

How reliable is endoscopic stone recognition?

Citation for published version (APA):

Henderickx, M. M. E. L., Stoots, S. J. M., de Bruin, D. M., Wijkstra, H., Freund, J. E., Wiseman, O., Ploumidis, A., Skolarikos, A., Somani, B. K., Sener, T. E., Emiliani, E., Dragos, L., Villa, L., Talso, M., Daudon, M., Traxer, O., Kronenberg, P., Doizi, S., Taily, T., ... Kamphuis, G. M. (2022). How reliable is endoscopic stone recognition? A comparison between visual stone identification and formal stone analysis. *Journal of Endourology*, 36(10), 1362-1370. Advance online publication. <https://doi.org/10.1089/end.2022.0217>

Document license:

TAVERNE

DOI:

[10.1089/end.2022.0217](https://doi.org/10.1089/end.2022.0217)

Document status and date:

Published: 01/10/2022

Document Version:

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.



How Reliable Is Endoscopic Stone Recognition? A Comparison Between Visual Stone Identification and Formal Stone Analysis

Michaël M.E.L. Henderickx, MD, FEBU,¹ Simone J.M. Stoots, MD,¹ Daniel M. De Bruin, MSc, PhD,¹
Hessel Wijkstra, MSc, PhD,^{1,2} Jan Erik Freund, MD, PhD,³ Oliver J. Wiseman, MA, FRCS (Urol),⁴
Achilles Ploumidis, MD, BSc, MSc, PhD, FEBU,⁵ Andreas Skolarikos, MD,⁶
Bhaskar K. Somani, MRCS, DM, FRCS (Urol), FEBU,⁷ Tarik Emre Şener, MD,⁸ Esteban Emiliani, MD, PhD,⁹
Laurian B. Dragos, MD,⁴ Luca Villa, MD,¹⁰ Michele Talso, MD,¹¹ Michel Daudon, MD, PhD,¹²
Olivier Traxer, MD, PhD,^{12,13} Peter Kronenberg, MD, FEBU,¹⁴ Steeve Doizi, MD, MSc,^{12,13}
Thomas Taily, MD, PhD, FEBU,¹⁵ Tzevat Tefik, MD,¹⁶ Nora Hendriks, MD,¹ Harrie P. Beerlage, MD, PhD,¹
Joyce Baard, MD, MSc,¹ and Guido M. Kamphuis, MD, MSc¹

Abstract

Objective: To assess the diagnostic accuracy and intra-observer agreement of endoscopic stone recognition (ESR) compared with formal stone analysis.

Introduction: Stone analysis is a corner stone in the prevention of stone recurrence. Although X-ray diffraction (XRD) and infrared spectroscopy are the recommended techniques for reliable formal stone analysis, this is not always possible, and the process takes time and is costly. ESR could be an alternative, as it would give immediate information on stone composition.

Materials and Methods: Fifteen endourologists predicted stone composition based on 100 videos from ureterorenoscopy. Diagnostic accuracy was evaluated by comparing the prediction from visual assessment with stone analysis by XRD. After 30 days, the videos were reviewed again in a random order to assess intra-observer agreement.

Results: The median diagnostic accuracy for calcium oxalate monohydrate was 54% in questionnaire 1 (Q1) and 59% in questionnaire 2 (Q2), whereas calcium oxalate dihydrate had a median diagnostic accuracy of 75% in Q1 and 50% in Q2. The diagnostic accuracy for calcium hydroxyphosphate was 10% in Q1 and 13% in Q2. The median diagnostic accuracy for calcium hydrogen phosphate dihydrate and calcium magnesium phosphate was 0% in both questionnaires. The median diagnostic accuracy for magnesium ammonium phosphate was 20%

¹Department of Urology, Amsterdam UMC Location University of Amsterdam, Amsterdam, the Netherlands.

²Biomedical Diagnostics Lab, Department of Electrical Engineering, Eindhoven University of Technology, Eindhoven, the Netherlands.

