Qualitative exploration of the renal stone patients' experience and development of the renal stone-specific Patient-Reported Outcome Measure "Development of the Renal Stone PROM"

Name	Affiliations	Email	Contribution
Mostafa Ragab*	Cambridge University Hospitals NHS Foundation Trust	m.ragab@nhs.net ORCID iD 0000-0003-2876-3395	Data collection, analysis and manuscript writing
Nikolay Baldin	University of Cambridge, Department of Applied Mathematics and Theoretical Physics	n.baldin@statslab.cam.ac.uk	Nikolay Baldin
Jane Collie	Cambridge University Hospitals NHS Foundation Trust	Jane.collie@addenbrookes.nhs.uk	Project development, Data collection
Maxine Gia Binh Tran	Royal Free London NHS Foundation Trust	m.tran@ucl.ac.uk	Project development, Data analysis and manuscript editing
Sami Al-Hayek	Cambridge University Hospitals NHS Foundation Trust	Sami.hayek@addenbrookes.nhs.uk	Data collection
Kasra Saeb Parsy	Cambridge University Hospitals NHS Foundation Trust	Kasra.saeb-parsy@addenbrookes.nhs.uk	Data Collection
James Armitage	Cambridge University Hospitals NHS Foundation Trust	james.armitage@addenbrookes.nhs.uk	Data collection
Oliver Wiseman	Cambridge University Hospitals NHS	Oliver.wiseman@addenbrookes.nhs.uk	Project development, data analysis and

| | Foundation Trust | | manuscript editing |

Corresponding Author: Mostafa Ragab, Address; Flat 2, 14-16 Harcourt Terrace, Salisbury, SP2 7SA. *Phone*; +44(0) 713959935. *Email*; m.ragab@nhs.net

Abstract (338 words)

Objectives To investigate the experience of patients living with renal calculi via a qualitative methodology, aiming to develop and validate a disease-specific Patient-reported outcome measure (PROM) for renal stones, the Cambridge Renal Stone PROM (CReSP).

Subjects and Methods Patients with radiologically proven renal calculi who had undergone a range of management options were invited to focus groups or semi-structured interviews to elicit patient input and generate the PROM content. The developed renal stone PROM undergone validity studies included Cronbach's alpha for internal consistency, Spearman's and Pearson's correlation coefficients for test–retest reliability. Discriminant validity was assessed by Pearson's correlation coefficients versus EQ5D5L. Our project has Health and Social Care Research Ethics Committee approval.

Results A total of 106 subjects participated in creating the newly developed PROM. 36 patients were invited to 22 semi-structured interviews and 4 focus groups, until reaching saturation. Major issues reported, and themes selected for the renal stone PROM included pain, anxiety, limitations to social life and tiredness, urinary symptoms, dietary changes' impacts and gastrointestinal symptoms. Reliability analysis for 30 patients to determine internal consistency using Cronbach's alpha with a mean of 0.91 (range 0.90 to 0.93) within domains and Cronbach's alpha between domains was 0.92. Average inter-item Pearson's and Spearman's correlation within domains was performed, with Pearson's correlation mean of 0.77 (range 0.73 to 0.85) and Spearman's correlation mean of 0.72 (range 0.63 to 0.77). Test-retest Pearson's correlation mean was 0.85 (range 0.57 to 0.95). Validity assessment was performed for 20 patients versus 20 controls. Pearson's correlation with EQ5D5L was -0.74, showing the newly developed PROM successfully discriminated patients suffering from kidney stones. Our final renal stone PROM consists of 14 questions which are rated on a Likert scale. The higher the score, the worse the effect on a patient's quality of life.

Conclusions Although pain was the most frequent symptom, other health-related and social well-being issues significantly impacted patients' lives. Our validated patient-derived CReSP is a new instrument, specifically tailored to measure renal stone disease health outcomes from the patient's point of view.

Key Words: *Renal Calculi, Patient-reported Outcome measure (PROM), Quality of life.*

1. Introduction (281 words)

Nephrolithasis is a global disease affecting both men and women with increasing incidence over the past few decades [1]. In addition, there is a high recurrence rate reaching 50% within 5 years and 75% by 20 years demonstrating the chronicity of the condition [2].

