis chest pain; 2) pericardial rub; 3) new widespread ST-elevation or PR depression on electrocardiogram; 4) pericardial effusion (new or worsening) on cardiac echocardiography.	
Renal histopathology	Histopathological findings on kidney biopsy consistent with scleroderma renal crisis which may include the following: small vessel (arcuate and interlobular arteries) changes that predominate over glomerular alterations. Glomerular changes of thrombotic microangiopathy may be present, with acute changes including fibrin thrombi and endothelial swelling, red blood cell fragments and mesangiolysis, and chronic changes including double contours of the glomerular basement membrane. Nonspecific ischemic changes with corrugation of the glomerular basement membrane, and even segmental or global sclerosis of glomeruli may occur. Early vascular abnormalities include intimal accumulation of myxoid material, thrombosis, fibrinoid necrosis, fragmented red blood cells, sometimes resulting in cortical necrosis. Narrowing and obliteration of the vascular lumen lead to glomerular ischemia. Juxtaglomerular apparatus hyperplasia, while relatively rare (10%), can be observed. Late changes are manifested by intimal thickening and proliferation (which lead to characteristic vascular "onion-skin" lesions), glomerulosclerosis and interstitial fibrosis. Nonspecific tubular changes may also occur, including acute tubular injury in the early stage of injury, and later interstitial fibrosis and tubular atrophy. Since none of these findings are specific for SRC, the pathological diagnosis must be supported by appropriate clinical and serological data.	

Table 7. Scleroderma renal crisis mimickers and signs and symptoms that differentiate the mimickers

	Signs and symptoms
Pre-renal causes (e.g. volume depletion, sepsis)	Volume loss (vomiting, diarrhea, bleeding), fever, hypotension, low urinary fractional excretion of sodium and response to fluid repletion
Renal artery stenosis	Chronic hypertension, acute kidney injury unusual except after initiation of angiotensin converting enzyme inhibitor or angiotensin receptor blocker, patient with diffuse atherosclerosis, asymmetry in renal size, unilateral small kidney, recurrent episodes of flash pulmonary edema
Drugs affecting glomerular hemodynamics (e.g. non-steroidal anti-inflammatories, calcineurin inhibitors, angiotensin converting enzyme inhibitors, radiocontrast)	Documented drug exposures
Acute tubular necrosis (eg. renal ischemia, sepsis, and nephrotoxins)	Muddy brown granular casts, epithelial cell casts, and free renal tubular epithelial cells
ANCA*-associated glomerulonephritis	Distinct upper and lower airway features, microscopic hematuria, red blood cell casts and dysmorphic red cells on urinalysis
Other vasculitides (e.g. polyarteritis nodosa, cryoglobulinemia, anti-glomerular basement membrane antibody syndrome)	Rash, neuritis, nephritic sediment, pulmonary hemorrhage
Thrombotic thrombocytopenic purpura and other primary thrombotic microangiopathies	Fever, gastrointestinal symptoms, purpura, profound thrombocytopenia
Membranous nephropathy	Nephrotic syndrome, severe hypertension less common, acute kidney injury uncommon, hypoalbuminemia and hyperlipidemia, oval fat bodies, lipid droplets and fatty casts on urinalysis, microscopic hematuria without red blood cell casts possible
Membranoproliferative nephropathy	Nephritic syndrome, hypocomplementemia, monoclonal gammopathy
Oxalate nephropathy	Recurrent calcium stones, oxalate crystals in the urine sediment, patients at risk for calcium oxalate precipitation
Pre-eclampsia/eclampsia	May be difficult to distinguish pre-eclampsia/eclampsia in a pregnant woman with SSc, although renal function is usually normal in pre-eclampsia/eclampsia and elevated liver enzymes may orient the diagnosis towards the HELLP syndrome (hemolysis, elevated liver enzymes and low platelets)
Isolated renal abnormalities	5% of diffuse cutaneous SSc patients have unexplained renal abnormalities (14)

CHAPTER 6. Discussion:

In this discussion, we will review the final core set proposed in this thesis project and discuss its face and content validity. We will then revisit the literature on definitions and criteria for SRC proposed thus far and compare our core set to these previous definitions. Limitations and strengths of this study will be presented. Finally, we will outline how this core set will be used to complete the development and validation of classification criteria for SRC.

6.1 Overview of findings

Scleroderma renal crisis (SRC) is a serious complication of systemic sclerosis (SSc) that lacks a gold standard. Definitions of SRC reported to date are thus heterogeneous and none has been validated. A scoping review of the literature identified 40 heterogeneous definitions of SRC using 48 items (16). To address this deficiency, we have undertaken a multi-phase project to develop and validate classification criteria for SRC. Using consensus methodology, including an online Delphi survey and a nominal group discussion, the purpose of this phase of the study was to generate a core set of items to define SRC. From an initial pool of 31 items, 13 reached consensus during the Delphi exercise and five domains with 13 items, each with standardized definitions, emerged from the nominal group discussion. The domains consisted of rise in blood pressure, acute kidney injury, microangiopathic hemolytic anemia and thrombocytopenia, target organ dysfunction (encephalopathy, retinopathy, cardiomyopathy) and histopathology. Published evidence and consultation with experts were used to generate the standardized definitions of the items in the core set.

This project made some progress towards validation, namely face and content validity, of SRC classification criteria. Content validity is defined as the extent an instrument, such as a core set of items, incorporates the relevant construct being examined, such as SRC (40,42,43). In Round 1 of the Delphi exercise, experts were asked to identify omissions and clarify ambiguities and, in Rounds 2 and 3, they were asked to rank the validity of items. In addition, the NGT meeting allowed a structured discussion to address issues with any of the items used for helping define SRC to ensure the core set was as inclusive as possible. The ability to reword, reclassify, remove and add items throughout the process also contributed to the validity of the criteria. Finally, the use of experts in fields outside of rheumatology allowed for items specific to

hematology, neurology, cardiology, ophthalmology, and pathology to be incorporated providing further validity to items and definitions included in the core set.

6.2 Comparison with previously proposed criteria

Previous definitions and criteria for SRC were introduced in Chapter 2. In this section, we compare and contrast our core set to the ANCONA criteria for SRC (5), the criteria proposed by Hudson et al. (6), and the UK Scleroderma Study Group criteria (73).

A set of variables to define SRC known as the ANCONA criteria was proposed by experts in a study by Steen et al. in 2003 (29). The variables included systolic and diastolic blood pressure, serum creatinine, proteinuria and hematuria. In addition, criteria for SRC including findings of microangiopathic hemolytic anemia and renal histopathology were proposed. Our core set has similarities to the ANCONA criteria, particularly with respect to blood pressure. However, there are also notable differences in defining acute kidney injury (such as our exclusion of proteinuria and hematuria). In addition, our core set includes target organ dysfunction, definitions of variables and a detailed histopathological description of SRC.

To date, only Hudson et al. (2014) attempted to validate the ANCONA criteria and another slightly different set of criteria for SRC, that included encephalopathy (5). This study proposed criteria for SRC, where hypertensive SRC was defined by hypertension in addition to at least one of the following items: increase in serum creatinine, proteinuria, hematuria, thrombocytopenia, hemolysis and encephalopathy. When normotensive, hypertension was not included in the characteristics but rather serum creatinine in addition to either proteinuria, hematuria, thrombocytopenia, hemolysis and encephalopathy was required for classification. Although the criteria for hypertensive SRC performed well compared to physician judgement, the criteria for normotensive SRC, which did not include renal biopsy findings, did not perform well for this subset of SRC. In comparison, our core set does not include either proteinuria (which is non-specific in SSc) or hematuria (which suggests the presence of a mimicker, rather than true SRC). Serum creatinine has been regrouped into the domain of AKI and further redefined using a validated definition. Thrombocytopenia has also been regrouped with items from hemolysis to create a new domain of MAHAT, again using validated definitions. We believe that our core set represents a significant advancement over these earlier definitions. In

addition, it defines target organ involvement which further includes encephalopathy and provides a detailed histopathological description to define the term “findings consistent with SRC”.

More recently, in 2016, the UK Scleroderma Study Group (UKSSG) proposed diagnostic criteria for SRC (77), which vary in purpose from our classification criteria. Overlap between items presented in diagnostic and classification criteria can occur; however, diagnostic criteria typically are much broader, with a focus on patient care, whereas classification criteria tend to be more specifically defined and are used for research purposes. Thus, classification criteria should not be directly used as diagnostic criteria. The UKSSG criteria were divided into diagnostic criteria (essential) and supportive evidence (desirable). Blood pressure and AKI were categorized as essential diagnostic criteria, while MAHAT, hypertensive retinopathy, hematuria, oliguria or anuria, renal biopsy consistent with SRC features and flash pulmonary edema were considered supportive evidence. Discrepancies with our proposed criteria include different cut-off values for blood pressure (150/85 mmHg versus 140/90 mmHg) and the lack of an item for rise in blood pressure. Further, the UKSSG incorporated hematuria. Oliguria and flash pulmonary edema were both proposed as stand-alone items, whereas in our core set these items are grouped into the AKI and acute heart failure domains, respectively. Finally, our proposed core set provides detailed definitions for all items presented.

6.3 Limitations

This study is not without limitations. First, only 99/216 experts invited to participate accepted and 77 (78%), 60 (61%) and 69 (70%) of these experts participated in Rounds 1-3 of the Delphi, respectively. Response bias may have occurred as a result of the individuals’ self-selection to participate. Part of the explanation for the observed response rates may have been that the Delphi exercise was conducted during the summer and early fall in the Northern hemisphere. Numerous out-of-office replies were returned. On the other hand, to mitigate this source of bias, reminder emails were sent to optimize participation rates and the final sample was still substantial and representative.

Second, there are large gaps in knowledge on SRC. Due to the nature of the Delphi, a further form of response bias can occur through judgement-based bias or participant bias of individuals when responding to the Delphi questions and ratings. Participants in the Delphi likely ranked validity based somewhat on experiential, rather than on purely literature-based scientific

evidence. Nevertheless, we provided the Delphi participants with the scoping review and all of the original papers included therein in every Round for easy access to the available literature. In addition, this phase of the project will be followed by a future, data-driven phase.

Third, recruitment of participants with a broad range of expertise is critical to the success of a consensus-building exercise. We recruited subjects for the Delphi exercise and NGT meeting through scleroderma research groups and established networks. However, it became clear at the NGT meeting that content expertise in certain items pertaining to histopathology, hematology, neurology, ophthalmology, and cardiology was lacking. We therefore recruited experts in all of these fields to help finalize the relevant items. While recruitment of these individuals late in the process (following the NGT exercise) provided valuable information, including these experts as participants in earlier phases of the study may have allowed better contribution of their knowledge during the development of the core set and promoted further discussion of the items involved. However, it should be noted that all finalized results were agreed upon by all participants in this study.

Fourth, the core items presented do not include any biomarkers for this disease. Biomarkers such as rapidly progressive diffuse SSc and the presence of anti-RNA III antibodies are known risk factors for SRC. These biomarkers may help improve the performance of classification criteria, and should be considered in future research phases.

A final possible limitation for this thesis project focuses on the participants of the study, for both the Delphi and NGT meeting. All participants for this research project were clinicians with interest in SRC and content experts. OMERACT recommends that these individuals be included to obtain the validation of the core set presented, since experts should have the greatest working knowledge in the field of SRC. However, there was no input from patients, which is also recommended by OMERACT. Patients have different viewpoints and knowledge, based on their experience living with SRC. Incorporating patients into this type of research may reveal key items not identified by clinicians, such as how a patient may feel prior to diagnosis or throughout the early stages of SRC onset. Patients living with the disease may have different experiences. Their signs and symptoms may present differently to them than to the diagnosing physician. Physicians may miss these possibly relevant nuances. Information and input from patients could benefit the development of classification criteria and may be of interest in future studies.

6.4 Strengths

The study has many strengths. The research methodology, consisting of a paired Delphi exercise and face-to-face structured NGT meeting, allowed both quantitative and qualitative data collection. The Likert scale ratings from the Delphi provided median and IQR values that later allowed for the calculation of disagreement in the quantitative data. The NGT meeting allowed for participants to vote on items, providing quantitative data, but additionally allowed for discussion on defining items, thus providing qualitative data.

This study provided validation, through the well-developed methods and participation of many experts in the field of SSc and SRC as well as content experts such as hematologists, neurologists, cardiologists, ophthalmologists, and pathologists. The input from these experts helped ensure that the final core set had face and content validity.

The extensive geographic range of participants is another strength, helping to ensure that the core set will be generalizable - to the broad spectrum of SRC, as well as internationally.

Finally, the rigorous process followed for this project, including a previous scoping review of the literature, followed by the complementary consensus-based and data-driven components will help ensure the usefulness of the classification criteria for future randomized trials and epidemiologic research of SRC. These methods complemented each other well. The Delphi provided a cost-effective approach, allowed for international expert participation with the ability to provide honest feedback in a confidential manner. The online platform for the Delphi exercise was flexible and allowed for a well-organized, visually pleasing and engaging process. The NGT meeting allowed for a highly structured, face-to-face discussion of international expert participants led by an experienced moderator that was time-efficient. These approaches allowed opinions to be thoroughly shared in multiple formats to arrive at consensus-based classification criteria for SRC.

6.5 Future steps

The generation of the core set is only part of a bigger project. As discussed, previous research presented through a scoping review by Hoa et al. (2017), laid out ground work for our Delphi and NGT meeting to achieve consensus on a core set of items. Future phases of research will be needed to develop, weight and validate the classification criteria for SRC, which are already in planning phases. Moreover, a few additional elements that arose during this project

will need to be addressed. These include SRC mimickers that should not be confused with SRC and differentiation between separate hypertensive and normotensive criteria.

The concept of SRC mimickers became a secondary objective for the NGT meeting. The definitions used in research to date surrounding SRC are heterogeneous and broad, with few or no definitive indicators. The resulting broad criteria items can in fact be indications of another disease and vice versa. After the NGT meeting, it was agreed that there are some mimickers of SRC that should be excluded prior to making a diagnosis. These mimickers of SRC share similar clinical presentations to SRC and are also associated with AKI (16,74). They are found in many SSc patients. but also present in individuals with otherrenal disorders. They are presented in Table 7 in Chapter five of this thesis. The knowledge of SRC mimickers will benefit future studies; cohorts inclusive of SRC mimickers will provide information on the specificity of the criteria, thus further strengthening the development of core items for SRC classification.

