BJU Int 2022 doi:10.1111/bju.15916

BJUI
BJU International

Original Article

The Uniform grading tooL for flexIble ureterorenoscoPes (TULIP-tool): a Delphi consensus project on standardised evaluation of flexible ureterorenoscopes

Michaël M.E.L. Henderickx[1], Nora Hendriks[1,2], Joyce Baard[1], Oliver J. Wiseman[3], Kymora B. Scotland[4], Bhaskar K. Somani[5], Tarik E. Şener[6], Esteban Emiliani[7], Laurian B. Dragos[3], Luca Villa[8], Michele Talso[9], Saeed Bin Hamri[10], Silvia Proietti[11], Steeve Doizi[12,13], Olivier Traxer[12,13], Ben H. Chew[14], Brian H. Eisner[15], Manoj Monga[16], Ryan S. Hsi[17], Karen L. Stern[18], David A. Leavitt[19], Marcelino Rivera[20], Daniel A. Wollin[21], Michael Borofsky[22], Noah E. Canvasser[23], Johann P. Ingimarsson[24], Marawan M. El Tayeb[25], Naeem Bhojani[26], Nariman Gadzhiev[27], Thomas Tailly[28], Otas Durutovic[29], Udo Nagele[30], Andreas Skolarikos[31], Barbara M.A. Schout[2], Harrie P. Beerlage[1], Rob C.M. Pelger[32] and Guido M. Kamphuis[1]

[1]Department of Urology, Amsterdam UMC Location University of Amsterdam, Amsterdam, [2]Department of Urology, Alrijne Hospital, Leiden, The Netherlands, [3]Department of Urology, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK, [4]Department of Urology, University of California, Los Angeles, CA, USA, [5]Department of Urology, University Hospitals Southampton NHS Trust, Southampton, UK, [6]Department of Urology, Marmara University School of Medicine, Istanbul, Turkey, [7]Department of Urology, Fundacio Puigvert, Barcelona, Spain, [8]Department of Urology, Università Vita – Salute San Raffaele, [9]Department of Urology, ASST Fatebenefratelli Sacco – Ospedale Luigi Sacco University Hospital, Milan, Italy, [10]Department of Urology, King Abdullah International Medical Research Center, College of Medicine, King Abdulaziz Medical City, King Saud Bin Abdulaziz University for Health Science, Riyadh, Saudi Arabia, [11]Department of Urology, IRCCS San Raffaele Hospital, Milan, Italy, [12]Department of Urology, Hopital Tenon, [13]Sorbonne Université, 27063, GRC n°20, Groupe de Recherche Clinique sur la Lithiase Urinaire, Paris, France, [14]Department of Urology, University of British Columbia, Vancouver, BC, Canada, [15]Department of Urology, Harvard Medical School, Massachusetts General Hospital, Boston, MA, [16]Department of Urology, University of California San Diego, San Diego, CA, [17]Department of Urology, Vanderbilt University Medical Center, Nashville, TN, [18]Department of Urology, Mayo Clinic, Phoenix, AZ, [19]Department of Urology, Vattikuti Urology Institute, Henry Ford Hospital, Detroit, MI, [20]Department of Urology, Indiana University School of Medicine, Indianapolis, IN, [21]Department of Urology, Harvard Medical School, Brigham and Women's Hospital, Boston, MA, [22]Department of Urology, University of Minnesota, Minneapolis, MN, [23]Department of Urology, University of California Davis Health System, Sacramento, CA, [24]Department of Urology, Maine Medical Center, South Portland, ME, [25]Department of Urology, Baylor Scott & White Medical Center – Temple, Temple, TX, USA, [26]Department of Urology, Centre Hospitalier de l'Université de Montréal, Université de Montréal, Montreal, QC, Canada, [27]Department of Urology, Endourology Academy, Istanbul, Turkey, [28]Department of Urology, University Hospital Ghent, Ghent, Belgium, [29]Department of Urology, University Clinical Center of Serbia, Belgrade, Serbia, [30]Department of Urology, General Hospital Hall I.T., Hall in Tirol, Austria, [31]Department of Urology, University of Athens, Athens, Greece, and [32]Department of Urology, Leids UMC, University of Leiden, Leiden, The Netherlands

M.M.E.L.H. and N.H. contributed equally to this paper and share first authorship.

Objective

To develop a standardised tool to evaluate flexible ureterorenoscopes (fURS).

Materials and Methods

A three-stage consensus building approach based on the modified Delphi technique was performed under guidance of a steering group. First, scope- and user-related parameters used to evaluate fURS were identified through a systematic scoping review. Then, the main categories and subcategories were defined, and the expert panel was selected. Finally, a two-step modified Delphi consensus project was conducted to firstly obtain consensus on the relevance and exact definition of each (sub)category necessary to evaluate fURS, and secondly on the evaluation method (setting, used tools and unit of outcome) of those (sub)categories. Consensus was reached at a predefined threshold of 80% high agreement.