³Department of Pathology, Amsterdam UMC, University of Amsterdam, Amsterdam, the Netherlands.

⁴Department of Urology, Addenbrooke's Hospital, Cambridge University Hospitals NHS Foundation Trust, Cambridge, United Kingdom.

⁵Department of Urology, Athens Medical Centre, Athens, Greece.

⁶Department of Urology, National and Kapodistrian University of Athens, Sismanogleion Hospital, Athens, Greece.

⁷Department of Urology, University Hospital Southampton NHS Trust, Southampton, United Kingdom.

⁸Department of Urology, Marmara University School of Medicine, Istanbul, Turkey.

⁹Department of Urology, Fundació Puigvert, Autonomus University of Barcelona, Barcelona, Spain.

¹⁰Department of Urology, IRCCS Ospedale San Raffaele, Milan, Italy.

¹¹Department of Urology, ASST Fatebenefratelli-Sacco-Luigi Sacco University Hospital, Milan, Italy.

¹²Sorbonne Université, GRC n°20, Groupe de Recherche Clinique sur la Lithiase Urinaire, Hôpital Tenon, Paris, France.

¹³Sorbonne Université, Service d'Urologie, AP-HP, Hôpital Tenon, Paris, France.

¹⁴Department of Urology, Hospital CUF Descobertas, Lisbon, Portugal.

¹⁵Department of Urology, UZ Ghent, Ghent, Belgium.

¹⁶Department of Urology, Istanbul Faculty of Medicine, Istanbul University, Istanbul, Turkey.

in Q1 and 40% in Q2. The median diagnostic accuracy for uric acid was 22% in both questionnaires. Finally, there was a diagnostic accuracy of 60% in Q1 and 80% in Q2 for cystine. The intra-observer agreement ranged between 45% and 72%.

Conclusion: Diagnostic accuracy of ESR is limited and intra-observer agreement is below the threshold of acceptable agreement.

Keywords: endourology, stone composition, urinary stone analysis, ureterorenoscopy, endoscopic stone recognition

Introduction

UROLITHIASIS IS A common medical condition worldwide. Prevalence ranges from 5% to 14% in Europe, 7% to 13% in the United States, and 1% to 5% in Asia and these numbers keep rising.¹

The most common types of urolithiasis can be divided into eight groups: calcium oxalate monohydrate (COM), calcium oxalate dihydrate (COD), calcium hydroxyphosphate (CHP), calcium hydrogen phosphate dihydrate (CHPD), calcium magnesium phosphate (CMP), magnesium ammonium phosphate (MAP), uric acid (UA), and cystine. Calcium oxalate stones (COD and COM) represent around 70% to 80% of all urolithiasis, whereas calcium phosphate (CaPh) stones (CHP, CHPD, and CMP) represent around 15%, UA stones 8%, cystine stones 1% to 2%, and MAP stones 1%.² The incidence of pure stones varies greatly between studies, ranging from 5% to 59% depending on the study.^{3–7}

Knowing the stone's composition leads to a better etiological approach to treat and prevent urolithiasis. If a stone sample is available, then X-ray diffraction (XRD) and infrared spectroscopy (IRS) are the recommended techniques to accurately identify stone composition and its relative proportions.² However, with the advent of dusting techniques used to break up stones, a sample is not always available to send for analysis. In 1993, Daudon and colleagues described a stone classification based on the morphoconstitutional aspects of urolithiasis and their possible pathophysiology.^{8–10} Morphoconstitutional assessment has an essential role in the etiological diagnosis of urolithiasis and although this classification is based on the appearance of *ex vivo*, dry urolithiasis, Estrade and colleagues showed that endoscopic stone recognition (ESR) is feasible.¹¹

Nevertheless, this was a single-surgeon study and intra-observer agreement was not assessed. Therefore, the present international multicenter study evaluates whether endourologists can accurately predict stone composition based on visual appearance during endoscopic stone treatment when compared with formal stone analysis and if their assessment is reproducible.