Finally, the differentiation of criteria for normotensive vs. hypertensive forms of SRC should also be studied in future phases of research. In this thesis project, SRC was considered broadly to ensure that the online surveys could be designed in a manner that encouraged increased participation rates and minimized incomplete surveys. The concept of separate criteria for normotensive and for hypertensive forms of SRC was discussed in the NGT meeting and circulated to experts outside of the rheumatology scope. It was found that the additional item of renal biopsy recommended by a physician for normotensive SRC should be included. However, distinction of SRC classification criteria for these two forms was deferred for future phases when supporting data can be collected.

Two future phases of this research are now being planned. The first, modeled on the *International Scleroderma Renal Crisis Survey* (5), will be to recruit an inception SRC cohort and collect the items in the core set. A comparison cohort consisting of subjects with conditions that mimic SRC will also be assembled. These data will be used to further develop and validate classification criteria for SRC. The second will be a forced-choice experiment using multi-criteria decision analysis methods to assign weights to the items in the criteria, and to set probability values for definite, probable and possible SRC.

6.6 Summary

This discussion has explored the core set developed in this thesis project and how face and content validity were established. Previously introduced research to date on SRC and the definitions used in the current literature were further explored in comparison to the core set proposed in this project. The limitations and strengths of this study and future steps were discussed. Using literature on current SRC definitions and criteria and, where appropriate, incorporating existing guidelines for select items as the foundation for our study, we were successful at achieving consensus on a core set of domains and items for SRC. With all of these factors explored, we believe that the proposed core set is the most valid list to date and recommend that future work be conducted with this core set to develop and validate classification criteria for SRC.

CHAPTER 7. Conclusion

In conclusion, using consensus methodology, this study developed a core set of items to be considered in the development of classification criteria for SRC. Future phases of this research are now being planned. The resulting classification criteria are expected to facilitate rigorous research in SRC. In the meantime, SSc researchers who are designing new studies (either observational or trials) are encouraged to collect the core set of items from the current project in their datasets. The inclusion of these items will be useful for future validation of the criteria.

References

1. Traub YM, Shapiro AP, Rodnan GP, Medsger TA, McDonald RH, Steen VD, et al. Hypertension and renal failure (scleroderma renal crisis) in progressive systemic sclerosis. Medicine (Baltimore). 1983;62(6):335–52.

2. Teixeira L, Mouthon L, Mahr A, Berezné A, Agard C, Mehrenberger M, et al. Mortality and risk factors of scleroderma renal crisis: a French retrospective study of 50 patients. Ann Rheum Dis. 2008;67:110–6.

3. Guillevin L, Berezne A, Seror R, Teixeira L, Pourrat J, Mahr A, et al. Scleroderma renal crisis: A retrospective multicentre study on 91 patients and 427 controls. Rheumatology. 2012;51:460–7.

4. Walker JG, Ahern MJ, Smith MD, Coleman M, Pile K, Rischmueller M, et al. Scleroderma renal crisis: Poor outcome despite aggressive antihypertensive treatment. Intern Med J. 2003;33(5–6):216–20.

5. Hudson M, Baron M, Tatibouet S, Furst DE, Khanna D, Hummers L, et al. Exposure to ACE inhibitors prior to the onset of scleroderma renal crisis-results from the international scleroderma renal crisis survey. Semin Arthritis Rheum. 2014;43:666–72.

6. Aggarwal R, Ringold S, Khanna D, Neogi T, Johnson SR, Miller A, et al. Distinctions Between Diagnostic and Classification Criteria? Arthritis Care Res. 2015;67(7):891–7.

7. Leroy E. Scleroderma (Systemic Sclerosis): Classificiation, Subsets and Pathogenesis. J Rheumatol. 1988;15(2):202–5.

8. Van Den Hoogen F, Khanna D, Fransen J, Johnson SR, Baron M, Tyndall A, et al. 2013 classification criteria for systemic sclerosis: An american college of rheumatology/European league against rheumatism collaborative initiative. Arthritis Rheum. 2013;65(11):2737–47.

9. Nikpour M, Stevens WM, Herrick AL, Proudman SM. Epidemiology of systemic sclerosis. Best Pract Res Clin Rheumatol. 2010;24(6):857–69.

10. Mayes MD. Scleroderma epidemiology. Rheum Dis Clin North Am. 2003;29(2):239–54.

11. Mayes MD, Lacey J V., Beebe-Dimmer J, Gillespie BW, Cooper B, Laing TJ, et al. Prevalence, incidence, survival, and disease characteristics of systemic sclerosis in a large US population. Arthritis Rheum. 2003;48(8):2246–55.

12. LeRoy E, Medsger T. Criteria for the Classification of Systemic Sclerosis (Scleroderma). J Rheumatol. 2001;23(7):1573–6.

13. Mouthon L, Bérezné A, Bussone G, Noël LH, Villiger PM, Guillevin L. Scleroderma renal crisis: A rare but severe complication of systemic sclerosis. Clin Rev Allergy Immunol. 2011;40(2):84–91.

14. Steen VD, Syzd A, Johnson JP, Greenberg A, Medsger TA. Kidney disease other than renal crisis in patients with diffuse scleroderma. J Rheumatol. 2005;32(4):649–55.

15. Caron M, Hudson M, Baron M, Nessim S, Steele R. Longitudinal Study of Renal Function in Systemic Sclerosis. J Rheumatol. 2012;39(9):1829–34.

16. Hoa S, Stern EP, Denton CP, Hudson M, Baron M, Frech T, et al. Towards developing criteria for scleroderma renal crisis: A scoping review. Autoimmun Rev. 2017;16(4):407–15.

17. Nihtyanova SI, Schreiber BE, Ong VH, Rosenberg D, Moinzadeh P, Coghlan JG, et al. Prediction of pulmonary complications and long-term survival in systemic sclerosis. Arthritis Rheumatol. 2014;66(6):1625–35.

18. DeMarco PJ, Weisman MH, Seibold JR, Furst DE, Wong WK, Hurwitz EL, et al. Predictors and outcomes of scleroderma renal crisis: The High-Dose Versus Low-Dose D-Penicillamine in early diffuse systemic sclerosis trial. Arthritis Rheum. 2002;46(11):2983–9.

19. Hesselstrand R, Scheja A, Wuttge D. Scleroderma renal crisis in a Swedish systemic sclerosis cohort: survival, renal outcome, and RNA polymerase III antibodies as a risk factor. Scand J Rheumatol. 2012;41:39–43.

20. Steen VD, Medsger T a, Osial T a, Ziegler GL, Shapiro a P, Rodnan GP. Factors predicting development of renal involvement in progressive systemic sclerosis. Am J Med. 1984;76(5):779–86.

21. Montagna G La, Baruffo A, Maja L, Tirri E, Matrone C, Vatti M. Scleroderma Renal Crisis: Analysis of Prevalence and Outcome in a Large Italia Series. J Clin Rheumatol. 1997;3(4):188–93.

22. Hamaguchi Y, Kodera M, Matsushita T, Hasegawa M, Inaba Y, Usuda T, et al. Clinical and immunologic predictors of scleroderma renal crisis in Japanese systemic sclerosis patients with anti-RNA polymerase III autoantibodies. Arthritis Rheumatol. 2015;67(4):1045–52.

23. Steen VD, Medsger TA. Case-control study of corticosteroids and other drugs that either precipitate or protect from the development of scleroderma renal crisis. Arthritis Rheum. 1998;41(9):1613–9.

24. Terras S, Hartenstein H, Höxtermann S, Gambichler T, Kreuter A. RNA polymerase III autoantibodies may indicate renal and more severe skin involvement in systemic sclerosis. Int J Dermatol. 2016;55(8):882–5.

25. Nguyen B, Assassi S, Arnett FC, Mayes MD. Association of RNA polymerase III antibodies with scleroderma renal crisis. J Rheumatol. 2010;37(5):1068.

26. Montanelli G, Beretta L, Santaniello A, Scorza R. Effect of dihydropyridine calcium channel blockers and glucocorticoids on the prevention and development of scleroderma renal crisis in an Italian case series. Clin Exp Rheumatol. 2013;31:135–9.

27. Nguyen B, Mayes MD, Arnett FC, Del Junco D, Reveille JD, Gonzalez EB, et al. HLA-DRB1*0407 and *1304 are risk factors for scleroderma renal crisis. Arthritis Rheum. 2011;63(2):530–4.

28. Yanaba K, Asano Y, Tada Y, Sugaya M, Kadono T, Hamaguchi Y, et al. Increased serum soluble CD147 levels in patients with systemic sclerosis: Association with scleroderma renal crisis. Clin Rheumatol. 2012;31(5):835–9.

29. Steen VD, Mayes MD, Merkel PA. Assessment of kidney involvement. Clin Exp Rheumatol. 2003;21.

30. Steen VD. Scleroderma and renal crisis. Rheum Dis Clin North Am. 2003;29:315–33.

31. Penn H, Howie AJ, Kingdon EJ, Bunn CC, Stratton RJ, Black CM, et al. Scleroderma renal crisis: Patient characteristics and long-term outcomes. Qjm. 2007;100(8):485–94.

32. Steen VD. Scleroderma renal crisis. Rheum Dis Clin North Am. 1996;22(4):861–78.

33. Denton CP, Khanna D. Systemic sclerosis. Lancet. 2017;390:1685–99.

34. Helfrich DJ, Banner B, Steen VD, Medsger TA. Normotensive renal failure in systemic sclerosis. Arthritis Rheum. 1989;32(9):1128–34.

35. Steen VD, Medsger TA. Changes in causes of death in systemic sclerosis, 1972-2002. Ann Rheum Dis. 2007;66(7):940–4.

36. Siragy HM, de Gasparo M, El-Kersh M, Carey RM. Angiotensin-Converting Enzyme Inhibition Potentiates Angiotensin II Type 1 Receptor Effects on Renal Bradykinin and cGMP. Hypertension. 2001;38:183–6.

37. Kellum J a, Lameire N, Aspelin P, Barsoum RS, Burdmann E a, Goldstein SL, et al. KDIGO Clinical Practice Guideline for Acute Kidney Injury. Kidney Int Suppl. 2012;2(1):1–138.

38. Scully M, Cataland S, Coppo P, de la Rubia J, Friedman KD, Kremer Hovinga J, et al. Consensus on the standardization of terminology in thrombotic thrombocytopenic purpura and related thrombotic microangiopathies. J Thromb Haemost. 2017;15(2):312–22.

39. Neunert C, Lim W, Crowther M, Cohen A, Solberg L, Crowther MA. The American Society of Hematology 2011 evidence-based practice guideline for immune thrombocytopenia. Blood. 2011;117(16):4190–207.

40. Polit DE, Beck CT. The content validity index: Are you sure you know what's being reported? Critique and rcommendations. Res Nurs Health. 2006;29:489–97.

41. Streiner DL, Norman GR, Cairney J. Health measurement scales: a practical guide to their development and use. Oxford University Press, USA; 2015.

42. Lynn MR. Determination and Quantification Of Content Validity. Vol. 35, Nursing Research. 1986. p. 382–6.

43. Lawshe C. A Quantitative Approach To Content Validity. Pers Psychol. 1975;(1):563–75.

44. Cronbach LJ, Meehl PE. Construct validity in psychological tests. 1955;(52):281–302.

45. Kaplan RM, Bush JW, Berry CC. Health Status : Types of Validity and the Index of Well-being. 1976;478–507.

46. Fink A, J. K, M. C, Brook RH. Consensus Methods: Characteristics and Guidelines for Use. 1991;74(9):1–9.

47. Gallagher M, Hares T, Spencer J, Bradshaw C, Webb I. The nominal group technique: a research tool for general practice? Fam Pract. 1993;10(1):76–81.

48. Potter M, Gordon S, Hamer P. The Nominal Group Technique: A useful consensus methodologyin physiotherapy research. NZ J Physiother. 2004;32(3):70–5.

49. Fitch K, Bernstein SJJ, Aguilar MDD, Burnand B, LaCalle JRR, Lazaro P, et al. The RAND / UCLA Appropriateness Method User's Manual. RAND. 2001. 109 p.

50. Dalkey N, Helmer O. Delphi Method of Convergence.Pdf. Vol. 9, Management Science. 1963. p. 458–67.

51. Pawlowski, Suzanne D, Okoli C. The Delphi Method as a Research Tool : An Example , Design Considerations and Applications 1 Introduction 2 Overview of the Delphi method. Inf Manag. 2004;42(1):15–29.

52. Furst DE. Outcome measures in rheumatologic clinical trials and systemic sclerosis. Rheumatology. 2009;47(Suppl 5):29–30.

53. Sinha IP, Smyth RL, Williamson PR. Using the Delphi technique to determine which outcomes to measure in clinical trials: Recommendations for the future based on a systematic review of existing studies. PLoS Med. 2011;8(1).

54. Strand V, Medicine S, Tugwell P. The OMERACT Filter for Outcome Measures in Rheumatology. 1998;25(2):198–9.

55. Boers M, Kirwan JR, Wells G, Beaton D, Gossec L, Agostino M, et al. Developing Core Outcome Measurement Sets for Clinical Trials : Omeract Filter 2.0. 2014;67:745–53.

56. Harvey N, Holmes CA. Nominal group technique: An effective method for obtaining group consensus. Int J Nurs Pract. 2012;18(2):188–94.

57. List D. The consensus group technique in social research. 2001;13(3):277–90.

58. Barber CEH, Marshall DA, Alvarez N, Mancini GBJ, Lacaille D, Keeling S, et al. Development of Cardiovascular Quality Indicators for Rheumatoid Arthritis: Results from an International Expert Panel Using a Novel Online Process. J Rheumatol. 2015;42(9):1548–55.

59. Schmajuk G, Hoyer B, Aringer M, Johnson S, Daikh D, Dörner T, et al. Multi-center Delphi Exercise Reveals Important Key Items for Classifying Systemic Lupus Erythematosus. Arthritis Care Res. 2017;

60. Fransen J, Johnson SR, Van Den Hoogen F, Baron M, Allanore Y, Carreira PE, et al. Items for developing revised classification criteria in systemic sclerosis: results of a consensus exercise with the ACR/EULAR working committee for classification criteria in systemic sclerosis. 2012;64(3):351–7.

61. Johnson SR, Khanna D, Cervera R, Costedoat-chalumeau N, Dafna D. Use of Nominal Group Technique to Determine Candidate Items for SLE Classi cation Criteria Development. Arthritis Rheumatol. 2016;68(suppl 10).

62. Whelton PK, Carey RM, Aronow WS, Casey DE, Collins KJ, Dennison Himmelfarb C, et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults. Hypertension. 2017.