Results

The panel consisted of 30 experts in the field of endourology. The first step of the modified Delphi consensus project consisted of two questionnaires with a response rate of 97% (n = 29) for both. Consensus was reached for the relevance and definition of six main categories and 12 subcategories. The second step consisted of three questionnaires (response rate of 90%, 97% and 100%, respectively). Consensus was reached on the method of measurement for all (sub)categories.

Conclusion

This modified Delphi consensus project reached consensus on a standardised grading tool for the evaluation of fURS – The Uniform grading tooL for flexIble ureterorenoscoPes (TULIP) tool. This is a first step in creating uniformity in this field of research to facilitate future comparison of outcomes of the functionality and handling of fURS.

Keywords

flexible ureterorenoscope, quality assessment, evaluation, tool, endourology, Delphi consensus, #EndoUrology, #UroStone, #Urology

Introduction

The first flexible ureterorenoscope (fURS) originally was derived from the gastroscope. During the 1950s, gastroenterologist Hirsowitch and a physics student joined forces to make what we would later call the first gastroscope. He tested it on himself and used it to treat a stomach ulcer in a patient the next day. This incited medical specialists in other fields, and so the first 9-F fURS was used by Marshall in the 1960s. Because of limitations in size and flexibility, as well as the fear of perforating the upper tract, it was only used as a diagnostic tool. The first purposeful use was in 1977 when a biopsy of a ureteric tumour was taken with concurrent coagulation. During the 1980s, a rapid development took place as urologists began to treat stones with ultrasonic lithotripters. As great safety concerns about the size of scopes and instruments prevailed, the movement to miniaturise the fURS commenced [1]. This drive to miniaturise has occurred alongside the desire to improve characteristics such as image quality, flexibility, irrigation, ease of use of different instruments and durability.

Securing high-quality standards while miniaturising the device, as well as introducing technological innovations, poses obvious challenges. Multiple manufacturers developed URS, all with specific trademarks and characteristics (such as flexible image transmission, active tip deflection or integration of miniaturised digital image sensors), leading to great variation in the quality of fURS [2–6]. Single-use fURS for example, are described to have inferior manoeuvrability when compared to reusable fURS. And digital fURS tend to have a larger calibre and less end-tip deflection than fibre-optic fURS [2,4]. Furthermore, manufacturers tend to advertise the outstanding quality of one specific characteristic such as bidirectional flexion, digital optics, dual working channels, or miniature size [2–6]. This may imply that the focus on one characteristic inhibits above average performance of other characteristics. This subsequently leads to a noticeable variance in the performance and quality of commercially available URS, especially as the number of URS has rapidly increased in recent years.

High-quality fURS are important to assure patient safety. A safe and efficient procedure depends not only on the urologist's skill and experience, but also on the tools the urologist uses. Clear vision may increase safety and is intimately related to irrigation flow and the size of the working channels. Furthermore, improved deflection may make access to steep calyces possible and thus increase the stone-free rate. More recent developments have included the integration of pressure control at the tip of the fURS, which could lead to a decreased risk of pyelovenous backflow, calyceal rupture or tubular and interstitial backflow as urologists can monitor the intrarenal pressure more accurately [3].

As urologists experience this variability of quality in daily practice, there has been a large increase in the number of published papers calling for a means of grading the quality of fURS [7]. One of the first articles evaluating a fURS was published in 1997 followed by a steep increase in the number of papers on this subject. More than 70% of all articles describing the evaluation of fURS have been published after 2010. Not only did the number of published papers increase significantly, but also the number of ureterorenoscopic procedures, scopes and manufacturers grew exponentially. England for example, saw an increase of 257% in ureterorenoscopies between 2000 and 2020. In the last 5 years, the number of ureterorenoscopies increased further by 18.9%, of which flexible ureterorenoscopies made up 20.4% or a growth from 7108 procedures per year in 2015 to 8558 procedures per year in 2020 [8].

While the interest and necessity to grade and compare scopes is portrayed by these numbers, our recently published systematic scoping review showed great heterogeneity in measurement methods, construct, definitions, and unit of outcome [7]. This makes it nearly impossible to compare quality of different fURS in different papers. However, after a process of regrouping, key parameters with a variety of subcategories could be distinguished [7]. Now that all important parameters and their methods of measurement have been mapped, categorised, and subcategorised, we feel that it is time to create a uniform grading tool. Therefore, we aimed to perform a modified Delphi consensus project in collaboration with the colleagues from the Collaborative for Research in Endourology (CoRE), Endourology Disease Group For Excellence (EDGE), European Association of Urology Section of Uro-Technology (ESUT) and Progress in Endourology, Technology and Research Association (PETRA) working groups to reach consensus on a standardised grading tool to evaluate the quality of fURS.

Materials and Methods

A three-stage consensus building approach based on the modified Delphi technique was performed under guidance of the study initiators (M.M.E.L.H.; N.H.; J.B. and G.M.K.) [9,10].

Stage 1: identifying scope-related and user-related parameters used to evaluate fURS through a systematic scoping review (August 2020–July 2021)

In order to assess all relevant literature on this topic, a systematic scoping review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) covering four different databases [11]. Specifics about the evidence acquisition are described in length in our previous paper [7].