Materials and Methods

Ethics

This study was performed according to the ethical standards described in the 1964 Declarations of Helsinki and its later amendments and was approved by the local ethics committee (W20_212 #20.245 on May 14, 2020).

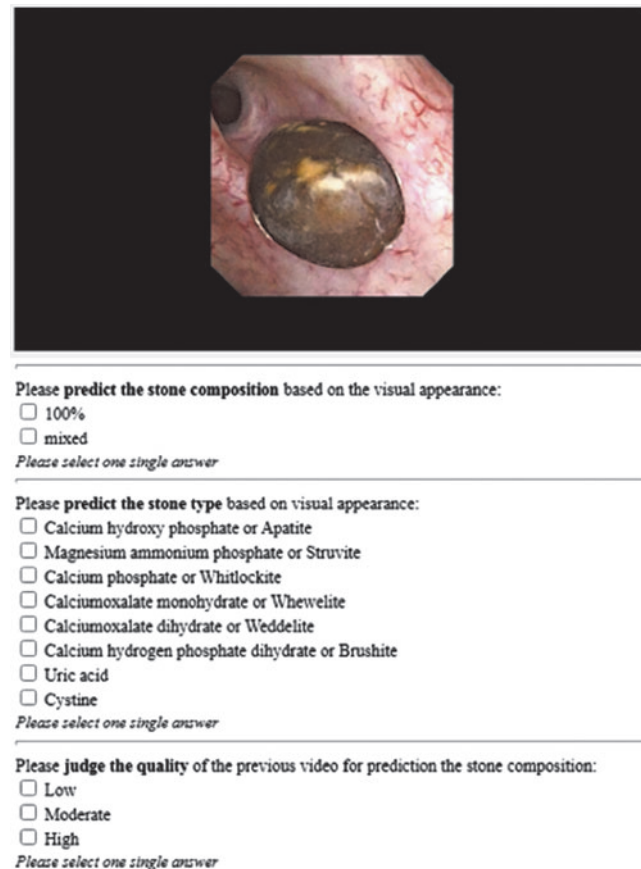
Study design

In this international multicenter study, 15 endourologists assessed stone composition based on the endoscopic ap-

pearance of urolithiasis in videos to evaluate diagnostic accuracy and intra-observer agreement. The videos were embedded in a questionnaire on an encrypted web-based platform (Data Management System; T&S Innovations, Maarssen, the Netherlands: <https://ts-innovations.com>). Figure 1 shows the questionnaires' interface.

The raters were instructed to view the videos on a laptop or stand-alone with a screen of minimally 13" in a quiet environment to avoid disturbance. Optimal video quality could be obtained by adjusting the size of the browser screen.

At the beginning of questionnaire 1 (Q1), information about the raters' experience (years practicing urology, years practicing stone treatment, and number of procedures per year) was collected. Then, each video was played in a loop and could be viewed unlimitedly. After assessment, the raters determined whether it was a pure or mixed stone. When raters decided on a pure composition, they had to identify the component subsequently. If they opted for a mixed composition, they had



The image shows a screenshot of a web-based questionnaire interface. At the top, there is a circular inset showing an endoscopic view of a dark, rounded stone within a urinary tract. Below the image, the text reads: "Please predict the stone composition based on the visual appearance:" followed by two radio button options: "100%" and "mixed". Below these options is the instruction "Please select one single answer".

The next section is titled "Please predict the stone type based on visual appearance:" and lists ten radio button options: "Calcium hydroxy phosphate or Apatite", "Magnesium ammonium phosphate or Struvite", "Calcium phosphate or Whitlockite", "Calciumoxalate monohydrate or Whewellite", "Calciumoxalate dihydrate or Weddelite", "Calcium hydrogen phosphate dihydrate or Brushite", "Uric acid", and "Cystine". Below this list is the instruction "Please select one single answer".