63. Williams B, Poulter NR, Brown MJ, Davis M, Gordon M, John P, et al. British Hypertension Society guidelines for hypertension management 2004 (BHS-IV): summary. BMJ. 2003;328(7440):634–40.

64. National Institute for Health and Care Excellence. Hypertension in adults: diagnosis and management. NICE Guidel. 2011;(August 2011).

65. Zini G, Onofrio GD, Briggs C, Erber W, Jou JM, Lee SH, et al. ICSH recommendations for identification , diagnostic value , and quantitation of schistocytes. Int J Lab Hematol. 2012;34:107–16.

66. Ho VT, Cutler C, Carter S, Martin P, Adams R, Horowitz M, et al. Blood and Marrow Transplant Clinical Trials Network Toxicity Committee Consensus Summary : Thrombotic Microangiopathy after Hematopoietic Stem Cell. Biol Blood Marrow Transpl. 2005;11(8):571–5.

67. Lamy C, Mas JL. Hypertensive Encephalopathy. Fifth Edit. Stroke. Elsevier; 2011. 734-740 p.

68. Keith NM, Wagener HP BN. Some Different Types of Essential Hypertension: their course and Prognosis. Vol. 197, The American Journal of the Medical Sciences. 1939. p. 332–43.

69. Scheie HG. Evaluation of ophthalmoscopic changes of hypertension and arteriolar sclerosis. AMA Arch Ophthalmol. 1953;49(2):117–38.

70. Ezekowitz JA, O'Meara E, McDonald MA, Abrams H, Chan M, Ducharme A, et al. 2017 Comprehensive Update of the Canadian Cardiovascular Society Guidelines for the Management of Heart Failure. Can J Cardiol. 2017;33(11):1342–433.

71. Yancy CW, Jessup M, Bozkurt B, Butler J, Casey DE, Drazner MH, et al. 2013 ACCF/AHA guideline for the management of heart failure: A report of the American college of cardiology foundation/american heart association task force on practice guidelines. J Am Coll Cardiol. 2013;62(16):e147–239.

72. Adler Y, Charron P, Imazio M, Badano L, Barón-Esquivias G, Bogaert J, et al. 2015 ESC Guidelines for the diagnosis and management of pericardial diseases. Eur Heart J. 2015;36(42):2921–64.

73. Lusco MA, Najafian B, Alpers CE, Fogo AB. AJKD Atlas of Renal Pathology: Systemic Sclerosis. Am J Kidney Dis. 2016;67(4):e19–20.

74. Batal I, Domsic RT, Medsger TA, Bastacky S. Scleroderma renal crisis: A pathology perspective. Int J Rheumatol. 2010;2010.

75. Mouthon L, Bussone G, Berezné A, Noël L, Guillevin L. Scleroderma Renal Crisis. 2014;41(6).

76. Johnson SR, Fransen J, Khanna D, Baron M, Den F Van, Jr TAM, et al. Validation of potential classification criteria for systemic sclerosis. 2012;64(3):358–67.

77. Lynch BM, Stern EP, Ong VH, Harber M, Burns A, Denton CP. UK Scleroderma Study Group (UKSSG) Guidelines on the Diagnosis and Management of Scleroderma Renal Crisis.

78. Johnson SR, Goek O, Singh-grewal D, Vlad SC, Feldman BM, Felson DT, et al. Classification Criteria in Rheumatic Diseases: A Review of Methodologic Properties. 2007;57(7):1119–33.

79. Johnson SR, Naden RP, Fransen J, Hoogen F Van Den, Pope JE, Baron M, et al. Multicriteria decision analysis methods with 1000Minds for developing systemic sclerosis classification criteria. J Clin Epidemiol. 2014;67(6):706–14.

Appendix 1

Items from pre-existing definitions

- Hypertension
  - Systolic blood pressure (>140, 150, 160 or 180 mmHg) (new onset)
  - Diastolic blood pressure (>85, 90, 100, 110 or 120 mmHg) (new onset)
  - Rise in systolic blood pressure >30 mmHg compared with baseline
  - Rise in diastolic blood pressure >20 mmHg compared with baseline
  - Abrupt onset or aggravation
  - Measured on at least 2 occasions
    - Minimum of 12 hours apart
    - Over a 24-hour period
  - Measured within 3 days of first event-associated observation
  - New onset of blood pressure >150/85 mmHg obtained at least twice over a 24-h period (i.e. significant hypertension as defined by the New York Health Association)
  - Independent of concomitant antihypertensive medication use
  - Responsive to ACE-inhibitors
- Renal insufficiency (or Azotemia)
  - Serum creatinine >120% of upper limit of normal for local laboratory
  - Serum creatinine ≥2.0 mg/day
  - Creatinine clearance ≤50 ml/min
  - Fall in creatinine clearance to <60 ml/min
  - Fall in estimated glomerular filtration rate (eGFR) by >30%
  - Increase in serum creatinine >50% over baseline
  - Increase in serum creatinine by ≥1.5 times baseline, known or presumed to have occurred within the prior 7 days, or increase in serum creatinine by ≥0.3 mg/dL (>26.5 μmol/L) within 48 hours (i.e. acute kidney injury according to KDIGO definitions)
  - Doubling of serum creatinine above the value at baseline
  - Rapid increase in serum creatinine
  - When possible, a repeat serum creatinine and recalculation of eGFR should be obtained to corroborate the initial results
  - Rapid deterioration of renal function (within a period of <1 month)
  - Rapidly progressive oliguric renal insufficiency
  - Oliguria or anuria
  - Absence of other defined cause
- Proteinuria
  - ≥ 2+ by dipstick
  - ≥ 1+ by dipstick
  - Protein: creatinine ratio > upper limit of normal
  - >500 mg in 24 hours
- Hematuria
  - >2+ by dipstick
  - >10 RBCs/HPF
  - New onset of urinary RBCs (excluding other causes)
  - Without menstruation
- Thrombocytopenia
  - <100 000 platelets/mm3
- Hemolysis (or Microangiopathic hemolytic anemia)
  - Schistocytes or other RBC fragments seen on blood smear
  - Increased reticulocyte count
  - Increase in LDH and indirect bilirubin
  - Haptoglobin consumption
  - Anemia not because of other causes
- Hypertensive encephalopathy
  - Seizures
- Hypertensive retinopathy
  - Grade III (flame-shaped hemorrhages and/or "cotton-wool" exudates) or IV (papilledema) retinopathy, according to Keith-Wagener classification
  - Retinopathy typical of acute hypertensive crisis
- Hyperreninemia
  - Elevation of plasma renin activity to twice the upper limit of normal or higher
- Abnormal kidney biopsy
  - Typical/characteristic changes of SRC (not further defined)
  - Findings consistent with SRC (microangiopathy) (not further defined)
  - Accumulation of mucin in interlobular arteries (indistinguishable from accelerated hypertension) and fibrinoid calcinosis of arteries
- Flash pulmonary edema

Predictors of SRC

- Patient-specific characteristics
  - Black race
  - Male sex
- Clinical characteristics
  - Shorter disease duration
  - Diffuse cutaneous subset
  - Skin score (>14 or 20)
  - Large joint contractures
  - Tendon friction rubs
  - Digital pitting scars
  - Cardiopulmonary involvement
    - Cardiac insufficiency
    - Pericarditis
    - FVC <75%
    - Lower DLCO
  - Muscle involvement
    - Muscle wekaness
    - High creatine kinase
    - Myalgias and myopathy
  - Arthralgias
- Medication history
  - Prednisone (prior or simultaneous use; higher dose; within prior 1 or 3 months; ≥15mg/d in prior 6 months)
  - Absence of calcium channel blocker
- Prednisone Serologies, Biomarkers and Genetics
  - Anti-RNA polymerase III positivity
  - Anti-RNA polymerase I/II/III positivity
  - ELISA anti-RNA polymerase III ≥157
  - Absence of anti-centromere
  - Anti-nRNP positivity
  - Speckled ANA
  - Anti-Scl70 positivity
- Biomarkers and Genetics
  - High lipocalin-2 levels
  - High sCD147 levels
  - High angiogenin levels
  - High endothelin-1 levels
  - HLA-DRB1*0407
  - HLA-DRB1*1304

Differential diagnoses to exclude

ANCA-associated glomerulonephritis
Thrombotic thrombocytopenic purpura /hemolytic uremic syndrome
Membranous nephropathy
Drug-induced nephropathies (e.g. cyclosporin A)
Other vasculitides (e.g. polyarteritis nodosa, mixed cryoglobulinemia, Goodpasture syndrome)
Oxalate nephropathy
Renal artery stenosis
Membranoproliferative nephropathy
Pre-renal causes (e.g. sepsis, dehydration, cardiac or pulmonary vascular involvement)
Isolated renal abnormalities

Other considerations

Hypertensive vs. Normotensive SRC
Definite vs. Probable (or Suspected or Possible) SRC
Classic vs. Subacute presentation of SRC
Restricted to SSc vs. expanded to SSc-spectrum of connective tissue diseases

Appendix 2

Confidential

Development of a consensus definition for scleroderma renal crisis (SRC)

Thank you for showing interest in this research study. Below is relevant information pertaining to the study. Please read all information before proceeding. If you have any questions, contact Dr Marie Hudson at marie.hudson@mcgill.ca.

Who is conducting the study?
The study is led by two principal investigators, Dr Marie Hudson, Jewish General Hospital and McGill University, Montreal, Canada, and Dr Christopher Denton, Royal Free Hospital, London, UK, under the auspices of the Scleroderma Clinical Trial Consortium (SCTC) Working Group on Scleroderma Renal Crisis.

Who is funding the research?
The study is being funded by a Scleroderma Clinical Trials Consortium (SCTC) grant to the Scleroderma Renal Crisis Working Group.

Why is the study being conducted?
Currently, there is no gold standard definition of SRC thereby allowing for important knowledge gaps in the understanding of this disease. Outcomes have been reported to vary widely but different studies have used different criteria to define SRC. Criteria for this disease have been proposed but none have been validated. We wish to develop and validate classification criteria for SRC and improve systematic research in this condition. This phase of the project aims to identify a core set of variables to be considered for these criteria.

What is expected of you as a research participant?
Participants are expected to complete all three online rounds of the Delphi exercise. Participants are expected to provide an answer to each question for the survey to be complete and are encouraged to provide feedback and comments when asked. Participants can take part in other studies during the course of this study. The Delphi exercise will consist of 3 rounds held approximately 6 weeks apart. Each round will be open for 7-14 days. Each survey should take approximately 20 minutes to complete.

Risks:
There are no risks associated with this study.

Benefits:
The direct individual benefits for participation in this study are minimal. Nevertheless, having a working definition will facilitate future research in SRC. In addition, participants who complete all 3 rounds of the survey will be included as investigators of the SCTC SRC working group.

Voluntary participation/withdrawal:
Your participation is voluntary, you may choose to withdraw from this study at any time. If you choose to withdraw, any information that has been collected up to the date of withdrawal may still be used for the study. You may be withdrawn from the study if you do not follow the instructions for participation in the study.

Confidentiality:
During your participation in the research study, we will collect and store personal identifiable information about you in a password protected account in a REDCap database. Only information necessary for the research study will be collected and it will remain confidential.

Marie Hudson and her research staff assigned to this project will have access to the REDCap account. Although your identity will be shared with other participants, your personal information and responses will remain confidential towards other participants.

After completion of each round of the Delphi exercise, summary statistics for each question will be returned to you in addition to your responses. However, no other participant will be given your responses; only summary data will be shared.

The aggregate results of the study may be printed/published or shared with other people in the scientific community. Aside from being acknowledged as an investigator of the SCTC SRC working group, your personal information and your responses to the survey will remain confidential.

Costs and compensation:
You will not be paid for your participation in this study. There will be no costs to you for participating in the 3 online rounds of the Delphi exercise.

Ethics

Ethics approval for this project was obtained from the Jewish General Hospital Research Ethics Board, Montréal, Quebec, Canada (Ethics Protocol # CODIM-MBM-17-104).

If you would like to read the research protocol for this research study, follow the link attached.

[Attachment: "Research protocol.pdf"]

Consent to Participate:

• I confirm that I have read the project based on the information provided and if I have any questions I can contact the principal investigators.

• My participation is voluntary and I am free to withdraw at any time without providing a reason and without any sort of penalty.

• Any data I provide will be treated securely and kept confidential.

• I agree to take part in the survey.

By continuing and completing this survey, you are providing consent and agreeing to the above statements.

It is recommended that you save a copy of this page for your records.

Demographic information

First name _____

Last name _____

Where are you from? _____
((City, Country))

What are your institutional affiliations? Please provide an address and postal code _____

What are your academic degrees?

What is your specialty?

- ◯ Rheumatologist
- ◯ Nephrologist
- ◯ Pathologist
- ◯ Internist
- ◯ Lab Scientist
- ◯ Other (Please specify)

What organization(s) are you affiliated with?

- ☐ Scleroderma Clinical Trials Consortium (SCTC)
- ☐ Canadian Scleroderma Research Group (CSRG)
- ☐ European Scleroderma Trials and Research (EUSTAR)
- ☐ Australian Scleroderma Interest Group (ASIG)
- ☐ Other (Please specify)

How many years have you been working as a clinician?

- ◯ 1-10
- ◯ 11-20
- ◯ 21-30
- ◯ >30
- ◯ I am not a clinician

How many unique scleroderma patients do you see each year?

- ◯ 1-30
- ◯ 31-60
- ◯ 61-100
- ◯ >100

How many new scleroderma renal crisis (SRC) patients do you see each year?

- ◯ 0
- ◯ 1-2
- ◯ 3-5
- ◯ >5 (Please specify)

How many returning SRC patients do you see each year?

- ◯ 0
- ◯ 1-5
- ◯ 6-10
- ◯ 10-15
- ◯ >15 (Please specify)

Comment field:

Below you will find the published scoping literature review on definitions for SRC, and the related papers discussing two sets of criteria for SRC that have been proposed to date.

[Attachment: "Hoa et al. 2017.pdf"]

[Attachment: "Steen et al. 2003.pdf"]

[Attachment: "Hudson et al. 2014.pdf"]

If you would like to review the literature pertaining to various definitions and classification criteria used for SRC identified in the above scoping review (Hoa et al. 2017) please download the attached zip file.

[Attachment: "Literature to be reviewed.zip"]

Introduction to Delphi Round 1

Please read the following carefully.

We have compiled the items that have been used to define SRC to date from a scoping review of the literature. This first round of the exercise is aimed at identifying items that may have been omitted or clarifying items that may be ambiguous. We have also included a few additional questions to explore alternative definitions. Please make comments in the spaces provided.