Stage 2: defining main evaluation categories and subcategories and selecting experts (July 2021)

The systematic scoping review resulted in the identification of five main categories with several subcategories. An additional sixth main category was identified, and some subcategories were redefined after steering group meetings in order to create a practical tool.

Participants for the third stage were selected based on their expertise in the field of endourology, their contribution to research on the evaluation of fURS and their affiliation with European and North American endourology working groups. A total of 31 experts from the PETRA, CoRE, EDGE and ESUT endourology working groups were invited to participate in this Delphi consensus project.

Stage 3: conducting a modified Delphi consensus project (August 2021–April 2022)

The consensus project was divided into the following steps:

Step 1: obtaining consensus on the relevance and exact definition of each (sub)category necessary to evaluate fURS (Questionnaire1 [Q1] and Questionnaire 2 [Q2]).

Step2: inventorying preference on an extensive range of methods of measurement for all (sub)categories to pre-select items for Step 3 (Questionnaire 3 [Q3]).

Step 3: obtaining consensus on the evaluation method of all (sub)categories (Questionnaire 4 [Q4] and Questionnaire 5 [Q5]).

Consecutive questionnaires were created on an on-line clinical trial platform (https://www.castoredc.com). The questionnaires and proposed definitions were checked for grammatical errors by two native English speakers (O.J.W. and K.B.S.). Prior to sending out the questionnaire to the experts, a pilot survey was sent to independent researchers in the field of urology (PhD candidates: Ben-Max de Ruiter, Luigi van Riel and Hilin Yildirim) in order to check understanding and ensure logical survey structure.

In every questionnaire, the experts were encouraged to provide suggestions and feedback. Results of previous questionnaires were summarised and incorporated in the following questionnaires. If consensus was not reached, the proposals were adjusted considering the suggestions of the experts. Agreement was evaluated on a visual analogue scale ranging from 1 to 9. A score of 1–3 was considered low agreement, 4–6 intermediate agreement and 7–9 high agreement. The cut-off for consensus was set at ≥80% high agreement [12].

Step 1. Consisted of Q1 and Q2 focused on assessing the relevance of different (sub)categories: 'Which (sub)categories are relevant in the assessment of the quality of fURS?'. As well as reaching consensus on the exact definition of each (sub)category. High agreement was reached on definitions and relevance in these two rounds.

Step 2. Due to the great heterogeneity in measurement methods, part of this modified Delphi consensus project consisted of an inventory round (Q3). This was necessary to preselect the right items and create grading tools based on the knowledge and preference of the expert panel. The expert panel was presented the full range of available methods of measurement for all (sub)categories found during the systematic scoping review. Additionally, extra options, found in comparable studies in other fields of research, were added, in order to be as thorough and unprejudiced as possible. The results of Q3 were the base for the proposed tools for the assessment of the different (sub)categories in the subsequent questionnaires.

Step 3. The aforementioned two subsequent questionnaires (Q4 and Q5) focused on establishing consensus for the evaluation method of the different (sub)categories: 'What method of measurement is the most user-friendly, reproducible and feasible in the assessment of the quality of fURS?'.

As consensus was reached after five questionnaires for both the definitions, as well as the methods of measurement, there was no need for an additional consensus meeting.

Results

An overview of the three-stage consensus building approach based on the modified Delphi technique is shown in Fig. 1.

Stage 1. Identifying scope- and user-related parameters used to evaluate fURS through a systematic scoping review

A total of 2386 articles were screened and finally 48 articles were included in the systematic scoping review. This resulted in the identification of five key parameters used to assess the quality of fURS: ‘Manoeuvrability’, ‘Optics’, ‘Irrigation’, ‘Handling’, and ‘Durability’. Nevertheless, within the different studies there was great heterogeneity in terms of measurement methods, construct, definitions, and measurement outcomes [7].

Stage 2. Defining main evaluation categories and subcategories and selecting experts

A total of 30 of the 31 invited experts agreed to participate in this modified Delphi consensus project. In all, 10 members of the PETRA group, nine of the CoRE group, six of the EDGE group and six of the ESUT endourology working group were involved in this project.

Fig. 1 Overview of the progress and results of the three stages.