Surgical outcomes for stone disease have traditionally been based on surgeon reported outcomes such as length of hospital stay, transfusion rates and 'stone free' rates. A paradigm shift in assessing health care from solely clinician reported outcomes to adding patient reported outcomes has emerged, increasing interest in patients' perspectives of the burden of their condition [3]. Since 2009, patient-reported outcome measures (PROMs) are collected for 4 elective procedures in the National Health Service England (hernia, varicose veins repair, hip and knee replacement) with the Department of Health working to extend the PROMs program to include a wider range of conditions and treatments [4]. PROMs can be generic, assessing overall quality of life (QoL), or disease-specific. Over the past decade the burden of bearing urinary calculi on the QoL has been investigated using generic tools such as the 36-item Short Form Health Survey (SF-36) [5-7]. These studies have shown decreased QoL for patients with urinary stones. The urolithasis-specific Health-related QoL (HRQoL) measure, Wisconsin Stone Quality of Life (WISQoL) Questionnaire was developed in North America, assessed the decrease of the HRQoL in patients with urinary stones along the course of the disease

[8]. To our knowledge, there is no readily available, validated tool to assess disease-specific PROMs of renal stones alone.

In this project, we explore the experiences of patients with renal stones aiming to define a focused, concise and validated renal stone-specific PROM tool via robust qualitative methods.

2. Patients and Methods (654 words)

Qualitative methodology was used to explore the burden of renal stones on patients. Item generation was performed by identifying patients with documented radiologically proven renal stones presenting to the stone clinic from February to June 2016, by an investigator and a gatekeeper (specialist nurse) separately and invite them to participate in semi-structured interviews and focus groups. Interviews were conducted first to explore the subject as a starting point of the enquiry. Interviews were one-on-one sessions between a trained investigator (a urologist not involved in the treatment process) and the participant. The discussions were conducted with broad, open-ended styled questions, allowing space for patients to express their experiences and probing the emergent information further. Themes generated were the pillars of our focus groups' discussions. Focus groups had a lead, assisted by a facilitator and a note keeper recording the meeting.

The transcript was analysed using thematic analysis in an inductive approach via the facilitator and the note keeper separately after each focus group. Further interviews and focus groups were conducted until no further concepts were identified reaching the saturation point guided by the saturation grid [9]. At this stage, respondent validation was performed by inviting the focus groups participants to revise and clarify earlier statements to confirm the credibility of the content. Performed via synthesized member checking, sending each participant a transcript report of their quotes with a free reply envelope allowing space for adjustments. This was preceded by checking with the gatekeeper on patients' status and suitability to receive their transcript reports [10]. Concepts generated were coded by medically and urologically defined symptoms (eg,

"haematuria", "renal pain", "nausea") and the relationships between them were examined, aiming to allocate similar codes under overarching potential themes. Additionally, two separate assessors evaluated the generated codes and whether they reflected the concepts within the original data set. Finally, themes were defined according to the general scope of each theme [11].

Frequency of prevalence of each theme in the focus groups was calculated and allocated a weight in the renal stone-specific PROM. Psychometrically sound items adhering to participants' words and reflecting the content of the qualitative statements were selected from existing measurement systems as the Patient-Reported Outcome Measurement Information System (PROMIS) and the Functional Assessment of Chronic Illness Therapy (FACIT) [12-13]. Number of items selected for each domain was determined by the weight allocated.

The primary PROM draft was validated using test-retest reliability and assessing agreement using the Bland-Altman plot. Internal consistency within domains was evaluated via Cronbach's alpha with a predefined statistical threshold of ≥0.70 [14] and average inter-item Pearson's and Spearman's correlation coefficients. The produced PROM was further validated by assessing the ability to discriminate between responses of a group of patients versus a control group. In addition, discriminant validity was performed between the responses of the patient group for the renal stone PROM versus the EQ5D5L using Pearson's Correlation coefficients. The renal stone PROM was graphically enhanced via a biomedical graphic designer, focusing on the colour, brightness, spatial arrangement, and consistency to enhance the visual presentation of questionnaire survey [15]. Ethical approval was granted as a non-substantial amendment of an existing project: Reference 14/NI/1111.