Of note, we are NOT YET interested in building consensus on validity or feasibility. In addition, there may be apparent redundancies. However, consensus and item reduction will be pursued in subsequent rounds of the exercise.

Hypertension

Please comment on ambiguities, inaccuracies or oversights of the following items to define hypertension in SRC.

1.1 New onset of high blood pressure ≥150/85 mmHg obtained at least twice over a 24-hour period ◯ Comment ◯ No comment

1.1 comment

1.2 New onset of systolic blood pressure (SBP) ≥140 mmHg or rise in SBP ≥30 mmHg compared with baseline ◯ Comment ◯ No comment

1.2 comment

1.3 New onset of diastolic blood pressure (DBP) ≥85 mmHg or rise in DBP ≥20 mmHg compared with baseline ◯ Comment ◯ No comment

1.3 comment

1.4 Blood pressure changes in both SBP and DBP ◯ Comment ◯ No comment

1.4 comment

1.5 Abrupt onset or aggravation of hypertension ◯ Comment ◯ No comment

1.5 comment

1.6 Blood pressure changes independent of concomitant antihypertensive medication use ◯ Comment ◯ No comment

1.6 comment

1.7 Blood pressure changes responsive to ACE-inhibitors ◯ Comment ◯ No comment

1.7 comment

1.8 What do you think would be appropriate cutoffs for high blood pressure?

1.8.1 Absolute SBP

◯ ≥ 150 mmHg
◯ ≥ 160 mmHg
◯ ≥ 180 mmHg
◯ No comment
◯ Other (Please specify)

1.8.1 comment

1.8.2Rise in SBP

◯ ≥ 20 mmHg
◯ ≥ 40 mmHg
◯ No comment
◯ Other (Please specify)

1.8.2 comment

1.8.3Absolute DBP

◯ ≥ 90 mmHg
◯ ≥ 100 mmHg
◯ ≥ 110 mmHg
◯ ≥ 120 mmHg
◯ No comment
◯ Other (Please specify)

1.8.3 comment

1.8.4Rise in DBP

◯ ≥ 10 mmHg
◯ ≥ 30 mmHg
◯ No comment
◯ Other (Please specify)

1.8.4 comment

1.9 Blood pressure changes should be measured:

1.9.1on at least 2 occasions a minimum of 12 hours apart ◯ Comment ◯ No comment

1.9.1 comment

1.9.2on at least 2 occasions over a 24-hour period ◯ Comment ◯ No comment

1.9.2 comment

1.9.3within 3 days of the first event-associated observation ◯ Comment ◯ No comment

1.9.3 comment

Comment field:

Renal Insufficiency (or Azotemia)

Please comment on ambiguities, inaccuracies or oversights of the following items to define renal insufficiency in SRC.

Note: creatinine clearance is not, whereas estimated glomerular filtration rate (eGFR) is, adjusted for body surface area. eGFR estimates can therefore be applied to determine level of kidney function, regardless of a patient's size.

2.1 Increase in serum creatinine by ≥1.5 times baseline, known or presumed to have occurred within the prior 7 days, or increase in serum creatinine by ≥0.3 mg/dL (≥26.5 µmol/L) within 48 hours ◯ Comment ◯ No comment

2.1 comment

2.2 Increase in serum creatinine ≥50% over baseline ◯ Comment ◯ No comment

2.2 comment

2.3 Doubling of serum creatinine above the value at baseline ◯ Comment ◯ No comment

2.3 comment

2.4 Serum creatinine ≥2.0 mg/dL (177 µmol/L) ◯ Comment ◯ No comment

2.4 comment

2.5 Serum creatinine ≥120% (or 1.2 times) the upper limit of normal for local laboratory ◯ Comment ◯ No comment

2.5 comment

2.6 Rapid increase in serum creatinine ◯ Comment ◯ No comment

2.6 comment

2.7 Fall in creatinine clearance to ≤60 mL/min ◯ Comment ◯ No comment

2.7 comment

2.8 Fall in estimated glomerular filtration rate (eGFR) by ≥30% ◯ Comment ◯ No comment

2.8 comment

2.9 Are there more suitable cutoffs for serum creatinine, creatinine clearance or eGFR than those proposed above? ◯ Comment ◯ No comment

2.9 comment

2.10 Rapid deterioration of renal function (within a period of ≤ 1 month) ◯ Comment ◯ No comment

2.10 comment

2.11 Rapidly progressive oliguric renal insufficiency ◯ Comment ◯ No comment

2.11 comment

2.12 Presence of oliguria or anuria ◯ Comment ◯ No comment

2.12 comment

2.13 A repeat serum creatinine and recalculation of renal function should be obtained to corroborate the initial results ◯ Comment ◯ No comment

2.13 comment

2.14 Absence of other defined cause of Acute Kidney Injury (AKI) ◯ Comment ◯ No comment

2.14 comment

Comment field:

Proteinuria

Please comment on ambiguities, inaccuracies or oversights of the following items to define proteinuria in SRC.

3.1 ≥ 2+ (100-300 mg/dL range) by dipstick ◯ Comment ◯ No comment

3.1 comment

3.2 ≥ 1+ (30-100 mg/dL range) by dipstick ◯ Comment ◯ No comment

3.2 comment

3.3 Protein:creatinine ratio > upper limit of normal (≤ 150 mg/day) ◯ Comment ◯ No comment

3.3 comment

3.4 ≥ 500 mg of albumin concentration in 24 hours ◯ Comment ◯ No comment

3.4 comment

Comment field:

Hematuria

Please comment on ambiguities, inaccuracies or oversights of the following items to define hematuria in SRC.

4.1 ≥ 2+ by dipstick ◯ Comment ◯ No comment

4.1 comment

4.2 ≥ 10 RBCs/HPF ◯ Comment ◯ No comment

4.2 comment

4.3 New onset of urinary RBCs (excluding other causes) ◯ Comment ◯ No comment

4.3 comment

4.4 Without menstruation ◯ Comment ◯ No comment

4.4 comment

Comment field:

Thrombocytopenia

Please comment on ambiguities, inaccuracies or oversights of the following items to define thrombocytopenia in SRC.

5.1 ≤ 100,000 platelets/mm3 ◯ Comment ◯ No comment

5.1 comment

Comment field:

Hemolysis (or Microangiopathic Hemolytic Anemia)

Please comment on ambiguities, inaccuracies or oversights of the following items to define hemolysis in SRC.

6.1 Presence of schistocytes or other RBC fragments on blood smear ◯ Comment ◯ No comment

6.1 comment

6.2 Increased reticulocyte count ◯ Comment ◯ No comment

6.2 comment

6.3 Increase in LDH and indirect bilirubin (to show breakdown of RBC) ◯ Comment ◯ No comment

6.3 comment

6.4 Haptoglobin consumption ◯ Comment ◯ No comment

6.4 comment

6.5 Anemia not because of other causes ◯ Comment ◯ No comment

6.5 comment

Comment field:

Encephalopathy

Please comment on ambiguities, inaccuracies or oversights of the following items to define encephalopathy in SRC.

7.1 Encephalopathy manifested by the presence of seizures ◯ Comment ◯ No comment

7.1 comment

Comment field:

Retinopathy

Please comment on ambiguities, inaccuracies or oversights of the following items to define retinopathy in SRC.

8.1 Grade III (flame-shaped hemorrhages and/or "cotton-wool" exudates) or IV (papilledema) retinopathy, according to Keith-Wagener classification ◯ Comment ◯ No comment

8.1 comment

8.2 Retinopathy typical of acute hypertensive crisis ◯ Comment ◯ No comment

8.2 comment

Comment field:

Hyperreninemia

Please comment on ambiguities, inaccuracies or oversights of the following items to define hyperreninemia in SRC.

9.1 Elevation of plasma renin activity ≥ 2 times the upper limit of normal ◯ Comment ◯ No comment

9.1 comment

Comment field:

Cardiac Dysfunction

Please comment on ambiguities, inaccuracies or oversights of the following items to define cardiac dysfunction in SRC.

10.1 Presence of flash pulmonary edema ◯ Comment ◯ No comment

10.1 comment

Comment field:

Abnormal Kidney Biopsy

Please comment on ambiguities, inaccuracies or oversights of the following items to define abnormal kidney biopsy in SRC.

11.1 Mucoid (myxoid) change in interlobular arteries and fibrinoid necrosis of arteries ◯ Comment ◯ No comment

11.1 comment

11.2 Typical/characteristic changes of SRC ◯ Comment ◯ No comment

11.2 comment

11.3 Findings consistent with SRC (microangiopathy) ◯ Comment ◯ No comment

11.3 comment

11.4 Proposed definition for "typical findings of SRC" as follows: ◯ Comment ◯ No comment

Small vessel (arcuate and interlobular arteries) changes predominate over glomerular alterations. Early vascular abnormalities include intimal accumulation of myxoid material, thrombosis, and/or fibrinoid necrosis**.** Intimal thickening and endothelial cell proliferation lead to characteristic vascular "onion-skin" lesions. Narrowing and obliteration of the vascular lumen lead to glomerular ischemia. Juxtaglomerular apparatus (JGA) hyperplasia, while relatively rare (10%), can be observed. Late changes are manifested by glomerulosclerosis and interstitial fibrosis.

11.4 comment

Comment field:

Normotensive versus Hypertensive SRC

Except for hypertension, do you think that the items used to define hypertensive SRC should be defined differently for normotensive SRC?

12.1 Hypertension ◯ Comment ◯ No comment

12.1 comment

12.2 Renal insufficiency ◯ Comment ◯ No comment

12.2 comment

12.3 Proteinuria ◯ Comment ◯ No comment

12.3 comment

12.4 Hematuria ◯ Comment ◯ No comment

12.4 comment

12.5 Thrombocytopenia ◯ Comment ◯ No comment

12.5 comment

12.6 Hemolysis ◯ Comment ◯ No comment

12.6 comment

12.7 Encephalopathy ◯ Comment ◯ No comment

12.7 comment

12.8 Retinopathy ◯ Comment ◯ No comment

12.8 comment

12.9 Hyperreninemia ◯ Comment ◯ No comment

12.9 comment

12.10 Cardiac dysfunction ◯ Comment ◯ No comment

12.10 comment

12.11 Abnormal kidney biopsy ◯ Comment ◯ No comment

12.11 comment

Comments

Please add any other comments that you feel would be helpful to define SRC

Appendix 3

Development of a consensus definition for SRC - Round 2

Thank you participating in the second round of this Delphi exercise.

The goal of this round is to begin to build consensus on a core set of items that could be used to develop classification criteria for SRC. The items in this survey have been identified from a scoping review of the literature (which is appended herewith, along with the original papers included in the review). Items that are identified as valid and feasible will be used to inform future data-driven development of the classification criteria.

Although the items included in this survey were identified from a scoping review of the literature, most have not been formally validated. Thus, for the purposes of this exercise, you will be asked to rate scientific (based on the literature provided) and empirical (based on your experience and knowledge of professional consensus) validity separately. In addition, since optimal cut-offs for several items (eg. blood pressure, azotemia) are not known, additional questions have been added to allow you to express your opinion on these.

You will also be asked to rate feasibility, based on whether the information necessary to identify the item is possible to find in an average medical record and is likely to be reliable.

The validity of items will be rated using Likert-type scales ranging from 1-9 with labeled endpoints (1= very invalid/unfeasible, 9 = very valid/feasible). Note that the midpoint of 5 is labelled “uncertain”, and may be used if you don’t know or are unsure of an item.

Finally, note that all items proposed in this survey assume that the findings are not explained by other medical conditions.

Below you will find the published scoping literature review on SRC, and two papers that have, to date, proposed criteria for SRC.

Below you will find the published scoping literature review on SRC, and two papers that have, to date, proposed criteria for SRC.

[Attachment: "Hoa et al. 2017.pdf"]

[Attachment: "Steen et al. 2003.pdf"]

[Attachment: "Hudson et al. 2014.pdf"]

In addition, if you would like to review the literature pertaining to various definitions and classification criteria used for SRC identified in the scoping review (Hoa et al. 2017) please download the attached zip file.

[Attachment: "Literature to be reviewed.zip"]

Hypertension

Please rate the scientific validity (based on the literature provided), empirical validity (based on your experience and knowledge of professional consensus) and feasibility (based on whether the information can be found in an average medical record and is likely to be reliable) of the following items.

1.1 New onset or deterioration of pre-existing hypertension, defined as any of the following:

1.1.1A) Systolic blood pressure ≥ 140 mmHg

Invalid/Unfeasible_____Uncertain_____Valid/Feasible

	1	2	3	4	5	6	7	8	9
Scientific Validity	◯	◯	◯	◯	◯	◯	◯	◯	◯
Empirical Validity	◯	◯	◯	◯	◯	◯	◯	◯	◯
Feasibility	◯	◯	◯	◯	◯	◯	◯	◯	◯

1.1.2B) Diastolic blood pressure ≥ 90 mmHg

	1	2	3	4	5	6	7	8	9
Scientific Validity	◯	◯	◯	◯	◯	◯	◯	◯	◯
Empirical Validity	◯	◯	◯	◯	◯	◯	◯	◯	◯
Feasibility	◯	◯	◯	◯	◯	◯	◯	◯	◯

1.1.3C) Rise in systolic blood pressure ≥ 30 mmHg

	1	2	3	4	5	6	7	8	9
Scientific Validity	◯	◯	◯	◯	◯	◯	◯	◯	◯
Empirical Validity	◯	◯	◯	◯	◯	◯	◯	◯	◯
Feasibility	◯	◯	◯	◯	◯	◯	◯	◯	◯

1.1.4D) Rise in diastolic blood pressure ≥ 20 mmHg

	1	2	3	4	5	6	7	8	9
Scientific Validity	◯	◯	◯	◯	◯	◯	◯	◯	◯
Empirical Validity	◯	◯	◯	◯	◯	◯	◯	◯	◯
Feasibility	◯	◯	◯	◯	◯	◯	◯	◯	◯

1.2 Increase in both systolic and diastolic blood pressure should be present.

	1	2	3	4	5	6	7	8	9
Scientific Validity	◯	◯	◯	◯	◯	◯	◯	◯	◯
Empirical Validity	◯	◯	◯	◯	◯	◯	◯	◯	◯
Feasibility	◯	◯	◯	◯	◯	◯	◯	◯	◯

1.3 In your opinion, what are the most appropriate cutoffs for high blood pressure?

1.3.1Absolute SBP

◯ 140 mmHg
◯ 150 mmHg
◯ 160 mmHg
◯ 170 mmHg
◯ 180 mmHg
◯ Other

1.3.1Other

_____

1.3.2Absolute DBP

◯ 90 mmHg
◯ 100 mmHg
◯ 110 mmHg
◯ 120 mmHg
◯ 130 mmHg
◯ Other

1.3.2Other

_____

1.3.3Increase in SBP

◯ 10 mmHg
◯ 20 mmHg
◯ 30 mmHg
◯ 40 mmHg
◯ 50 mmHg
◯ Other

1.3.3Other

_____

1.3.4Increase in DBP

◯ 10 mmHg
◯ 20 mmHg
◯ 30 mmHg
◯ 40 mmHg
◯ 50 mmHg
◯ Other

1.3.4Other

_____

1.4 In the absence of signs and symptoms, blood pressure should be measured on at least 2 occasions

Invalid/Unfeasible_____Uncertain_____Valid/Feasible

	1	2	3	4	5	6	7	8	9
Scientific Validity	◯	◯	◯	◯	◯	◯	◯	◯	◯
Empirical Validity	◯	◯	◯	◯	◯	◯	◯	◯	◯
Feasibility	◯	◯	◯	◯	◯	◯	◯	◯	◯

1.5 In your opinion, what are the most appropriate frequency and intervals for repeated measurements?

1.5.1Frequency

◯ Only once is enough
◯ 2 times
◯ 3 times
◯ 4 times
◯ Other

1.5.1Other

_____

1.5.2Intervals

◯ 12 hours apart
◯ 24 hours apart
◯ 48 hours apart
◯ 72 hours apart
◯ 1 week apart
◯ Other

1.5.2Other

_____

Comments:

_____

Renal Insufficiency (or Azotemia)

Please rate the scientific validity (based on the literature provided), empirical validity (based on your experience and knowledge of professional consensus) and feasibility (based on whether the information can be found in an average medical record and is likely to be reliable) of the following items.