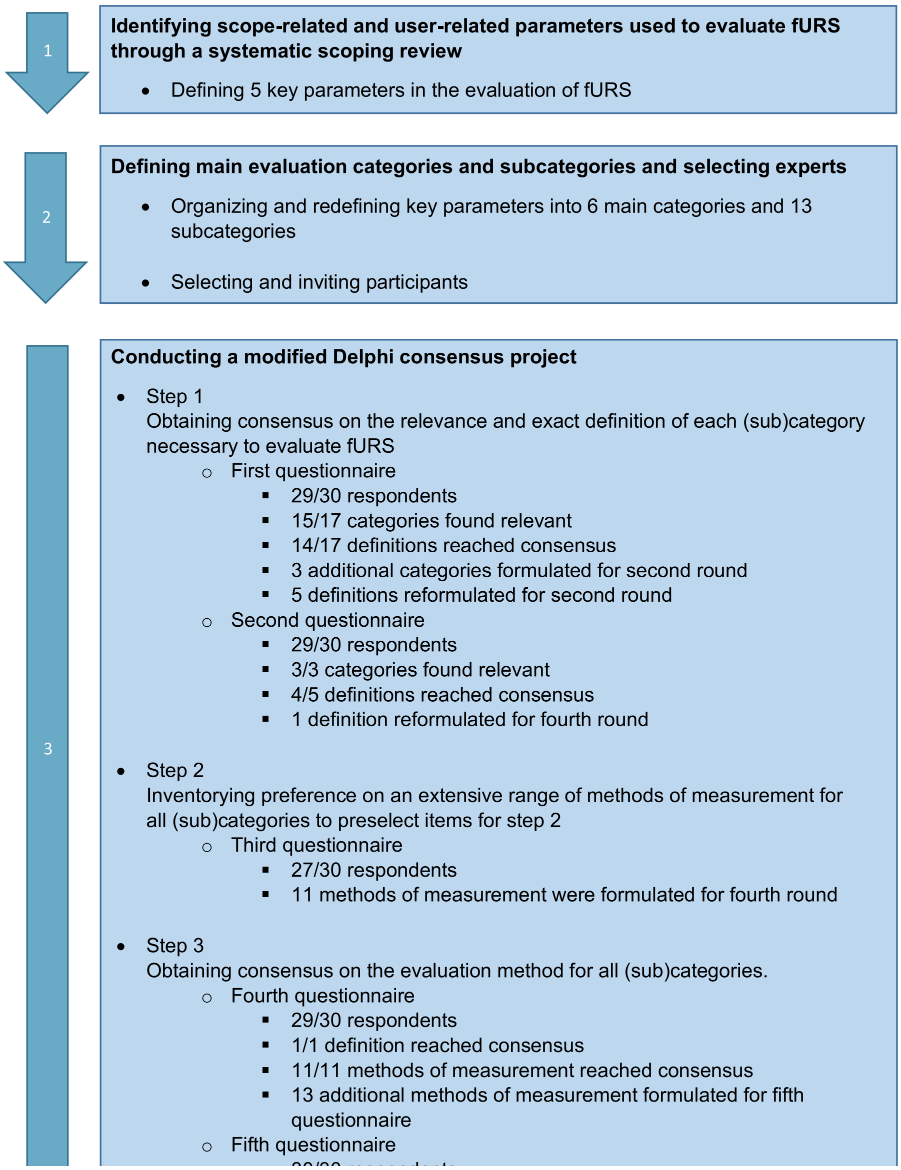


Table 1 Definitions for all included (sub)categories and their respective percentage of high agreement.

Category	Definition	High agreement, %
1. Manoeuvrability	The quality of movement and directionality of a fURS	97
1.1 Flexion	The assessment of maximum upward and downward flexion from the tangent of the shaft to the distal tip of the fURS	90
1.2 Control	The degree to which movement of the handle and the lever was transmitted and precisely reproduced at the distal part of the scope	96
2. Optics	The assessment of visual characteristics of a fURS	90
2.1 Resolution	The potential detail of an image through a fURS	90
2.2 Brightness	The assessment of the quality and/or quantity of light produced by or transmitted through a fURS	86
2.3 Colour and greyscale	The ability and reliability to differentiate between colours and greyscales in an image produced by a fURS	100
2.4 Field of view	The area that is visible through a fURS	90
2.5 Depth of field	The distance between the closest and the farthest image that appear acceptably sharp through a fURS	90
2.6 Visibility	The assessment of the overall quality of the image of a fURS as perceived by the surgeon	86
3. Irrigation	The flow of fluid through the working channel of a fURS	90
4. Ergonomics	The physical comfort and convenience in the use of a fURS	83
4.1 Comfort	The physical comfort and convenience in the use of a fURS	80
4.2 Scope design	The convenience experienced with regards to the positioning of buttons, ease of using the handle and weight of the fURS	86
5. Durability	The ability of fURS to withstand wear, pressure and damage while still performing adequately	83
5.1 Wear	The loss of quality of a specific characteristic (manoeuvrability, optics, ...) of a fURS over a period of time	83
5.2 Damage	The impairment of a characteristic of the fURS to the extent that the scope is not fit for use during surgery anymore	82
6. Satisfaction	The overall satisfaction of use of a fURS as assessed by the surgeon during a procedure	93

Stage 3. Conducting a modified Delphi consensus project

Step 1. Obtaining consensus on the relevance and exact definition of each (sub)category necessary to evaluate fURS.

This first step consisted of two questionnaires with a response rate of 97% (n = 29) for both.

Consensus was reached for the definition of all but one of these (sub)categories after two questionnaires as presented in Table 1. The definition for the subcategory 'Ergonomics – Comfort' required an additional round to reach the predefined threshold for high agreement of 80% (Fig. 1).

Consensus was reached for the inclusion of six main categories and 12 subcategories in the assessment of quality of a fURS, as presented in Fig. 2. Two subcategories, 'Manoeuvrability – Access' (76%) and 'Optics – Distortion' (79%), were excluded as their relevance, according to the expert panel, did not reach the predefined threshold of 80% for high agreement.