The final section is titled "Please judge the quality of the previous video for prediction the stone composition:" and has three radio button options: "Low", "Moderate", and "High". Below this is the instruction "Please select one single answer".

FIG. 1. Questionnaires' interface. Color images are available online.

to identify the main and secondary component, as well as a tertiary component if deemed applicable. The exact numbers and relative proportions of the included urolithiasis as well as any other clinical information were unknown to the raters. Finally, they were asked to score the video quality on a 3-point Likert scale (1 = low, 2 = moderate, 3 = high) for each video.¹²

After a wash-out period of 30 days, the raters assessed the same videos in a random order in a second questionnaire (Q2) to assess intra-observer agreement.¹²

Data acquisition

All videos were prospectively collected during consecutive endoscopic stone treatment with a digital flexible ureterorenoscope (FLEX-XC[®]; Karl Storz, Tuttlingen, Germany) at a tertiary referral center (Amsterdam UMC, the Netherlands). The procedures were recorded with RVC Clinical Assistant[®] (RVC, Amersfoort, the Netherlands). After surgery, the recordings were screen captured with Snagit[™] (TechSmith, Okemos, MI) and edited with Apple iMovie V.10.2.3. (Apple, Inc., Cupertino, CA) to obtain qualitative full HD (1080p) videos of 5 to 15 seconds.

Stones were treated with a combination of dusting (20 Hz–0.2 J) and fragmentation (10 Hz–1 J) with a laser VersaPulse[®] PowerSuite[™] 100W laser (Lumenis Ltd., Borehamwood, United Kingdom) in the majority of the cases. However, some of the included stones were also extracted in toto. The extracted stone fragments were collected and sent to the department of Clinical chemistry of the Erasmus MC (Rotterdam, the Netherlands) for formal stone analysis with XRD to analyze the stone's composition. XRD was performed according to the general guidelines. First, the extracted fragments were pulverized into dust, after which monochromatic X-rays were used to identify the constituents of the sample.¹³

Main component, as well as secondary and tertiary components if applicable, were determined. Pure stones were defined as stones consisting of only one composition on formal stone analysis with XRD. Not all stone compositions and relative proportions are as common. The respective prevalence in this study reflects the prevalence in an academic clinical practice, as data from consecutive procedures was collected. This is comparable with studies describing the prevalence in France, Norway, and the United States.^{3,6,7}

Statistical analysis

After consulting our institution's statistician, we concluded that a power-analysis to determine the sample size was not possible, due to the lack of publications on this topic. Based on practical grounds and on a similar diagnostic study by Freund and colleagues, we decided to include 100 videos.¹²

SPSS V.28 (IBM Corp, Armonk, NY) was used to perform the statistical analysis. Figures and tables were created with Microsoft[®] Excel (Microsoft Corp., Redmond, WA).

Descriptive analysis was used to assess rater experience and quality of the videos as a median (range) for the whole group.

The diagnostic accuracy for each stone composition was assessed by calculating the concordance of the rater's visual assessment for the main component with the main chemical composition identified with XRD. Results are described as a median (range) for the whole group. In addition, diagnostic accuracy for pure and mixed stones was calculated.

The intra-observer agreement was defined as the percentage of cases where the raters predicted the same main component in both questionnaires. Acceptable agreement was defined as minimally 80% agreement.^{12,14}

Spearman's rank correlation coefficient was calculated for the experience of each rater and the diagnostic accuracy.

The influence of video quality was assessed by comparing the diagnostic accuracy for all videos with the diagnostic accuracy for videos rated of a high or intermediate quality. A Mann–Whitney *U*-test was applied to determine statistical significance between these groups.

A two-sided *p*-value ≤ 0.05 was considered statistically significant.

Results

Rater experience

At the time of assessment, the raters had been in practice for a median of 10 years (5–20 years), with a median endourological experience of 8 years (5–18 years) and an average caseload of ~120 ureteroscopies for stone treatment per year (70–300 procedures).