2.1 Increase in serum creatinine ≥50% over baseline or, if no baseline available, serum creatinine ≥ 120% (or 1.2 times) the upper limit of normal for local laboratory (with measurement repeated if necessary to rule out lab error).

Invalid/Unfeasible_____Uncertain_____Valid/Feasible

	1	2	3	4	5	6	7	8	9
Scientific Validity	◯	◯	◯	◯	◯	◯	◯	◯	◯
Empirical Validity	◯	◯	◯	◯	◯	◯	◯	◯	◯
Feasibility	◯	◯	◯	◯	◯	◯	◯	◯	◯

2.2 In your opinion, what are the most appropriate cutoffs for increase in serum creatinine?

2.2.1 Increase above baseline

◯ 20%
◯ 30%
◯ 40%
◯ 50%
◯ 60%
◯ 70%
◯ 80%
◯ 90%
◯ 100% (doubling)
◯ Other

2.2.1 Other

_____

2.2.2 Increase above upper limit of normal for local laboratory

◯ 120% (1.2 times)
◯ 130% (1.3 times)
◯ 140% (1.4 times)
◯ 150% (1.5 times)
◯ 175% (1.75 times)
◯ 200% (double)
◯ Other

2.2.2 Other

_____

Comments:

_____

Proteinuria

Please rate the scientific validity (based on the literature provided), empirical validity (based on your experience and knowledge of professional consensus) and feasibility (based on whether the information can be found in an average medical record and is likely to be reliable) of the following items.

3.1 New proteinuria defined as ≥ 1+ (30-100 mg/dL range) by urine dipstick or worsening proteinuria defined as a ≥ 1 point increase in protein on urine dipstick (1+ to ≥ 2+, 2+ to ≥ 3+, etc).

Invalid/Unfeasible_____Uncertain_____Valid/Feasible

	1	2	3	4	5	6	7	8	9
Scientific Validity	◯	◯	◯	◯	◯	◯	◯	◯	◯
Empirical Validity	◯	◯	◯	◯	◯	◯	◯	◯	◯
Feasibility	◯	◯	◯	◯	◯	◯	◯	◯	◯

3.2 New proteinuria defined as ≥ 2+ (100-300 mg/dL range) by urine dipstick or worsening proteinuria defined as a ≥ 1 point increase in protein on urine dipstick (2+ to ≥ 3+, 3+ to ≥ 4+).

	1	2	3	4	5	6	7	8	9
Scientific Validity	◯	◯	◯	◯	◯	◯	◯	◯	◯
Empirical Validity	◯	◯	◯	◯	◯	◯	◯	◯	◯
Feasibility	◯	◯	◯	◯	◯	◯	◯	◯	◯

3.3 Proteinuria should be confirmed by urine protein:creatinine ratio

	1	2	3	4	5	6	7	8	9
Scientific Validity	◯	◯	◯	◯	◯	◯	◯	◯	◯
Empirical Validity	◯	◯	◯	◯	◯	◯	◯	◯	◯
Feasibility	◯	◯	◯	◯	◯	◯	◯	◯	◯

3.4 Proteinuria should be confirmed by 24-hour urine collection

	1	2	3	4	5	6	7	8	9
Scientific Validity	◯	◯	◯	◯	◯	◯	◯	◯	◯
Empirical Validity	◯	◯	◯	◯	◯	◯	◯	◯	◯
Feasibility	◯	◯	◯	◯	◯	◯	◯	◯	◯

3.5 In your opinion, what are the most appropriate cutoffs for new proteinuria?

3.5.1Dipstick

◯ 1+
◯ 2+
◯ 3+
◯ 4+
◯ Other

3.5.1Other

_____

3.5.2Urine protein:creatinine ratio

◯ ≥ 0.15 g/day
◯ ≥ 0.5 g/day
◯ ≥ 1.0 g/day
◯ ≥ 2.0 g/day
◯ Other

3.5.2Other

_____

3.6 In your opinion, what are the most appropriate cutoffs for worsening proteinuria?

3.6.1Dipstick

- ◯ a ≥ 1 point increase in protein on urine dipstick (1+ to ≥ 2+)
- ◯ a ≥ 2 point increase in protein on urine dipstick (1+ to ≥ 3+)
- ◯ Other

3.6.1Other

_____

3.6.2Urine protein:creatinine ratio

- ◯ Doubling
- ◯ Tripling
- ◯ Quadrupling
- ◯ Other

3.6.2Other

_____

Comments:

_____

Hematuria

Please rate the scientific validity (based on the literature provided), empirical validity (based on your experience and knowledge of professional consensus) and feasibility (based on whether the information can be found in an average medical record and is likely to be reliable) of the following items.

4.1 New hematuria defined as ≥ 1+ by urine dipstick or worsening hematuria defined as a ≥ 1 point increase on urine dipstick (1+ to ≥ 2+, 2+ to ≥ 3+, etc).

Invalid/Unfeasible_____Uncertain_____Valid/Feasible

	1	2	3	4	5	6	7	8	9
Scientific Validity	◯	◯	◯	◯	◯	◯	◯	◯	◯
Empirical Validity	◯	◯	◯	◯	◯	◯	◯	◯	◯
Feasibility	◯	◯	◯	◯	◯	◯	◯	◯	◯

4.2 New hematuria defined as ≥ 2+ by urine dipstick or worsening hematuria defined as a ≥ 1 point increase on urine dipstick (2+ to ≥ 3+, 3+ to ≥ 4+).

	1	2	3	4	5	6	7	8	9
Scientific Validity	◯	◯	◯	◯	◯	◯	◯	◯	◯
Empirical Validity	◯	◯	◯	◯	◯	◯	◯	◯	◯
Feasibility	◯	◯	◯	◯	◯	◯	◯	◯	◯

4.3 New hematuria defined as ≥ 10 RBCs/HPF on urine microscopy or worsening hematuria defined as a doubling of baseline hematuria on urine microscopy.

	1	2	3	4	5	6	7	8	9
Scientific Validity	◯	◯	◯	◯	◯	◯	◯	◯	◯
Empirical Validity	◯	◯	◯	◯	◯	◯	◯	◯	◯
Feasibility	◯	◯	◯	◯	◯	◯	◯	◯	◯

4.4 In your opinion, what are the most appropriate cutoffs for new hematuria?

4.4.1 Dipstick

◯ 1+
◯ 2+
◯ 3+
◯ 4+
◯ Other

Other

_____

4.4.2 Microscopy

◯ ≥ 10 RBCs/HPF
◯ ≥ 20 RBCs/HPF
◯ ≥ 30 RBCs/HPF
◯ ≥ 50 RBCs/HPF
◯ Other

Other

_____

4.5 In your opinion, what are the most appropriate cutoffs for worsening hematuria?

4.5.1Dipstick

- ◯ a ≥ 1 point increase in protein on urine dipstick (1+ to ≥ 2+)
- ◯ a ≥ 2 point increase in protein on urine dipstick (1+ to ≥ 3+)
- ◯ Other

Other

_____

4.5.2Microscopy

- ◯ Doubling
- ◯ Tripling
- ◯ Quadrupling
- ◯ Other

Other

_____

Comments:

_____

Thrombocytopenia

Please rate the scientific validity (based on the literature provided), empirical validity (based on your experience and knowledge of professional consensus) and feasibility (based on whether the information can be found in an average medical record and is likely to be reliable) of the following items.

5.1 ≤ 100,000 platelets/mm3

Invalid/Unfeasible_____Uncertain_____Valid/Feasible

	1	2	3	4	5	6	7	8	9
Scientific Validity	◯	◯	◯	◯	◯	◯	◯	◯	◯
Empirical Validity	◯	◯	◯	◯	◯	◯	◯	◯	◯
Feasibility	◯	◯	◯	◯	◯	◯	◯	◯	◯

5.2 In your opinion, what is the most appropriate cutoff for thrombocytopenia?

5.2.1Range 50,000 to 140,000 platelets/mm3

◯ 50 000 platelets/mm3
◯ 60 000 platelets/mm3
◯ 70 000 platelets/mm3
◯ 80 000 platelets/mm3
◯ 90 000 platelets/mm3
◯ 100 000 platelets/mm3
◯ 110 000 platelets/mm3
◯ 120 000 platelets/mm3
◯ 130 000 platelets/mm3
◯ 140 000 platelets/mm3
◯ Other

Other

_____

5.3 Thrombocytopenia should be confirmed by manual blood smear.

Invalid/Unfeasible_____Uncertain_____Valid/Feasible

	1	2	3	4	5	6	7	8	9
Scientific Validity	◯	◯	◯	◯	◯	◯	◯	◯	◯
Empirical Validity	◯	◯	◯	◯	◯	◯	◯	◯	◯
Feasibility	◯	◯	◯	◯	◯	◯	◯	◯	◯

Comments:

_____

Hemolysis (or Microangiopathic Hemolytic Anemia)

Please rate the scientific validity (based on the literature provided), empirical validity (based on your experience and knowledge of professional consensus) and feasibility (based on whether the information can be found in an average medical record and is likely to be reliable) of the following items.

6.1 MAHA defined as new or worsening anemia not due to other causes and supported by the presence of one of the following:

6.1.1A) Schistocytes or other RBC fragments on blood smear.

Invalid/Unfeasible_____Uncertain_____Valid/Feasible

	1	2	3	4	5	6	7	8	9
Scientific Validity	○	○	○	○	○	○	○	○	○
Empirical Validity	○	○	○	○	○	○	○	○	○
Feasibility	○	○	○	○	○	○	○	○	○

6.1.2B) Reticulocyte count above normal range for local laboratory.

	1	2	3	4	5	6	7	8	9
Scientific Validity	○	○	○	○	○	○	○	○	○
Empirical Validity	○	○	○	○	○	○	○	○	○
Feasibility	○	○	○	○	○	○	○	○	○

6.1.3C) Serum LDH and/or indirect bilirubin above normal ranges for local laboratory.

	1	2	3	4	5	6	7	8	9
Scientific Validity	○	○	○	○	○	○	○	○	○
Empirical Validity	○	○	○	○	○	○	○	○	○
Feasibility	○	○	○	○	○	○	○	○	○

6.1.4D) Serum haptoglobin below normal range for local laboratory.

	1	2	3	4	5	6	7	8	9
Scientific Validity	○	○	○	○	○	○	○	○	○
Empirical Validity	○	○	○	○	○	○	○	○	○
Feasibility	○	○	○	○	○	○	○	○	○

6.2 MAHA defined as new or worsening anemia not due to other causes and supported by the presence of at least two lab abnormalities (RBC fragments, elevated reticulocyte count, elevated serum LDH/indirect bilirubin, low haptoglobin).

	1	2	3	4	5	6	7	8	9
Scientific Validity	○	○	○	○	○	○	○	○	○
Empirical Validity	○	○	○	○	○	○	○	○	○
Feasibility	○	○	○	○	○	○	○	○	○

6.3 A direct Coombs test should be documented to rule out autoimmune hemolytic anemia.

	1	2	3	4	5	6	7	8	9
Scientific Validity	◯	◯	◯	◯	◯	◯	◯	◯	◯
Empirical Validity	◯	◯	◯	◯	◯	◯	◯	◯	◯
Feasibility	◯	◯	◯	◯	◯	◯	◯	◯	◯

Comments:

Encephalopathy

Please rate the scientific validity (based on the literature provided), empirical validity (based on your experience and knowledge of professional consensus) and feasibility (based on whether the information can be found in an average medical record and is likely to be reliable) of the following items.

7.1 Encephalopathy defined by the American Academy of Neurology as follows: 'Any diffuse disease of the brain that alters brain function or structure. The hallmark of encephalopathy is an altered mental state. Depending on the type and severity of encephalopathy, common neurological symptoms are progressive loss of memory and cognitive ability, subtle personality changes, inability to concentrate, lethargy, and progressive loss of consciousness. Other neurological symptoms may include myoclonus (involuntary twitching of a muscle or group of muscles), nystagmus (rapid, involuntary eye movement), tremor, muscle atrophy and weakness, dementia, seizures, and loss of ability to swallow or speak'.

Invalid/Unfeasible_____Uncertain_____Valid/Feasible

	1	2	3	4	5	6	7	8	9
Scientific Validity	○	○	○	○	○	○	○	○	○
Empirical Validity	○	○	○	○	○	○	○	○	○
Feasibility	○	○	○	○	○	○	○	○	○

Comments:

Retinopathy

Please rate the scientific validity (based on the literature provided), empirical validity (based on your experience and knowledge of professional consensus) and feasibility (based on whether the information can be found in an average medical record and is likely to be reliable) of the following items.