Step 2. Inventory round: inventorying preference on an extensive range of methods of measurement for all (sub)categories to preselect items for Step 2.

The inventory questionnaire (Q3) had a response rate of 90% (Fig. 1). Based on the results of this inventory questionnaire, we were able to create proposals for the method of measurement for 11 (sub)categories ('Manoeuvrability – Control: Degree of torque', 'Optics – Field of view', 'Optics – Depth of field', 'Optics – Visibility', 'Ergonomics – Scope design: Position of the lever', 'Ergonomics - Scope design: Position of the buttons', 'Ergonomics – Scope design: Working channel access', 'Ergonomics – Scope design: Grip', 'Durability – Damage: Leakage', 'Durability: Damage: Shaft damage', 'Satisfaction').

Step 3. Obtaining consensus on the evaluation method of all (sub)categories.

This step consisted of two questionnaires (Q4 and Q5), which had a response rate of 97% and 100%, respectively (Fig. 1). In Q4, consensus was reached for the proposed methods of measurement of all 11 (sub)categories. Additional preferences for the remaining 13 (sub)categories were inventoried in Q4. Proposals for a method of measurement for these 13 (sub)categories were created with these data and presented for consensus in Q5. Consensus was reached for all (sub)categories after Q5 (Figs. 1 and 3). The final result, The Uniform grading tooL for flexIble ureterorenoscoPes (TULIP) tool, is provided in Appendix 1.

Discussion

The characteristics of fURS continue to evolve, resulting in a broad range of scopes that excel in different characteristics. However, improving one characteristic can pose challenges

Fig 2 Relevance of the (sub)categories.

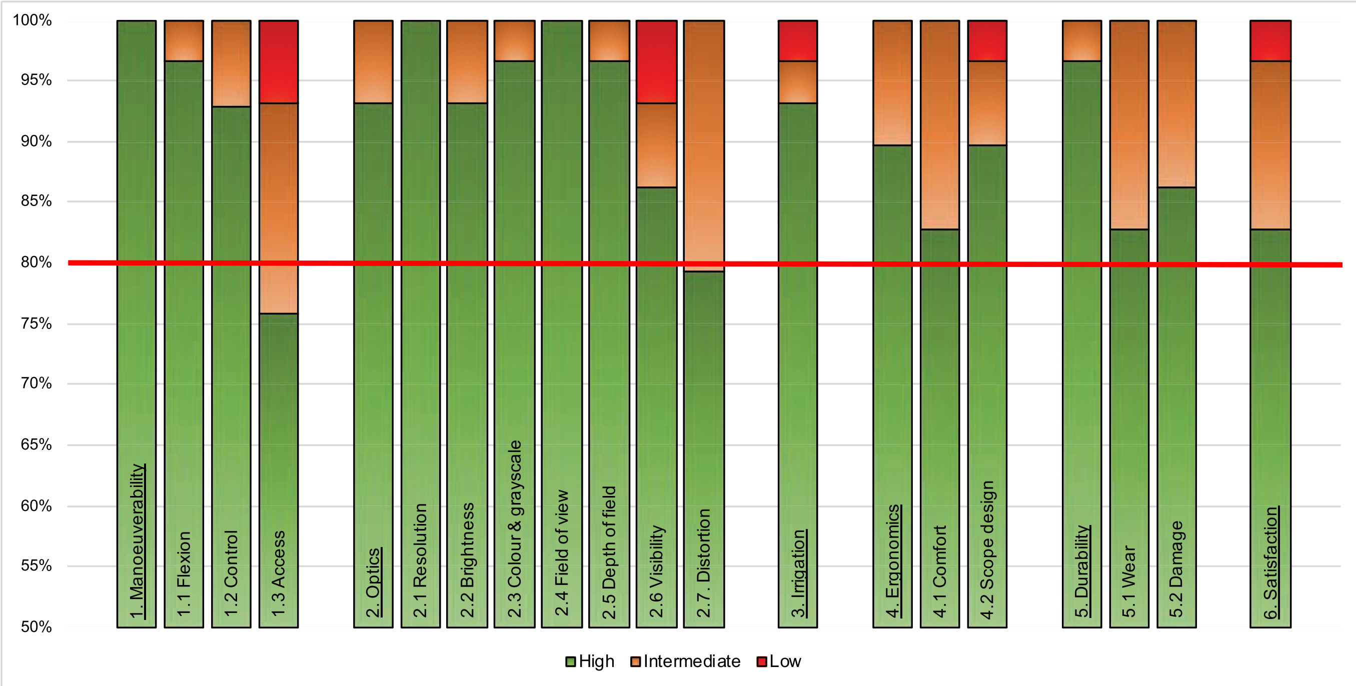
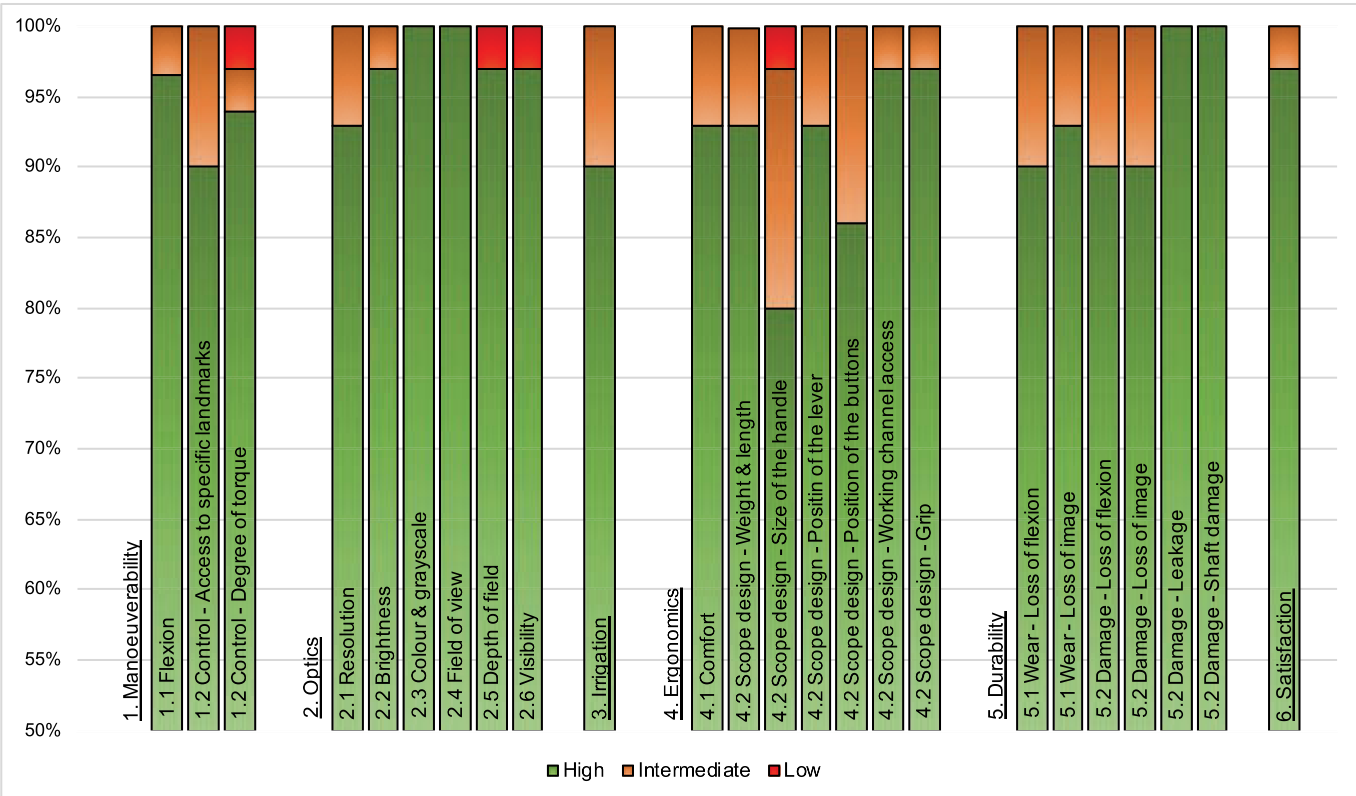