Diagnostic accuracy

The diagnostic accuracy for each chemical composition is presented in Table 1 and as simple boxplots in Figure 2.

The median diagnostic accuracy for stones with COM as main component ($n=39$) was 54% (44%) in Q1 and 59% (61%) in Q2. Stones with COD as main component ($n=4$) had a median diagnostic accuracy of 75% (100%) in Q1 and 50% (75%) in Q2.

Stones with CHP as main component ($n=30$) had a median diagnostic accuracy of 10% (43%) in Q1 and 13% (53%) in Q2. The median diagnostic accuracy for CHPD as main component ($n=8$) was 0% (38%) in Q1 and 0% (50%) in Q2. One stone had CMP as main component; the median diagnostic accuracy was 0% (100) in both questionnaires.

The median diagnostic accuracy for stones with MAP as main component ($n=4$) was 20% (80%) in Q1 and 40% (60%) in Q2.

TABLE 1. MEDIAN (RANGE)
OF THE DIAGNOSTIC ACCURACY

Composition	n	Median (range)	
		Q1	Q2
COM	39	54% (44)	59% (61)
COD	4	75% (100)	50% (75)
CHP	30	10% (43)	13% (53)
CHPD	7	0% (38)	0% (50)
CMP	1	0% (100)	0% (100)
MAP	5	20% (80)	40% (60)
UA	9	22% (45)	22% (45)
Cystine	5	60% (100)	80% (80)

CHP=calcium hydroxyphosphate; CHPD=calcium hydrogen phosphate dihydrate; CMP=calcium magnesium phosphate; COD=calcium oxalate dihydrate; COM=calcium oxalate monohydrate; MAP=magnesium ammonium phosphate; Q1=questionnaire 1; Q2=questionnaire 2; UA=uric acid.