8.1 Retinopathy typical of malignant hypertension

Invalid/Unfeasible_____Uncertain_____Valid/Feasible

	1	2	3	4	5	6	7	8	9
Scientific Validity	○	○	○	○	○	○	○	○	○
Empirical Validity	○	○	○	○	○	○	○	○	○
Feasibility	○	○	○	○	○	○	○	○	○

8.2 Grade III (flame-shaped hemorrhages and/or "cotton-wool" exudates) or IV (papilledema) retinopathy, according to Keith-Wagener classification

	1	2	3	4	5	6	7	8	9
Scientific Validity	○	○	○	○	○	○	○	○	○
Empirical Validity	○	○	○	○	○	○	○	○	○
Feasibility	○	○	○	○	○	○	○	○	○

Comments:

Hyperreninemia

Please rate the scientific validity (based on the literature provided), empirical validity (based on your experience and knowledge of professional consensus) and feasibility (based on whether the information can be found in an average medical record and is likely to be reliable) of the following items.

9.1 Elevation of plasma renin activity ≥ 2 times the upper limit of normal

Invalid/Unfeasible_____Uncertain_____Valid/Feasible

	1	2	3	4	5	6	7	8	9
Scientific Validity	○	○	○	○	○	○	○	○	○
Empirical Validity	○	○	○	○	○	○	○	○	○
Feasibility	○	○	○	○	○	○	○	○	○

Comments:

Cardiac Dysfunction

Please rate the scientific validity (based on the literature provided), empirical validity (based on your experience and knowledge of professional consensus) and feasibility (based on whether the information can be found in an average medical record and is likely to be reliable) of the following items.

10.1 Presence of flash pulmonary edema based on all available information and clinical judgement.

Invalid/Unfeasible_____Uncertain_____Valid/Feasible

	1	2	3	4	5	6	7	8	9
Scientific Validity	◯	◯	◯	◯	◯	◯	◯	◯	◯
Empirical Validity	◯	◯	◯	◯	◯	◯	◯	◯	◯
Feasibility	◯	◯	◯	◯	◯	◯	◯	◯	◯

10.2 Presence of symptomatic pericardial effusion based on all available information and clinical judgement.

	1	2	3	4	5	6	7	8	9
Scientific Validity	◯	◯	◯	◯	◯	◯	◯	◯	◯
Empirical Validity	◯	◯	◯	◯	◯	◯	◯	◯	◯
Feasibility	◯	◯	◯	◯	◯	◯	◯	◯	◯

Comments:

Abnormal Kidney Biopsy

Please rate the scientific validity (based on the literature provided), empirical validity (based on your experience and knowledge of professional consensus) and feasibility (based on whether the information can be found in an average medical record and is likely to be reliable) of the following items.

11.1 Findings consistent with SRC (microangiopathy)

Invalid/Unfeasible_____Uncertain_____Valid/Feasible

	1	2	3	4	5	6	7	8	9
Scientific Validity	○	○	○	○	○	○	○	○	○
Empirical Validity	○	○	○	○	○	○	○	○	○
Feasibility	○	○	○	○	○	○	○	○	○

11.2 Accumulation of mucoid (myxoid) in interlobular arteries (indistinguishable form accelerated hypertension) and/or fibrinoid necrosis of arteries

	1	2	3	4	5	6	7	8	9
Scientific Validity	○	○	○	○	○	○	○	○	○
Empirical Validity	○	○	○	○	○	○	○	○	○
Feasibility	○	○	○	○	○	○	○	○	○

11.3 Proposed definition for findings consistent with SRC:

Histopathological findings on kidney biopsy consistent with SRC may include the following: small vessel (arcuate and interlobular arteries) changes predominate over glomerular alterations. Early vascular abnormalities include intimal accumulation of myxoid material, thrombosis, fibrinoid necrosis, sometimes resulting in cortical necrosis. Narrowing and obliteration of the vascular lumen lead to glomerular ischemia. Juxtaglomerular apparatus (JGA) hyperplasia, while relatively rare (10%), can be observed. Late changes are manifested by intimal thickening and proliferation (which lead to characteristic vascular "onion-skin" lesions), glomerulosclerosis and interstitial fibrosis. Since none of these findings are specific for SRC, the pathological diagnosis must be supported by appropriate clinical and serological data.

	1	2	3	4	5	6	7	8	9
Scientific Validity	○	○	○	○	○	○	○	○	○
Empirical Validity	○	○	○	○	○	○	○	○	○
Feasibility	○	○	○	○	○	○	○	○	○

Comments:

Thank you for completing the second round of the online delphi exercise. You will be notified when the third and final round is avaiable.

Appendix 4

Accepted manuscript for publication in *Arthritis and Rheumatology*

From Butler, E. , Baron, M. , Fogo, A. B., Frech, T. , Ghossein, C. , Hachulla, E. , Hoa, S. , Johnson, S. R., Khanna, D. , Mouthon, L. , Nikpour, M. , Proudman, S. , Steen, V. , Stern, E. , Varga, J. , Denton, C. , Hudson, M., Scleroderma Clinical Trials Consortium Scleroderma Renal Crisis Working Group , (2019), *Generation of a core set of items to develop classification criteria for Scleroderma renal crisis using consensus methodology*. Arthritis Rheumatol. Accepted Author Manuscript. doi:10.1002/art.40809.

For permissions see Appendix 5

DR. TRACY FRECH (Orcid ID : 0000-0002-5472-3840)
DR. SABRINA HOA (Orcid ID : 0000-0002-9466-5729)

Article type : Brief Report

GENERATION OF A CORE SET OF ITEMS TO DEVELOP CLASSIFICATION CRITERIA FOR SCLERODERMA RENAL CRISIS USING CONSENSUS METHODOLOGY

Emily-Ann Butler, BS (Hons)[1]; Murray Baron, MD[2,3]; Agnes B. Fogo, MD[4]; Tracy Frech, MD MS[5]; Cybele Ghossein, MD[6]; Eric Hachulla, MD[7]; Sabrina Hoa, MD[3,8]; Sindhu R. Johnson, MD PhD[9]; Dinesh Khanna, MD MS[10]; Luc Mouthon, MD PhD[11]; Mandana Nikpour, MD PhD[12]; Susanna Proudman, MB BS (Hons)[13]; Virginia Steen, MD[14]; Edward Stern, MBBS[15]; John Varga, MD[6]; Christopher Denton, PhD FRCP[15]; Marie Hudson, MD MPH[2,3,8]; Scleroderma Clinical Trials Consortium Scleroderma Renal Crisis Working Group*

Institutional affiliations: [1]Dalhousie University, Halifax, Canada; [2]Jewish General Hospital, Montreal, Canada; [3]McGill University, Montreal, Canada; [4]Vanderbilt University Medical Center, Nashville, USA; [5]University of Utah, Salt Lake City, USA; [6]Northwestern University - Feinberg School of Medicine, Chicago, USA; [7]University of Lille - Hôpital Claude Huriez, Lille, France; [8]Lady Davis Institute, Montreal, Canada; [9]Toronto Scleroderma Program, Division of Rheumatology, Department of Medicine, Toronto Western Hospital, Mount Sinai Hospital, University of Toronto, Toronto, Ontario, Canada; [10]University of Michigan, Ann Arbor, USA; [11]Department of Internal Medicine, Cochin Hospital, Paris-Descartes University, Paris, France; [12]University of Melbourne at St. Vincent's Hospital, Victoria, Australia; [13]Royal Adelaide Hospital and University of Adelaide, Adelaide,

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/art.40809

Australia; [14]Georgetown University, Washington DC, USA; [15]Royal Free Hospital, London, UK .

***Investigators of the Scleroderma Clinical Trials Consortium Scleroderma Renal Crisis Working Group:** April Barnado (Vanderbilt University Medical Center), Murray Baron (Jewish General Hospital), Elana J. Bernstein (Columbia University Medical Center), Francesco Boin (University of California San Francisco), Yolanda Braun-Moscovici (Rheumatology Institute, Rambam Health Care Campus, Rappaport Faculty of Medicine, Technion, Haifa, Israel), Flavia V. Castelino, M.D. (Massachusetts General Hospital, Harvard Medical School), Luis J. Catoggio (Hospital Italiano de Buenos Aires), Marco Matucci-Cerinic (University of Florence), Lorinda Chung (Palo Alto VA Health Care System, Stanford University), Philip Clements (University of California, Los Angeles), Mary Ellen Csuka (Medical College of Wisconsin), Ellen De Langhe (University Hospital Leuven, Rheumatology Department), Christopher Denton (Royal Free Hospital), Jörg Distler (Friedrich-Alexander-University, Erlangen-Nürnberg), Oliver Distler (University Hospital Zurich), Dominique Claire Farge (AP-HP, St Louis Hospital, Université Denis Diderot, France), Aryeh Fischer (University of Colorado), Agnes Fogo (Vanderbilt University), Tracy Frech (University of Utah, Salt Lake Veterans Affairs Medical Center), Armando Gabrielli (Università Politecnica delle Marche), Cybele Ghossein (Northwestern University - Feinberg School of Medicine), Eric Hachulla (National Referral Centre for Rare Systemic Auto-immune Diseases North and North-West of France, University of Lille, Huriez Hospital), Minoru Hasegawa (University of Fukui), Samina Hayat (Louisiana State University), Ariane Herrick (University of Manchester), Roger Hesselstrand (Lund University Hospital), Sabrina Hoa (McGill University), Vivien Hsu (Rutgers-RWJ Medical School), Marie Hudson (McGill University), Michael Hughes (University of Manchester), Nicolas Hunzelmann (University of Cologne), Laura Hummers (Johns Hopkins University), Florenzo Iannone (University of Bari), Francesca Ingegnoli (Università degli Studi di Milano), Soren Jacobsen (Rigshospitalet), Sindhu Johnson (University of Toronto), Yasushi Kawaguchi (Tokyo Women's Medical University), Dinesh Khanna (University of Michigan), Martial

Koenig (University of Montreal), Masataka Kuwana (Nippon Medical School), Jan Lenaerts (University Hospital Leuven, Rheumatology Department), Thierry Martin (National referral center for systemic autoimmune diseases, Strasbourg University Hospital), Maureen D. Mayes (University of Texas McGovern Medical School, Houston), Zsuzsanna McMahan (Johns Hopkins University), Thomas Medsger (University of Pittsburgh), Peter Merkel (University of Pennsylvania), Sonali Narain (Northwell Health), Mandana Nikpour (University of Melbourne, St. Vincent's Hospital), Voon Ong (University College London), John D. Pauling (Royal National Hospital for Rheumatic Diseases , University of Bath), Janet Pope (University of Western Ontario, St. Joseph's Health Care), Susanna Proudman (Royal Adelaide Hospital, University of Adelaide), Carlos de la Puente Bujidos (Hospital Ramón y Cajal), Maureen Rischmueller (Queen Elizabeth Hospital, University of Adelaide), Tatiana Sofía Rodríguez-Reyna (Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán), Joanne Sahhar (Monash Health, Monash University), Lesley Ann Saketkoo (New Orleans Scleroderma and Sarcoidosis Patient Care and Research Center, Tulane University Lung Center), Jean-Luc Senécal (University of Montreal Scleroderma Research Chair, Centre Hospitalier de l'Université de Montréal), Ankoor Shah (Duke University), Ami A. Shah (Johns Hopkins University), Walter Alberto Sifuentes-Giraldo (Ramón y Cajal University Hospital), Richard Silver (Medical University of South Carolina), Virginia Steen (Georgetown University Medical Center), Edward Stern (University College London), Wendy Stevens (St. Vincent's Hospital), Evelyn Sutton (Dalhousie University), Vivek Thakkar (Macquarie University), Gabriele Valentini (Università degli Studi della Campania), Jeska de Vries-Bouwstra (Leiden University), Madelon Vonk (Radboud University), Ulrich A. Walker (Basel University Hospital)

Funding: This study was funded by a grant from the Scleroderma Clinical Trials Consortium. Dr Hudson is funded as a Clinician Scientist by the Fonds de Recherche du Québec – Santé. Dr Stern was supported by MRC grant MR/K015230/1. The funding sources had no role in the design and conduct

of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Corresponding author: Marie Hudson, Jewish General Hospital, 3755 Côte Ste Catherine Road, Montreal, Quebec, Canada, H3T 1E4; tel. (514) 340-8222 ext. 23476; E-mail: marie.hudson@mcgill.ca

Abstract

Background: This project was undertaken to generate a core set of items to develop classification criteria for scleroderma renal crisis (SRC) using consensus methodology.

Methods: An international, multidisciplinary panel of experts was invited to participate in a 3-round Delphi exercise developed using a survey based on items identified by a scoping review. In Round 1, participants were asked to identify omissions and clarify ambiguities regarding the items in the survey. In Round 2, participants were asked to rate the validity and feasibility of the items using Likert-type scales ranging from 1-9 (1= very invalid/unfeasible, 5 = uncertain, 9 = very valid/feasible). In Round 3, participants reviewed the results and comments of Round 2, and were asked to provide final ratings. Items rated as highly valid and feasible (both median scores ≥ 7) in Round 3 were selected as the provisional core set of items. A consensus meeting using nominal group technique (NGT) followed to further reduce the core set of items.

Results: Ninety-nine experts from 16 countries participated in the Delphi exercise. Of the 31 items in the survey, consensus was achieved on 13, including hypertension, renal insufficiency, proteinuria and hemolysis. Eleven experts took part in the NGT discussion, where consensus was achieved in 5

domains: blood pressure, acute kidney injury, microangiopathic hemolytic anemia, target organ dysfunction, and renal histopathology.

Conclusions: A core set of items that characterize SRC was identified using consensus methodology. This core set will be used in future data-driven phases of this project to develop classification criteria for SRC.

Introduction

Scleroderma renal crisis (SRC) is a life-threatening complication of systemic sclerosis (SSc) (1–4). It is usually characterized by malignant hypertension and acute kidney injury (3). However, the clinical spectrum of SRC is broad, ranging from full-blown disease presenting as new onset accelerated arterial hypertension and rapidly progressive oliguric renal failure, to more modest elevations in blood pressure and renal dysfunction, and at times normotensive presentations. On the other hand, hypertension without uraemia, urinary abnormalities and/or mild uraemia attributable to other factors (e.g., concomitant comorbidities such as diabetes or exposure to nephrotoxic medications) are common in SSc (4,5). These conditions should not be confused with SRC.