2.1. Sample inclusion and exclusion criteria

Purposeful sampling recruiting patients covering broad spectrum of the disease ensured capturing all relevant perspectives of renal stones burden. The sample included patients with different renal stone sizes, locations and undergoing variable treatment modalities, such as;

Percutaneous nephrolithotomy, flexible ureterorenoscopy, shock wave lithotripsy & active surveillance. We included patients at different phases of treatment, with a proportion stone free, followed up post-operatively and a proportion with a recent stone event, thereby gathering the whole experience. Patients were English speaking, of 18-80 years of age, and able to give informed consent to participate in the study. Exclusion criteria was concurrent urological pathology, including episodes of documented radiologically proven ureteral stones, pre-existing anticholinergic, alpha blocker, calcium channel antagonists, or phosphodiesterase 5 inhibitor therapy and pre-existing chronic pain syndrome.

3.   Results (909 words)

A total of 106 participants contributed in forming the renal stone PROM. Starting by the qualitative phase, 36 patients were invited to 22 semi-structured interviews and 4 focus groups (14 patients) in Cambridge, UK. Table 1 summarises the demographic and clinical characteristics of the sample. Interviews were conducted first reaching saturation after 14 interviews (out of 22). Concepts generated were used as a basis for the discussion in the focus groups. 14 patients attended four focus groups, reaching saturation guided by a saturation grid (Table 2). Frequency of the items generated was calculated reflecting the weight of domains in the primary draft of the PROM tool. Six of the 14 participants (n= 6, 42.85%), responded to the transcript reports sent and confirmed their statements with no amendments and no additional comments.

3.1.   Generated themes

18 themes were generated and mentioned 83 times, with variable frequencies. Thematic maps were particularly useful at this stage as the codes were reviewed globally and adjusted according to the suggested themes (Figure S1, Figure S2).

Pain was the commonest symptom to be mentioned among all the raised symptoms (n=25, 30%), with variable degrees of severity elicited. Some described it as minimal intermittent pain,

while others mentioned it as continuous and limiting, in addition to extremely severe pain being reported as well. On the other hand, some patients with non-obstructing renal stones denied any pain. Finally, various degrees of limitations because of the pain were observed.

Limitations to work, daily activities and travel plans were noticed in a considerable percentage of patients (n=16, 19.3%). The disease affected the patients socially as well, including family relations. Anxiety was expressed (n=14, 16.9%), where several patients felt fearful and worried during various stages of the disease and its management in some cases due to specific symptoms such as pain and haematuria. Variable complications from multiple interventions such as pain, infection and bleeding were mentioned, causing fear and anxiety from pursuing treatment. Also, many patients described their negative experience of having a ureteric stent. Urinary symptoms, were next to be discussed by patients (n= 13, 15.7%) and passing blood in urine was the commonest urinary symptom, in addition to dysuria and frequency. Nausea was also addressed (n=2, 2.4%).

Another domain that impacted the lives of our participants was the dietary changes recommendations to reduce risk of acquiring further urolithasis (n=11, 13.3%). Alterations to the patients' diet as advised by information leaflet for stone formers given to patients upon diagnosis, such as; feeling a need to be hydrated constantly and avoiding certain foods impacted the daily routine of our participants.

3.2. *Item selection of the primary draft of the Cambridge Renal Stone PROM (CReSP) and assessing internal consistency, reliability and validity*

3.2.1. *Prevalence of themes and item selection*

Frequency of occurrence of each theme in the focus groups was calculated (Fig. 1) and given a weight in the renal stone-specific PROM, with a number of items allocated per theme accordingly (Table 3). Psychometrically validated items were selected from the PROMIS and FACIT measurement

systems databases where available. Major issues reported and number of items selected for the PROM included pain (6), anxiety (4), limitations to social life and tiredness (4), urinary symptoms (3), dietary changes impacts (3) and Gastrointestinal symptoms (1).

3.2.2. Content Validity and acceptability

Patients participating in the reliability pilot study answered all questions of the primary draft of CRESP with no nonresponses recorded. Also, the PROM was presented in multiple urological scientific meetings and to the British Association of Urological Surgeons Endourology Committee with general acceptability of the tool.

3.2.3. Test-retest reliability

Test-retest data was collected for 30 patients during their clinic appointment. Test-retest time interval was 2-3 hours. All questions yielded very high test-retest correlation (Mean 0.85, range 0.57-0.95) apart from 2 questions in the Anxiety domain (0.66 & 0.57) (Fig. 2). To further illustrate test-retest reliability, the Bland-Altman plots for separate domains are depicted (Fig. 3).