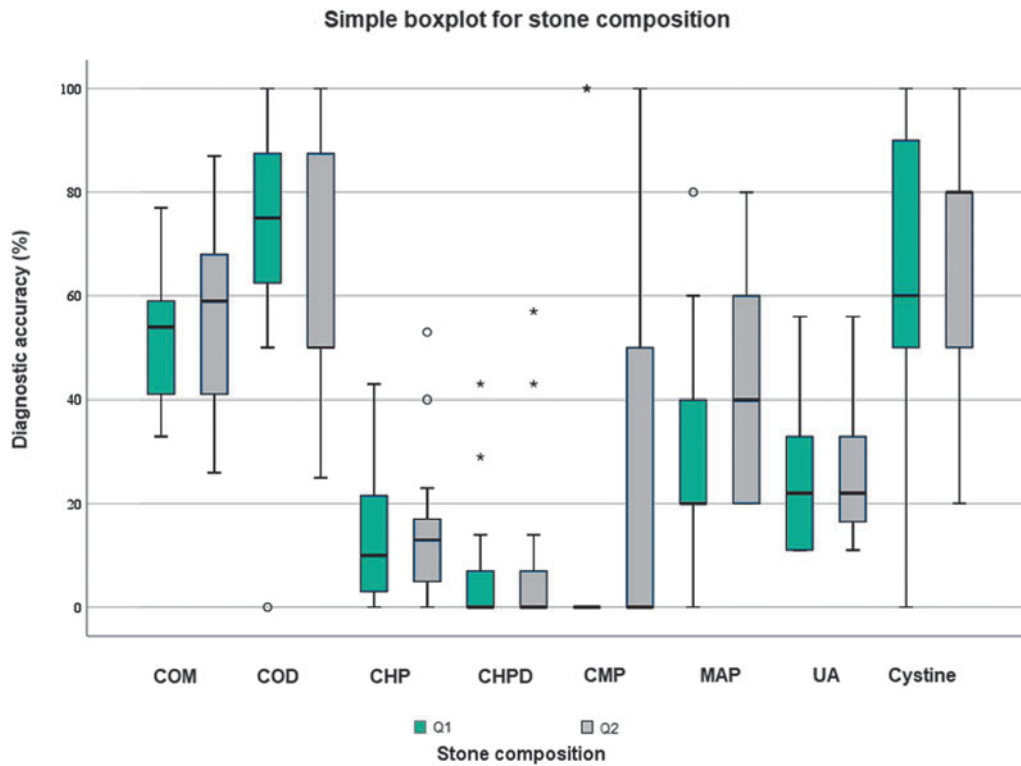


FIG. 2. Simple boxplot for correctly predicted composition per chemical component. CHP=calcium hydroxyphosphate; CHPD=calcium hydrogen phosphate dihydrate; CMP=calcium magnesium phosphate; COD=calcium oxalate dihydrate; COM=calcium oxalate monohydrate; MAP=magnesium ammonium phosphate; Q1=questionnaire 1; Q2=questionnaire 2; UA=uric acid. Color images are available online.

The median diagnostic accuracy for UA as main component ($n=9$) was 22% (45%) in both questionnaires.

Finally, stones with cystine as main component ($n=5$) revealed a median diagnostic accuracy of 60% (100%) in Q1 and 80% (80%) in Q2.

A sub-analysis for pure stones ($n=33$) was performed. Results are shown in Table 2 and Figure 3.

A sub-analysis for mixed stones ($n=67$) was also performed. Results are presented in Table 3 and Figure 4.

Additional analysis for subgroups (diagnostic accuracy for pure and mixed stones, calcium (COM, COD, CHP, CHPD, and/or CMP) and noncalcium (MAP, UA, and/or cystine) stones, calcium oxalate (COM and/or COD) and calcium phosphate (CHP, CHPD and/or CMP) stones, as well as for infection (CHP and/or MAP) stones) is presented in Supplementary Appendix SA1.¹⁵ Supplementary Appendix SA2 shows the results per rater.

Intra-observer agreement

The median intra-observer agreement calculated as simple percentage agreement was 56% (27%) (Fig. 5). This is below the predefined threshold of 80% to classify as acceptable agreement.^{12,14} The Cohen's kappa coefficient was 0.461 (SE 0.015, $p=0.000$). Hence, the overall inter-observer agreement was moderate.

Experience and diagnostic accuracy

Experience with endourology was not correlated with a higher diagnostic accuracy. Specifics can be found in Supplementary Appendix SA3.

Video quality

The overall median video quality based on the 3-point Likert scale for Q1 was 2.2 (1.6–2.8) and 2.1 (1.4–2.9) for Q2. No statistically significant difference was found in diagnostic accuracy when videos rated as low quality were omitted (Supplementary Appendix SA4).

Discussion

According to the EAU guidelines, medical treatment and prevention of urolithiasis is based on basic metabolic evaluation and reliable stone-analysis. XRD or IRS are recommended for stone analysis.² While corroborating this

TABLE 2. MEDIAN (RANGE) OF THE DIAGNOSTIC ACCURACY IN PURE STONES

		<i>Diagnostic accuracy for pure stones</i>	
		<i>Median (range)</i>	
<i>Composition</i>	<i>n</i>	<i>Q1</i>	<i>Q2</i>
COM	17	65% (47)	65% (65)
COD	0	NA	NA
CHP	2	0% (100)	0% (100)
CHPD	2	0% (50)	0% (100)
CMP	0	NA	NA
MAP	0	NA	NA
UA	7	29% (57)	29% (43)
Cystine	5	60% (100)	80% (80)