Fig 3 Agreement for the method of measurement of the (sub)categories.



for the quality of other characteristics. This consequently led to the need to evaluate the quality of fURS. Until now, as there were no guidelines or recommendations on how to evaluate fURS, researchers decided themselves what to evaluate and which method they used to assess that particular characteristic. As a consequence, our systematic scoping review showed great heterogeneity in terms of measured subjects, their definitions, and their method of measurement [7]. This results in a range of different outcome variables and complicates comparison of studies, which consequently leads to unstructured fragmented data on the quality of fURS. With this modified Delphi consensus project, we aimed to select and define essential (sub)categories and identify the best possible method of measurement in order to create a uniform grading tool for fURS in association with the CoRE, EDGE, ESUT and PETRA working groups. We established a high response rate (minimum of 90%) and reached consensus (high agreement >80%) on all definitions and methods of measurement on all (sub)categories in five rounds. Consequently, we were able to create the TULIP tool, a user-friendly, efficient, and reproducible grading tool for the evaluation of fURS (Appendix 1).

The TULIP Tool

During the creation process of the TULIP tool, emphasis was laid on the user-friendliness of the tool as well as its reproducibility. A user-friendly tool will increase the chances of it actually being brought into practice later on. One of the consequences of this approach was that the tool had to be created for three different environments: 'in vitro – on bench'; 'in vitro – artificial model of the urinary tract'; and an 'in vivo' environment. This gives researchers the possibility to evaluate fURS in all stages of product development or evaluation (before and after Conformité Européene [CE]/United States Food and Drug Administration [FDA] certification) and in line with the aim of their study (reproducible in vitro measurements vs real life in vivo measurements). As a result of this emphasis, the experts did not always choose for the 'optimal' environment, such as a human or porcine cadaver, but sometimes opted for the more user-friendly environment. Reproducibility on the other hand was guaranteed by offering an in vitro on bench option. This allows researchers from all over the world, independent of their facilities, to evaluate fURS in a standardised manner and to compare their outcomes with other studies from fellow researchers.