3.2.4. Internal Consistency

Internal consistency was calculated within each domain via Cronbach's alpha with a predefined statistical threshold of ≥0.70 and average inter-item correlation (Pearson's and Spearman's) as well. After exclusion of redundant questions, the Cronbach's alpha had a mean of 0.91 (range 0.90 to 0.93) and Cronbach's alpha between domains was 0.92. Inter-item correlation within domains was positive, with Pearson's correlation mean of 0.77 (range 0.73 to 0.85) and Spearman's correlation mean of 0.72 (range 0.63 to 0.77) (Table 4).

3.2.5 Discriminant validity

Discriminant validity was assessed by comparing the averages of the scores of 2 groups of 40 patients and controls for each domain of the renal stone PROM, characteristics of both groups were analysed (Table 5). The box plots show the scores of participants per study allocation group and

domain (Fig. 4). For all domains, the spread of the scores of the patient group were larger than the one of the control group. Analysis of the patient group, completing the EQ5D5L simultaneously with the renal stone PROM showed a Pearson's correlation of -0.74, showing the newly developed CReSP successfully discriminated patients suffering from kidney stones (Fig. 5).

3.2.6   Design of the produced PROM tool

Eight designs were proposed and members of the research team were asked to select their first two preferences, reaching the current design by >80% majority. The developed instrument fits a single paper with acceptable font size (10-11) and appropriate spatial arrangement. CReSP consists of six domains containing 14 items with a score range of 14-75 (Fig 6a, 6b), ready for clinical use.

4.   Discussion (772 words)

The inclusion of PROMs instruments in urological practice has been increasing over the past decade, aiding clinicians in viewing the experience lived by the patients [16]. PROMs allow comparison between different therapies, hospitals and clinicians, benchmark the performance of health care providers, and enable health care professionals to monitor feedback on practices provided [4]. Previous studies assessed the HRQoL of renal stones patients by generic tools or non-stone specific tools, showing decrease in HRQoL [5-7] and inducing stress and depression [17-18]. We aimed to develop a PROMs tool to assess HRQoL of kidney stones solely. WISQOL is a urolithasis specific PROM tool, developed in North America which demonstrated decreased HRQoL along the disease continuum. WISQOL is currently in the testing phase to confirm its broad applicability and clinical significance [19]. While WISQOL examines the whole experience of acquiring urolithasis, this study explored experiences of patients specifically with renal stones and defined a renal stone-specific PROM tool. Despite being a disease continuum, ureteric and renal stones are separate entities with different presentations and treatment modalities. The minority of patients with non-obstructing renal calculi experience symptoms [20-21] which is embraced by the European Association of Urology grade C recommendation for active surveillance for asymptomatic, non-

obstructing renal stone [22]. This is different for ureteric stones where majority of patients are symptomatic [23], highlighting the need to assess the experience of the patients with renal stone separately from ureteric stones. In conjunction with developing a ureteric stone disease-specific PROM tool (Cambridge Ureteric Stone PROM) [24], we allow for a sensitive, tailored approach to assess the experience of two clinically different phases of the disease. Studying the correlation between CReSP and WISQOL will guide on complimentary benefits of using both tools.

In developing this new instrument, we followed multiple strategies to ensure the trustworthiness of our qualitative methodology. Starting by utilising different methods, such as; semi-structured interviews and focus groups for data collection. Information were produced by asking questions in a broad, open-ended manner and incorporating the emergent information to be explored in subsequent interviews and focus groups [25]. Interviews provided a private atmosphere enabling the opportunity for in-depth discussion of sensitive issues that may not have been forthcoming within a focus group discussion [9]. Each focus group had a leader, accompanied by a facilitator allowing time for all participants to be involved and ensuring the discussion is not dominated by few members and a note keeper to record the meeting [26]. Using these two complementary techniques ensured triangulation. Supplemented by respondent validation, debriefing sessions between involved subspecialist urinary stone clinicians and peer scrutiny in specialised conferences, which contributed to the overall credibility of our methodology [27].