SRC is relatively rare, occurring in about 5% of all SSc patients (3). It is more common in patients with rapidly progressing diffuse cutaneous SSc (dcSSc) (11%) as compared to patients with limited cutaneous SSc (lcSSc) (4%) (6). SRC can be further sub-categorized into hypertensive or normotensive forms, representing approximately 90% and 10% of SRC cases, respectively (7,8). Historically, SRC was the leading cause of death in SSc (9). However, with the advent of angiotensin converting enzyme (ACE) inhibitors, mortality rates have decreased significantly (10,11). Nevertheless, one-year outcomes remain poor, with over 30% mortality and 25% of patients remaining dialysis-dependent (12). There is an urgent need to undertake research to identify novel treatments and to improve outcomes of SRC.

In addition to heterogeneity and rarity, the absence of a gold standard and classification criteria are important challenges for research on SRC. To date, most studies of SRC have used *ad hoc* criteria that have varied considerably from study to study. In a scoping review of the literature, 40 original definitions of SRC, with significant heterogeneity among them, were identified (13). Only one study to date has partially validated criteria for SRC (12).

The Scleroderma Clinical Trials Consortium (SCTC) SRC Working Group was created to develop classification criteria for SRC. The objective of this phase of the study was to generate a core set of items to define SRC using consensus methodology. Future studies using data-driven methods will be required to develop and validate classification criteria for SRC.

Methods

A scoping review of the literature to identify items used to define SRC has been published (13). The results of this review were used to inform this project, which consisted of two phases: 1) a modified online Delphi exercise to develop provisional consensus on a core set of items to define SRC and 2) a consensus meeting using nominal group technique (NGT) to further reduce the core set. Ethics approval for this project was obtained from the Jewish General Hospital Research Ethics Board, Montréal, Quebec, Canada (Protocol # CODIM-MBM-17-104).

Phase 1: Delphi

A modified, online, 3-round Delphi exercise was conducted (14,15). Experts from the SCTC, European Scleroderma Trials and Research Group (EUSTAR), Canadian Scleroderma Research Group (CSRG) and Australian Scleroderma Interest Group (ASIG) were invited to participate. In addition, pathologists and nephrologists known through these organizations with interest in SRC were also invited to participate. Individuals interested in participating were asked to accept the invitation by

return email. All individuals who accepted were then considered study participants, and thereby constituted the denominator for the participation rates.

The Delphi survey was developed and managed through the REDCap platform (Vanderbilt University, Nashville, Tennessee). In Round 1, consent to participate was obtained and demographic and personal information was collected on participants. Subsequently, Round 1 asked participants to consider the items identified in the scoping review and requested them to clarify ambiguities, identify omissions and provide comments. Items were modified accordingly.

In Round 2, participants were asked to rate the scientific validity, empirical validity and feasibility of the items using Likert-type scales ranging from 1-9 (1= very invalid/unfeasible, 5 = uncertain, 9 = very valid/feasible) and to provide comments. Participants were provided links to full-text copies of the scoping review and all of the papers included therein. Scientific validity was defined as items supported by published literature and empirical validity as items supported by personal experience and knowledge of professional consensus. Feasibility was defined in terms of whether the item could be performed/tested in an easy or convenient matter.

In Round 3, the results of Round 2 were presented using summary statistics, including medians and interquartile ranges, and bar graphs. Participants were also shown their answers and anonymized comments from other participants from Round 2. The participants were then asked to provide their final rating on scientific validity, empirical validity and feasibility of the items.

Consensus was defined as items rated highly scientifically valid and feasible (both median scores $\geq$ 7) in Round 3, and for which there was no disagreement, calculated using the RAND/UCLA Appropriateness Method formula. Disagreement exists when the inter-percentile range (IPR: difference between the 30^{th} and 70^{th} percentiles) is larger than the IPR adjusted for symmetry (IPRAS), calculated as follows:

IPRAS = 2.35 + [Asymmetry Index x 1.5]

Derivation of the formula is shown in the RAND/UCLA Appropriateness Method handbook (16).

Phase 2: NGT meeting

The second phase of this study was to reduce the number of items and achieve consensus using NGT (17). International experts, including rheumatologists, internists and nephrologists, were invited to participate in a 2-hour face-to-face meeting held in November 2017 in San Diego (California, USA). Dr. Dinesh Khanna moderated the discussion based on expertise and previous experience in the fields of SRC and NGT techniques (17,18). Each item from the Delphi was discussed in turn. Each panelist was invited to provide comments. At the end of the discussion, the panelists were asked to vote by a show of hands if the items should be included in the core set. A simple majority was required to include the item.

During the NGT meeting, it became clear that some items required content expertise beyond rheumatology, internal medicine and nephrology. Thus, some items were conditionally included, pending further review with content experts. Experts in hematology, neurology, ophthalmology, and cardiology were then contacted and asked to provide input and published evidence to define items in those domains.

A final list of core set items (and their definitions) was compiled and circulated among the participants of the NGT meeting for final approval.

Results

Phase 1: Delphi

We contacted 216 people with an interest in SRC of which 99 agreed to participate in the modified online Delphi exercise. Of those, 77 (78%), 60 (61%) and 69 (70%) participated in Rounds

1, 2 and 3, respectively, and 49 (49%) completed all three rounds of the exercise. Participants were mainly rheumatologists (86%) with some internists, nephrologists and pathologists. Most participants worked as clinicians for >11 years, with only a few having less than 10 years of experience (13%). The majority of participants were from the United States (35%) followed by Canada (11%); 16 other countries were also represented.

A total of 31 items in 11 categories were included in the Delphi exercise. Of these, 13 items in 4 categories (hypertension, renal insufficiency, proteinuria and hemolysis) achieved consensus in Round 3 (median ratings $\geq$ 7 on scientific validity and feasibility with no disagreement). Disagreement on feasibility was only present for hyper-reninemia. In any case, that item had not achieved consensus on feasibility either. Of note, all items that reached consensus in Round 2, also reached consensus in Round 3 with no additional items reaching consensus in Round 3. However, the IQR for the majority of items became smaller in Round 3, demonstrating growing consensus. The median ratings and IQR for each item for Rounds 2 and 3 are presented in Table 1.

Phase 2: Nominal Group Technique meeting

Seventeen international experts were invited to participate in a face-to-face NGT meeting. Six were not available. Thus, the panel consisted of 11 participants, 10 rheumatologists and 1 nephrologist, from the USA, Canada, United Kingdom, France, Netherlands and Australia. Prior to the NGT meeting, the 11 categories from the Delphi exercise were re-organized into 5 domains (hypertension, renal dysfunction [renal insufficiency, proteinuria, hematuria and hyper-reninemia], microangiopathic hemolytic anemia with thrombocytopenia, target organ dysfunction [encephalopathy, retinopathy and cardiac dysfunction] and renal histopathology). Prior to and at the meeting, it was agreed that items should be defined as much as possible according to evidence and/or international guidelines. Content experts in hematology, neurology, ophthalmology, and cardiology were contacted to provide input on definitions of items included in the core set.

The final core set of items and their definitions are presented in Table 2, and were approved by the NGT participants.

Discussion

In this study, we generated a core set of items to classify SRC using consensus methodology. This core set includes 5 domains and 14 items. The definitions for each item were evidence-based or, in the absence of evidence, determined in consultation with content experts.

The progress made to date to develop classification criteria for SRC demonstrates the importance of using the best evidence available. A scoping review of the literature identified 40 heterogeneous definitions of SRC using more than 40 items with variable definitions (13). The Delphi exercise led to consensus on 13 of these items. However, the need to go beyond consensus in the rheumatology community and to get the input of content experts emerged as a critical factor at the NGT meeting. Thus, the input from content experts was sought to finalize the core set. Proteinuria is a perfect example of how this approach allowed the core set to evolve. Indeed, low-level proteinuria is common in SSc (4), dipstick and urine protein-to-creatinine ratio are not reliable in AKI, proteinuria is not part the Kidney Disease Improving Global Outcomes (KDIGO) definition of AKI (19), and proteinuria would compromise specificity of SRC criteria. Thus, despite the fact that there was consensus to include proteinuria in the core set after the Delphi exercise, this item was excluded after the NGT meeting and discussion with nephrologists.

A core set of variables to define SRC was proposed by experts in 2003 (7). It included items for systolic and diastolic blood pressure, serum creatinine, proteinuria, hematuria, microangiopathic hemolytic anemia and renal histopathology. These are known as the Ancona criteria for SRC. Our core set has similarities to the Ancona criteria in particular with respect to blood pressure. However, there are also notable differences in defining acute kidney injury (including the exclusion of

proteinuria and hematuria). In addition, our core set includes target organ dysfunction and a detailed histopathological description of SRC.

In 2016, the UK Scleroderma Study Group proposed criteria for the diagnosis of SRC (20). The criteria were divided into categories: diagnostic criteria (essential) and supportive evidence (desirable) with blood pressure and AKI as the former, MAHAT, hypertensive retinopathy, hematuria, oliguria or anuria, renal biopsy consistent with SRC features and flash pulmonary edema as the latter. Discrepancies with our proposed criteria are found in the slightly modified cut-off values for blood pressure (150/85 mmHg versus 140/90 mmHg) and additionally, there is no noted rise in diastolic blood pressure, only $\geq$20 mmHg for systolic blood pressure which is lower than $\geq$30 mmHg proposed in this study. Further, the UK criteria included hematuria. Additionally, oliguria and flash pulmonary edema were proposed as stand-alone items whereas in our list, these items are grouped into the AKI and acute heart failure definitions, respectively. Our core set provides a more in depth detailed definition for each item, specifically for AKI, MAHAT and renal histopathology.

Only one study to date has attempted to validate the Ancona criteria and another slightly different set of criteria for SRC that included encephalopathy (12). In that study, a diagnosis of SRC confirmed by a study physician was used as the gold standard for SRC. Compared to the gold standard, the two sets of criteria identified 70/70 subjects with hypertensive, but only 2/5 subjects with normotensive SRC. We believe that our core set, which was developed using robust consensus methodology and evidence-based content, represents a significant advancement over these definitions. In addition, it defines target organ involvement and provides a detailed histopathological description to define the term "findings consistent with SRC".

This study has some limitations. First, only 99/216 experts invited to participate accepted and 77 (78%), 60 (61%) and 69 (70%) of these participated in Rounds 1-3 of the Delphi, respectively. We cannot exclude some response bias. Part of the reason for the low response rates may have been that the Delphi exercise was conducted during the summer and early fall in the Northern hemisphere. Numerous out of office replies were returned. On the other hand, to mitigate this source of bias,

reminder emails were sent to optimize participation rates and the final sample was still substantial and representative. Second, there are large gaps in knowledge on SRC. Hence, participants in the Delphi may have rated validity based more on empirical, rather than on scientific evidence. Nevertheless, we provided the Delphi participants with the scoping review and all of the original papers included therein in every Round for easy access to the available literature. Third, recruitment of participants with a broad range of expertise is critical to the success of a consensus-building exercise. Although there were a few specialists other than rheumatologists who participated in the Delphi, it became clear at the NGT meeting that content expertise in hematology, neurology, ophthalmology, and cardiology was lacking. We therefore recruited experts in all of these fields to help finalize the relevant items.

This study has substantial strengths. The emphasis on evidence and input from content experts ensured that the final core set had face and content validity. The geographic range of participants contributed to the generalizability of the results. There was important complementarity in the use of both a Delphi exercise and a semi-structured NGT consensus meeting. The Delphi provided a cost-effective approach to survey a larger sample of international experts working anonymously. The NGT meeting allowed for a time-efficient, face-to-face discussion of a smaller sample of experts led by an experienced moderator.

Conclusion and future steps

In conclusion, using consensus methodology, we generated a core set of items, and the definition of those items, to be used in the development of classification criteria for SRC. To determine if and how these items should be incorporated into classification criteria for SRC, two future phases of this research project are now in planning. The first, modeled on the *International Scleroderma Renal Crisis Survey* (12), will be to recruit an inception SRC cohort and collect the items in the core set. A comparison cohort consisting of subjects with conditions that mimic SRC will also

be assembled. These data will be used to develop and validate classification criteria for SRC. The second will be a forced choice study using multi-criteria decision analysis methods to assign weights to the items in the criteria and to set probability values for definite, probable and possible SRC. The resulting classification criteria will facilitate rigorous research in SRC. In the meantime, SSc researchers who are designing new studies (either observational or trials) are encouraged to collect these items in their datasets. These will be useful for future external validation of the criteria.

Acknowledgements

We would like to thank all the participants of the Delphi survey, the participants of the Nominal Group Meeting (Murray Baron, Mary Ellen Csuka, Jeska de vries Bouwstra, Christopher Denton, Tracy Frech, Cybele Ghossein, Sindhu Johnson, Luc Mouthon, Mandana Nikpour, Richard Silver, Virginia Steen) and the content experts (hematology: Mark Blotstein, Stephen Caplan, Nathalie Johnson; cardiology: Vartan Mardigyan, Richard Sheppard; neurology: Robert Altman, Colin Chalk, Rami Massie; ophthalmology: Julius Gomolin) who contributed to this study.

References

1. Denton CP, Khanna D. Systemic sclerosis. Lancet. 2017;390:1685–99.
2. Helfrich DJ, Banner B, Steen VD, Medsger TA. Normotensive renal failure in systemic sclerosis. Arthritis Rheum. 1989;32:1128–34.
3. Mouthon L, Bérezné A, Bussone G, Noël LH, Villiger PM, Guillevin L. Scleroderma renal crisis: A rare but severe complication of systemic sclerosis. Clin Rev Allergy Immunol. 2011;40:84–91.
4. Steen VD, Syzd A, Johnson JP, Greenberg A, Medsger TA. Kidney disease other than renal

crisis in patients with diffuse scleroderma. J Rheumatol. 2005;32:649–55.

5. Caron M, Hudson M, Baron M, Nessim S, Steele R. Longitudinal Study of Renal Function in Systemic Sclerosis. J Rheumatol. 2012;39:1829–34.

6. Nihtyanova SI, Schreiber BE, Ong VH, Rosenberg D, Moinzadeh P, Coghlan JG, et al. Prediction of pulmonary complications and long-term survival in systemic sclerosis. Arthritis Rheumatol. 2014;66:1625–35.

7. Steen VD, Mayes MD, Merkel PA. Assessment of kidney involvement. Clin Exp Rheumatol. 2003;21:S29-S31.

8. Steen VD. Scleroderma and renal crisis. Rheum Dis Clin North Am. 2003;29:315–33.

9. Traub YM, Shapiro AP, Rodnan GP, Medsger TA, McDonald RH, Steen VD, et al. Hypertension and renal failure (scleroderma renal crisis) in progressive systemic sclerosis. Medicine (Baltimore). 1983;62:335–52.