The expert panel decided not to define which instruments (i.e., basket, biopsy forceps) should be used during the evaluation, but instead created a setting in which a range of instruments can be used as long as they are described in detail in the 'Materials and Methods' section. This contributes to the longevity of the grading tool, as future instruments can be included in the evaluation as well.

Strengths and Limitations

A strength of this study is the expert panel consisting of 30 experienced endourologists from some of the most eminent endourology working groups in Europe and North America. Furthermore, we reached a minimal response rate of 90% for all five questionnaires. This portrays the necessity of such a tool as well as the commitment of the experts to this project.

Finally, consensus was set at a threshold of minimum 80% high agreement. This is higher than in most modified Delphi consensus projects, where thresholds are usually set at 50–70% [13,14]. As our absolute priority was to make a tool that would actually be used, we deemed it necessary to set a high threshold.

Despite its strengths, there are some limitations. The composition of the expert panel may be subject to discussion. Even though we included experts from Europe and North America, we did not incorporate working groups situated in other continents. As the researched subject is a niche within endourology, our primary focus was on those experts who have already published on this subject. Furthermore, we did not involve experts from other fields, such as physicists or engineers. Nor did this project involve representation from the industry. As we wanted the TULIP tool to be applicable for all fURS and stay away from possible conflicts of interest, we chose not to included representatives from companies producing fURS. However, a number of co-authors are or have been consultants for companies producing fURS. And in that role, they have aided in the design and development of different scopes. Additionally, the 'in vitro – model of the urinary tract' was not defined as there is a wide range of endourological simulators and every model has its pros and cons [15,16]. Ideally, a standardised model should be created through a validation study. Unfortunately, this was outside the scope of this study. We therefore opted to leave the choice of the artificial model of the urinary tract up to the users of the tool as long as it was described in detail in their manuscript.

Future Perspectives

As technology evolves, new fURS will be developed that might have characteristics that are not yet evaluated by this tool. Therefore, the TULIP tool may need to be refined or updated in the future. Nonetheless, according to all available literature to date, the TULIP tool implements all important scope characteristics in a single user-friendly tool.

Our hope is that the TULIP tool will create a more structured and systematic approach for future research. The next step would be to validate this tool to see if the outcome is reproducible and if the incorporated non-validated questionnaires are correctly interpreted. Extensive use of the

TULIP tool will facilitate meta-analysis of study outcomes. Furthermore, larger databases will create more meaningful insight into the quality of all characteristics of a fURS and enable comparison of different scopes. Extensive use of the proposed standardised methods of evaluation, which take the specifics that are provided in the TULIP tool on settings, tools and units of outcome into account, can lead to a more 'objective' standardised assessment. Consequently this might compensate for surgeon-specific confounding factors in (sub)categories that are inherently subjective. In this way, further insight will be given on the quality of different fURS.

Conclusion

This Delphi consensus project resulted in full consensus (>80% high agreement) on a standardised grading tool for the evaluation of fURS and the creation of the TULIP tool. This is a first step in creating uniformity in this field of research and to facilitate future comparison of outcomes.

Acknowledgements

We would like to thank the CoRE, EDGE, ESUT and PETRA working groups and their members for the participation in this modified Delphi consensus project. Furthermore, we would like to express our sincere gratitude towards our fellow PhD candidates Ben-Max de Ruiter, Luigi van Riel and Hilin Yildirim for reviewing the questionnaires.

Author Contributions

Both Michaël M.E.L. Henderickx and Nora Hendriks contributed equally and share first authorship. Michaël M.E.L. Henderickx: protocol/project development, data collection or management, data analysis, manuscript writing/editing. Nora Hendriks: protocol/project development, data collection or management, data analysis, manuscript writing/editing. Joyce Baard: protocol/project development, manuscript writing/editing. Oliver J. Wiseman: data collection or management, manuscript writing/editing. Kymora B. Scotland: data collection or management, manuscript writing/editing. Bhaskar K. Somani: data collection or management, manuscript writing/editing. T. Emre Şener: data collection or management, manuscript writing/editing. Esteban Emiliani: data collection or management, manuscript writing/editing. Laurian B. Dragos: data collection or management, manuscript writing/editing. Luca Villa: data collection or management, manuscript writing/editing. Michele Talso: data collection or management, manuscript writing/editing. Saeed Bin Hamri: data collection or management, manuscript writing/editing. Silvia Proietti: data collection or management, manuscript writing/editing. Steeve Doizi: data collection or management, manuscript writing/editing. Olivier Traxer: data collection or management, manuscript writing/editing. Ben H. Chew: data collection or management, manuscript writing/editing. Brian H. Eisner: data collection or management, manuscript writing/editing. Manoj Monga: data collection or management, manuscript writing/editing. Ryan S. Hsi: data collection or management, manuscript writing/editing. Karen L. Stern: data collection or management, manuscript writing/editing. David A. Leavitt: data collection or management, manuscript writing/editing. Marcelino Rivera: data collection or management, manuscript writing/editing. Daniel A. Wollin: data collection or management, manuscript writing/editing. Michael Borofsky: data collection or management, manuscript writing/editing. Noah E. Canvasser: data collection or management, manuscript writing/editing. Johann P. Ingimarsson: data collection or management, manuscript writing/editing. Marawan M. El Tayeb: data collection or management, manuscript writing/editing. Naeem Bhojani: data collection or management, manuscript writing/editing. Nariman Ghadziev: data collection or management, manuscript writing/editing. Thomas Tailly: data collection or management, manuscript writing/editing. Otas Durutovic: data collection or management, manuscript writing/editing. Udo Nagele: data collection or management, manuscript writing/editing. Andreas Skolarikos: data collection or management, manuscript writing/editing. Barbara M.A. Schout: protocol/project development, manuscript writing/editing. Harrie P. Beerlage: manuscript writing/editing. Rob C.M. Pelger: manuscript writing/editing. Guido M. Kamphuis: protocol/project development, data collection or management, manuscript writing/editing.