Developing CReSP, we selected items from the PROMIS and FACIT measurement systems databases, which provided items with good content validity and coverage of a wide variety of symptoms and social well-being which was a frequent theme in our study. Selected items reflected the words and meanings participants described, in addition to easily understood Likert Scale responses [28-29]. A limitation of this approach was the need to measure the internal consistency of the selected items within each domain and between domains. Besides, the impact of dietary changes theme was not represented in the aforementioned databases, thus assessment of the

reliability and validity of the impact of dietary changes domain's items together with the whole tool was needed.

Reliability and internal consistency values were acceptable per predefined thresholds (≥0.7), after excluding redundant questions for the sample of patients. The short test-retest interval ensured clinical stability of the disease, which is essential in the early phase of developing the PROM tool and decreasing the chance of a significant change in symptoms that could change the response to the retest questionnaire [30]. Simultaneously, conducting the clinic consultation between both times creates an interruption preventing familiarisation with the tool. All questions yielded very high test-retest correlation, apart from 2 questions in the anxiety domain. The majority of the researchers decided to include the items, as patients scored lower for these questions in the retest questionnaire after eluding their anxiety as a result of attending their consultation. Discriminant validity showed the produced PROM's ability to detect deterioration in QoL against the EQ5D5L and differentiate between the patient and control groups. While results appear promising, a longer test-retest duration needs addressing in the next phase of confirming the reliability for a larger, more diverse patient sample with representatives from other ethnicities and different marital status. The effects of which will be measured in the upcoming validation studies as part of the UK national PUrE study, in addition to translating CReSP to different languages to assess international applicability.

5.   Conclusion (86 words)

This study explored patients' experiences of living with a renal stone and provides a detailed approach for obtaining a trustworthy qualitative approach for developing a PRO instrument. The developed disease-specific renal stone PROM has the potential of assessing the experience of renal stone patients, after establishing its clinical applicability. It will allow comparison between the various possible treatment options for kidney stones from a patient's viewpoint, and will also allow comparison of outcomes between individual clinicians and hospitals for a particular stone size, site or procedure.

Word Count 2702

Abbreviations

PROMs = Patient Reported Outcome Measures

CReSP = Cambridge Renal Stone PROM

HrQoL = Health related Quality of Life

SF-36 = 36-item Short Form Health Survey

WISQoL = Wisconsin Stone Quality of Life

PROMIS = Patient-Reported Outcome Measurement Information System

FACIT = Functional Assessment of Chronic Illness Therapy

Compliance with Ethical Standards

Conflict of Interest

Oliver Wiseman is a consultant to Boston Scientific and Porges Coloplast; Education for Boston

Scientific, Porges Coloplast, Olympus Corp, EMS. Research study for Porges Coloplast. All other

authors declare that they have no conflicting interests.

Ethical Approval

All procedures performed in studies involving human participants were in accordance with the

ethical standards of the institutional and/or national research committee and with the 1964 Helsinki

declaration and its later amendments or comparable ethical standards.

Informed consent

Informed consent was obtained from all individual participants included in the study.

References:

1. Sorokin I, Mamoulakis C, Miyazawa K, Rodgers A, Talati J, Lotan Y. Epidemiology of stone disease across the world. *World Journal of Urology* 2017;35:1301-1320

2. Moe O. Kidney stones pathophysiology and medical management. *The Lancet* 2006;367:333-344

3. Bren L. The importance of patient-reported outcomes…it's all about the patients. *FDA Consumer Magazine* 2006;40:26-32

4. Devlin N, Appleby J. *Getting the most out of PROMs*. London, King's Fund, 2010.

5. Penniston K, Nakada S. Health Related Quality of Life Differs Between Male and Female Stone Formers. *The Journal of Urology* 2007;178:2435-2440

6. Bensalah K, Tuncel A, Gupta A, Raman J, Pearle M, Lotan Y. Determinants of Quality of Life in Patients with Kidney Stones. *The Journal of Urology* 2008;179:2238-2243

7. Bryant M, Angell J, Tu H, Goodman M, Pattaras J, Ogan K. Health Related Quality of Life for Stone Formers. *The Journal of Urology* 2012;188:436-440

8. Penniston K, Nakada S. Development of an Instrument to Assess the Health Related Quality of Life of Kidney Stone Formers. *The Journal of Urology* 2013;189:921-930