10. Guillevin L, Berezne A, Seror R, Teixeira L, Pourrat J, Mahr A, et al. Scleroderma renal crisis: A retrospective multicentre study on 91 patients and 427 controls. Rheumatology. 2012;51:460–7.

11. Teixeira L, Mouthon L, Mahr A, Berezné A, Agard C, Mehrenberger M, et al. Mortality and risk factors of scleroderma renal crisis: a French retrospective study of 50 patients. Ann Rheum Dis. 2008;67:110–6.

12. Hudson M, Baron M, Tatibouet S, Furst DE, Khanna D, Hummers L, et al. Exposure to ACE inhibitors prior to the onset of scleroderma renal crisis-results from the international scleroderma renal crisis survey. Semin Arthritis Rheum. 2014;43:666–72.

13. Hoa S, Stern EP, Denton CP, Hudson M, Baron M, Frech T, et al. Towards developing criteria for scleroderma renal crisis: A scoping review. Autoimmun Rev. 2017;16:407–15.

14. Barber CEH, Marshall DA, Alvarez N, Mancini GBJ, Lacaille D, Keeling S, et al.

Development of Cardiovascular Quality Indicators for Rheumatoid Arthritis: Results from an International Expert Panel Using a Novel Online Process. J Rheumatol. 2015;42:1548–55.

15. Schmajuk G, Hoyer BF, Aringer M, Johnson SR, Daikh DI, Dorner T, SLE classification criteria steering committee and the international SLE expert panel of the initiative. Multi-center Delphi Exercise Reveals Important Key Items for Classifying Systemic Lupus Erythematosus. Arthritis Care and Research. 2018; 70:1488-1494.

16. Fitch K, Bernstein SJJ, Aguilar MDD, Burnand B, LaCalle JRR, Lazaro P, et al. The RAND / UCLA Appropriateness Method User's Manual. RAND. 2001. 109 p.

17. Fransen J, Johnson SR, Van Den Hoogen F, Baron M, Allanore Y, Carreira PE, et al. Items for developing revised classification criteria in systemic sclerosis: results of a consensus exercise with the ACR/EULAR working committee for classification criteria in systemic sclerosis. 2012;64:351–7.

18. Johnson SR, Khanna D, Cervera R, Costedoat-Chalumeau N, Gladman, DD, Hahn BH, Hiepe F, Sanchez-Guerrero J, Massarotti E, Boumpas, Costenbader KH, Daikh D, Jayne D, Dörner T, Kamen D, Mosca M, Ramsey-Goldman R, Smolen JS, Wofsy D, Aringer M. Use of Consensus Methodology to Determine Candidate Items for Systemic Lupus Erythematosus Classification Criteria. Journal of Rheumatology. 2018. *In press.*

19. Kellum J, Lameire N, Aspelin P, Barsoum RS, Burdmann E, Goldstein SL, et al. KDIGO Clinical Practice Guideline for Acute Kidney Injury. Kidney Int Suppl. 2012;2:1–138.

20. Lynch BM, Stern EP, Ong VH, Harber M, Burns A, Denton CP. UK Scleroderma Study Group (UKSSG) Guidelines on the Diagnosis and Management of Scleroderma Renal Crisis. Clin Exp Rheumatol. 2016;34 Suppl 100:106-9

Table 1. Results from Rounds 2 and 3 of the Delphi exercise and consensus achieved after Round 3.

Criteria Category	Question		Round 2		Round 3		Consensus
			Scientific Validity	Feasibility	Scientific Validity	Feasibility	
Hypertension	New onset or deterioration of pre-existing hypertension, defined as any of the following:	Systolic blood pressure ≥ 140 mmHg	7(2)*	8(2)	7(1)	8(1)	yes
		Diastolic blood pressure ≥ 90 mmHg	7(2)	8(1)	7(0.5)	8(1)	yes
		Rise in systolic blood pressure ≥ 30 mmHg	7(2)	8(1)	7(1)	8(1)	yes
		Rise in diastolic blood pressure ≥ 20 mmHg	7(2)	8(2)	7(1)	8(0)	yes
	Increase in both systolic and diastolic blood pressure should be present.		6(3)	8(2)	6(2)	8(0.5)	no
	In the absence of signs and symptoms, blood pressure measurements should be measured on at least 2 occasions.		7(3)	8(1)	7(1)	8(1)	yes
Renal insufficiency	Increase in serum creatinine ≥50% over baseline or, if no baseline available, serum creatinine ≥120% (or 1.2 times) the upper limit of normal for local laboratory (with measurement repeated if necessary to rule out lab error).		7(2)	8(2)	7(1)	8(1)	yes
Proteinuria	New proteinuria defined as ≥ 1+ (30-100 mg/dL range) by urine dipstick or worsening proteinuria defined as a ≥ 1 point increase in protein on urine (1+ to ≥ 2+, 2+ to ≥ 3+, etc).		5(2)	7(2)	5(1)	7(1)	no
	New proteinuria defined as ≥ 2+ (100-300 mg/dL range) by urine dipstick or worsening proteinuria defined as a ≥ 1 point increase in protein on urine (2+ to ≥ 3+, 3+ to ≥ 4+, etc).		7(2)	8(1)	7(1)	8(1)	yes
	Proteinuria should be confirmed by urine protein:creatinine ratio.		7(2)	8(2)	7(1)	8(0)	yes
	Proteinuria should be confirmed by 24-hour urine collection.		6(4)	6(3)	6(2)	6(2)	no
Hematuria	New hematuria defined as ≥ 1+ by urine dipstick or worsening hematuria defined as a ≥ 1 point increase on urine dipstick (1+ to ≥ 2+, 2+ to ≥ 3+, etc).		6(3)	8(1)	6(1)	8(1)	no
	New hematuria defined as ≥ 2+ by urine dipstick or worsening hematuria defined as a ≥ 1 point increase on urine dipstick (2+ to ≥ 3+, 3+ to ≥ 4+, etc).		6(3)	8(1)	6(1)	8(1)	no
	New hematuria defined as ≥ 10 red blood cells per high powered field on urine microscopy or worsening hematuria defined as a doubling of baseline hematuria on urine microscopy.		6(2)	7(2)	6(2)	7(1)	no
Thrombocytopenia	≤ 100,000 platelets/mm3		6(3)	8(1)	6(1)	8(1)	no
	Thrombocytopenia should be confirmed by manual blood smear.		6(2)	6(2)	6(2)	6(1)	no
Hemolysis	Microangiopathic hemolytic anemia defined as new or worsening anemia not due to other causes and supported by the presence of one of the following:	Schistocytes or other red blood cell fragments on blood smear.	8(1)	8(1)	8(0)	8(0)	yes
		Reticulocyte count above normal range for local laboratory.	7(3)	7(1)	7(1)	7(1)	yes
		Serum lactate dehydrogenase and/or indirect bilirubin above normal ranges for local laboratory.	6(2)	8(2)	6(1)	8(1)	no
		Serum haptoglobin below normal range for local laboratory.	7(2)	8(2)	7(1)	8(1)	yes
	Microangiopathic hemolytic anemia defined as new or worsening anemia not due to other causes and supported by the presence of at least two lab abnormalities (red blood cell fragments, elevated reticulocyte count, elevated serum lactate dehydrogenase/indirect bilirubin, low haptoglobin).		8(1)	8(1)	8(0)	8(0)	yes
	A direct anti-globulin test should be documented to rule out autoimmune hemolytic anemia.		7(3)	7(2)	7(0)	7(1)	yes

* Median values (inter-quartile range)

Table 1. Results from Rounds 2 and 3 of the Delphi exercise and consensus achieved after Round 3 - Continued

Criteria Category	Question	Round 2		Round 3		Consensus
		Scientific Validity	Feasibility	Scientific Validity	Feasibility	
Encephalopathy	Encephalopathy defined by the American Academy of Neurology as follows: 'Any diffuse disease of the brain that alters brain function or structure. The hallmark of encephalopathy is an altered mental state. Depending on the type and severity of encephalopathy, common neurological symptoms are progressive loss of memory and cognitive ability, subtle personality changes, inability to concentrate, lethargy, and progressive loss of consciousness. Other neurological symptoms may include myoclonus (involuntary twitching of a muscle or group of muscles), nystagmus (rapid, involuntary eye movement), tremor, muscle atrophy and weakness, dementia, seizures, and loss of ability to swallow or speak'.	6(3)*	7(2)	6(1)	7(1)	no
Retinopathy	Retinopathy typical of malignant hypertension	7(2)	6(3)	7(1)	6(1)	no
	Grade III (flame-shaped hemorrhages and/or "cotton-wool" exudates) or IV (papilledema) retinopathy, according to Keith-Wagener classification	7(3)	6(3)	7(1)	6(2)	no
Hyperreninemia	Elevation of plasma renin activity ≥ 2 times the upper limit of normal	7(3)	4(4)	7(1)	5(2)	no
Cardiac dysfunction	Presence of flash pulmonary edema based on all available information and clinical judgement.	6(2)	7(2)	6(1)	7(0)	no
	Presence of symptomatic pericardial effusion based on all available information and clinical judgement.	6(2)	6(2)	6(1)	6(1)	no
Abnormal kidney biopsy	Findings consistent with scleroderma renal crisis (microangiopathy)	8(2)	6(4)	8(0)	6(2)	no
	Accumulation of mucoid (myxoid) in interlobular arteries (indistinguishable from accelerated hypertension) and/or fibrinoid necrosis of arteries	7(2)	6(4)	7(1)	6(2)	no
	Histopathological findings on kidney biopsy consistent with SRC may include the following: small vessel (arcuate and interlobular arteries) changes predominate over glomerular alterations. Early vascular abnormalities include intimal accumulation of myxoid material, thrombosis, fibrinoid necrosis, sometimes resulting in cortical necrosis. Narrowing and obliteration of the vascular lumen lead to glomerular ischemia. Juxtaglomerular apparatus hyperplasia, while relatively rare (10%), can be observed. Late changes are manifested by intimal thickening and proliferation (which lead to characteristic vascular "onion-skin" lesions), glomerulosclerosis and interstitial fibrosis. Since none of these findings are specific for scleroderma renal crisis, the pathological diagnosis must be supported by appropriate clinical and serological data.	8(2)	6(3)	8(0)	6(2)	no

* Median values (inter-quartile range)

Table 2. Final core set of items to develop classification criteria for SRC

Domain	Item	
Blood pressure	Acute rise in blood pressure defined as any of the following:	Systolic blood pressure ≥ 140 mmHg Diastolic blood pressure ≥ 90mmHg A rise in systolic blood pressure ≥ 30 mmHg above normal A rise in diastolic blood pressure ≥ 20 mmHg above normal
	Blood pressure measurement should be taken twice separated by at least 5 minutes. If blood pressure readings are discordant, repeat readings should be obtained until 2 consistent readings are obtained.	
Kidney injury*	Acute kidney injury defined as any of the following:	Increase in serum creatinine by ≥ 26.5 umol/L (≥ 0.3 mg/dl) within 48 hours Increase in serum creatinine to ≥1.5 times baseline, which is known or presumed to have occurred within the prior 7 days Urine volume < 0.5 ml/kg/h for 6 hours
Microangiopathic hemolytic anemia and thrombocytopenia	New or worsening anemia not due to other causes. Schistocytes or other red blood cell fragments on blood smear. Thrombocytopenia ≤ 100,000, confirmed by manual smear. Laboratory evidence of hemolysis, including elevated lactate dehydrogenase, reticulocytosis and/or low/absent haptoglobin A negative direct anti-globulin test.	
Target organ dysfunction	*Hypertensive retinopathy* (hemorrhages, hard and soft (cotton wool) exudates, and/or disc edema, not attributable to other causes), confirmed by an ophthalmologist. *Hypertensive encephalopathy*, characterized by headache, altered mental status, seizures, visual disturbances and/or other focal or diffuse neurologic signs not attributable to other causes. *Acute heart failure*, characterized by typical symptoms (e.g. breathlessness, ankle swelling and fatigue) that may be accompanied by signs (e.g. elevated jugular venous pressure, pulmonary crackles and peripheral edema). *Acute pericarditis*, diagnosed with at least 2 of the 4 following criteria: 1) pericarditis chest pain; 2) pericardial rub; 3) new widespread ST-elevation or PR depression on electrocardiogram; 4) pericardial effusion (new or worsening) on cardiac echocardiography.	
Renal histopathology	Histopathological findings on kidney biopsy consistent with scleroderma renal crisis which may include the following: small vessel (arcuate and interlobular arteries) changes that predominate over glomerular alterations. Glomerular changes of thrombotic microangiopathy may be present, with acute changes including fibrin thrombi and endothelial swelling, red blood cell fragments and mesangiolysis, and chronic changes including double contours of the glomerular basement membrane. Nonspecific ischemic changes with corrugation of the glomerular basement membrane, and even segmental or global sclerosis of glomeruli may occur. Early vascular abnormalities include intimal accumulation of myxoid material, thrombosis, fibrinoid necrosis, fragmented red blood cells, sometimes resulting in cortical necrosis. Narrowing and obliteration of the vascular lumen lead to glomerular ischemia. Juxtaglomerular apparatus hyperplasia, while relatively rare (10%), can be observed. Late changes are manifested by intimal thickening and proliferation (which lead to characteristic vascular "onion-skin" lesions), glomerulosclerosis and interstitial fibrosis. Nonspecific tubular changes may also occur, including acute tubular injury in the early stage of injury, and later interstitial fibrosis and tubular atrophy. Since none of these findings are specific for SRC, the pathological diagnosis must be supported by appropriate clinical and serological data.	

*This is the definition of acute kidney injury from the Kidney Disease Improving Global Outcomes (KDIGO) guidelines (19)

Appendix 5

Wiley Global Permissions

Dear Emily-Ann,

Thank you for your email.

Permission is granted for you to use the material requested for your thesis/dissertation subject to the usual acknowledgements (author, title of material, title of book/journal, ourselves as publisher) and on the understanding that you will reapply for permission if you wish to distribute or publish your thesis/dissertation commercially.

You should also duplicate the copyright notice* that appears in the Wiley publication in your use of the Material. Permission is granted solely for use in conjunction with the thesis, and the material may not be posted online separately.

Any third-party material is expressly excluded from this permission. If any material appears within the article with credit to another source, authorisation from that source must be obtained.

Should you require any further authorisation, please do not hesitate to contact me.

Kind regards,

Wiley
www.wiley.com