Disclosure of Interests

The authors declare that they have no disclosure of interests.

Funding

None.

References

1 **Georgescu DR, Multescu R, Geavlete PA, Geavlete B.** *Retrograde Ureteroscopy*, 1st edn. San Diego, CA: Elsevier, 2016

2 **Dragos LB, Somani BK, Keller EX et al.** Characteristics of current digital single-use flexible ureteroscopes versus their reusable counterparts: an in-vitro comparative analysis. *Transl Androl Urol* 2019; 8: S359–70

3 **Keller EX, De Coninck V, Traxer O.** Next-generation fiberoptic and digital ureteroscopes. *Urol Clin North Am* 2019; 46: 147–63

4 **Zeng G, Traxer O, Zhong W et al.** International alliance of urolithiasis guideline on retrograde intrarenal surgery. *BJU Int* 2022. https://doi.org/10.1111/bju.15836

5 **Mazzucchi E, Marchini GS, Berto FCG et al.** Single-use flexible ureteroscopes: update and perspective in developing countries. A narrative review. *Int Braz J Urol* 2022; 48: 456–67

6 **Dragos LB, Somani BK, Sener ET et al.** Which flexible ureteroscopes (digital vs. fiber-optic) can easily reach the difficult lower pole calices and have better end-tip deflection: in vitro study on K-box. A PETRA evaluation. *J Endourol* 2017; 31: 630–7

7 **Hendriks N**, **Henderickx MMEL**, **Schout BMA et al.** How to evaluate a flexible ureterorenoscope? Systematic mapping of existing evaluation methods. *BJU Int* 2021; 128: 408–23

8 **Jour I**, **Lam A**, **Turney B**. Urological stone disease: a 5-year update of stone management using hospital episode statistics. *BJU Int* 2022; 130: 364–9

9 **Jones J**, **Hunter D**. Qualitative research: consensus methods for medical and health services research. *BMJ* 1995; 311: 376–80

10 **Linstone HA**, **Turoff M**. *The Delphi Method: Technique and Applications*, 1st edn. London: Addison Wesley Publishing Company, 1975

11 **Liberati A**, **Altman DG**, **Tetzlaff J et al.** The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ* 2009; 339: b2700

12 **Zondervan PJ**, **Wagstaff PGK**, **Desai MM et al.** Follow-up after focal therapy in renal masses: an international multidisciplinary Delphi consensus project. *World J Urol* 2016; 34: 1657–65

13 **von der Gracht HA**. Consensus measurement in Delphi studies. *Technol Forecast Soc Change* 2012; 79: 1525–36

14 **Zeng G**, **Zhao Z**, **Mazzon G et al.** European association of urology section of urolithiasis and international alliance of urolithiasis joint consensus on retrograde intrarenal surgery for the management of renal stones. *Eur Urol Focus* 2021. (in press). https://doi.org/10.1016/j.euf.2021.10.011

15 **Villa L**, **Somani BK**, **Sener TE et al.** Comprehensive flexible ureteroscopy (FURS) simulator for training in endourology: the K-box model. *Cent Eur J Urol* 2016; 69: 118–20

16 **Cloutier J**, **Traxer O**. Do high-fidelity training models translate into better skill acquisition for an endourologist? *Curr Opin Urol* 2015; 25: 143–52

Correspondence: Michaël M.E.L. Henderickx, Department of Urology, Amsterdam UMC Location University of Amsterdam, Meibergdreef 9, Amsterdam, The Netherlands.

e-mail: m.m.henderickx@amsterdamumc.nl

Abbreviations: AERG, Amsterdam Endourology Research Group; CoRE, Collaborative for Research in Endourology; EDGE, Endourology Disease Group for Excellence; ESUT, European Association of Urology Section of Uro-Technology; fURS, flexible ureterorenoscope(s); PETRA, Progress in Endourology, Technology and Research Association; Q1, Questionnaire 1; Q2, Questionnaire 2; Q3, Questionnaire 3; Q4, Questionnaire 4; Q5, Questionnaire 5; TULIP, The Uniform grading tooL for flexIble ureterorenoscoPes.