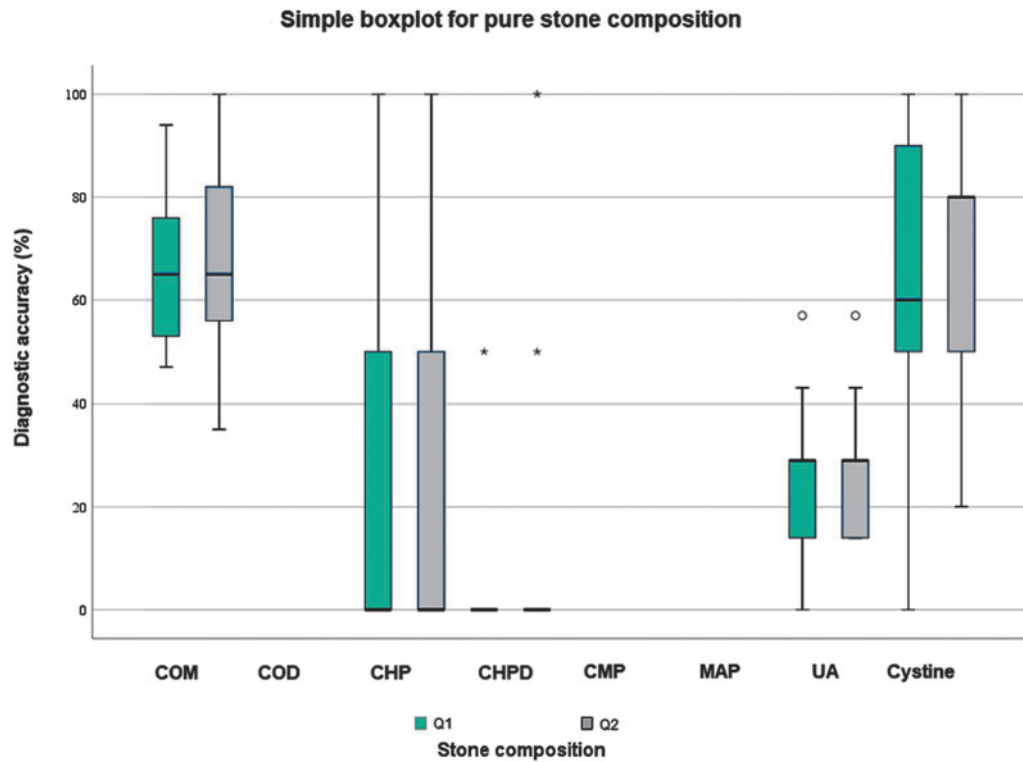


FIG. 3. Simple boxplot for correctly predicted composition per chemical component in pure stones. Color images are available online.

recommendation, a recent international consensus meeting emphasized the importance of visual identification of stone morphology to avoid sampling error.¹⁶

The present study demonstrates the diagnostic limitations of ESR. COM, COD, and cystine were correctly recognized in more than 50% of the cases. In contrast, CHP, CHPD, CMP, MAP, and UA all had a diagnostic accuracy of less than 40%. Further, intra-observer agreement was below the threshold of acceptable agreement.

The ESR could be beneficial to preserve an etiological approach in the treatment and prevention of urolithiasis. This approach might not be possible when postoperative analysis of stone dust provides insufficient information.¹⁷ Further, there could be a difference between the composition of the stone surface and core. The composition of the surface rep-

resents recent stone formation, and stone treatment with laser will unavoidably lead to loss of information on the etiology of the stone.¹¹

Moreover, Keller and colleagues described that high-frequency dusting with the Thulium fiber laser can lead to changes in the infrared spectra, thus possibly biasing stone analysis through IRS.^{17,18} Not only postoperative prevention of urolithiasis can be aided by ESR, but also the preoperative treatment strategy (laser settings, acceptance of residual fragments, postoperative policy, and further direct medical management) might be influenced depending on the composition. Laser setting, for example, will depend on the hardness of the urolithiasis. Even though high-power lasers can destroy any type of urolithiasis, harder stone will require greater potency. ESR can help the surgeon to choose the right settings to treat the urolithiasis.^{18,19}

Another possible preoperative influence of ESR is the choice to accept residual fragments. Endourologists might accept small residual fragments in case they are treating a calcium oxalate stone, whereas they will not accept residual fragments when they are treating infection stones. This is because the risk of recurrence is higher in patients with residual fragments after the treatment of infection stones than for other stones.^{2,20}

Daudon and colleagues developed a classification system to summarize the principal etiological causes for urolithiasis formation based on their morphoconstititional aspects that has been a guide for urologists all over the world.^{8,10} This classification system, however, was based solely on the subjective assessment of the macroscopic appearance of *ex vivo*, dry urolithiasis. Estrade and colleagues evaluated the feasibility ESR of commonly encountered urolithiasis.