9. Brod M, Tesler L, Christensen T. Qualitative research and content validity: developing best practices based on science and experience. *Quality of Life Research* 2009;18:1263-1278

10. Birt L, Scott S, Cavers D, Campbell C, Walter F. Member Checking: A Tool to Enhance Trustworthiness or Merely a Nod to Validation? *Qualitative Health Research* 2016;26:1802-1811

11. Braun V, Clarke V. Using thematic analysis in psychology. *Qualitative Research in Psychology* 2006;3:77-101

12. Jackson M, Sciberras J, Mangera A, et al. Defining a Patient-Reported Outcome Measure for Urethral Stricture Surgery. *European Urology* 2011;60: 60-68

13. Schroter S, Lamping DL. Coronary revascularisation outcome questionnaire (CROQ): development and validation of a new, patient based measure of outcome in coronary bypass surgery and angioplasty. *Heart* 2004;90: 1460-1466

14. Smith S, Cano S, Lamping D, et al. Patient-Reported Outcome Measures (PROMs) for routine use in Treatment Centres: recommendations based on a review of the scientific evidence. London, London School of Hygiene and Tropical Medicine, 2005.

15. McColl E, Jacoby A, Thomas L, et al. Design and use of questionnaires: a review of best practice applicable to surveys of health service staff and patients. *Health Technology Assessment* 2001;5:1-256

16. Narang G, Pannell S, Laviana A, et al. Patient-reported outcome measures in urology. *Current Opinion in Urology* 2017;27:366-374

17. Miyaoka R, Ortiz-Alvarado O, Kriedberg C, Alanee S, Chotikawanich E, Monga M. Correlation Between Stress and Kidney Stone Disease. *Journal of Endourology* 2012;26:551-555

18. Angell J, Bryant M, Tu H, Goodman M, Pattaras J, Ogan K. Association of Depression and Urolithiasis. *Urology* 2012;79:518-525

19. Penniston K, Antonelli J, Viprakasit D, et al. Validation and Reliability of the Wisconsin Stone Quality of Life Questionnaire. *The Journal of Urology* 2017;197:1280-1288

20. Keeley F, Tilling K, Elves A, et al. Preliminary results of a randomized controlled trial of prophylactic shock wave lithotripsy for small asymptomatic renal calyceal stones. *BJU International* 2001;87:1-8

21. Dropkin B, Moses R, Sharma D, Pais, Jr. V. The Natural History of Nonobstructing Asymptomatic Renal Stones Managed with Active Surveillance. *The Journal of Urology* 2015;193:1265-1269

22. Türk C, Petřík A, Sarica K, et al. EAU Guidelines on Diagnosis and Conservative Management of Urolithiasis. *European Urology* 2016;69:468-474

23. Wimpissinger F, Türk C, Kheyfets O, Stackl W. The Silence of the Stones: Asymptomatic Ureteral Calculi. *The Journal of Urology* 2007;178:1341-1344

24. Tran M, Sut M, Collie J, et al. Development of a Disease-Specific Ureteral Calculus Patient Reported Outcome Measurement Instrument. *Journal of Endourology* 2018;32(6):548-558

25. Lingard L, Albert M, Levinson W. Grounded theory, mixed methods, and action research. *BMJ* 2008;337:a567

26. Onwuegbuzie A, Leech N, Collins K. The Qualitative report: Innovative Data Collection Strategies in Qualitative Research: 2010. http://nsuworks.nova.edu/tqr/vol15/iss3/12. (Visited January 25, 2017).

27. Shenton A. Strategies for ensuring trustworthiness in qualitative research projects. *Education for Information* 2004;22:63-75

28. Riley W, Rothrock N, Bruce B, et al. Patient-reported outcomes measurement information system (PROMIS) domain names and definitions revisions: further evaluation of content validity in IRT-derived item banks. *Quality of Life Research* 2010;19:1311-1321

29. Webster K, Cella D, Yost K. The Functional Assessment of Chronic Illness Therapy (FACIT) Measurement System: properties, applications, and interpretation. *Health and Quality of Life Outcomes* 2003;1:79

30. Carvajal A, Centeno C, Watson R, Bruera E. A comprehensive study of psychometric properties of the Edmonton Symptom Assessment System (ESAS) in Spanish advanced cancer patients. *European Journal of Cancer* 2011;47:1863-1872