TABLE 3. MEDIAN (RANGE) OF THE DIAGNOSTIC ACCURACY IN MIXED STONES

<i>Diagnostic accuracy for mixed stones</i>			
<i>Composition</i>	<i>n</i>	<i>Median (range)</i>	
		<i>Q1</i>	<i>Q2</i>
COM	22	45% (46)	50% (68)
COD	4	75% (100)	50% (75)
CHP	28	7% (39)	7% (54)
CHPD	6	0% (40)	0% (40)
CMP	1	0% (100)	0% (100)
MAP	4	20% (80)	40% (60)
UA	2	0% (50)	50% (50)
Cystine	0	NA	NA

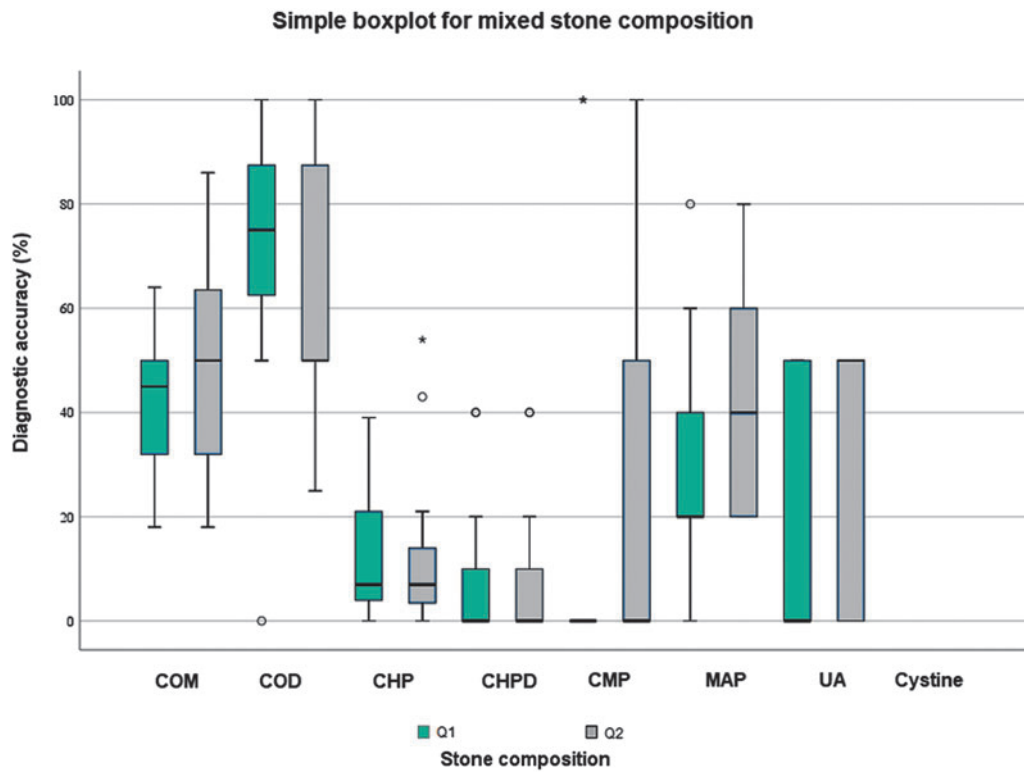


FIG. 4. Simple boxplot for correctly predicted stone composition per chemical component in mixed stones. Color images are available online.

This study had one experienced urologist evaluate 399 stones during endoscopic stone treatment and found a concordance between microscopy and IRS of 81.6%. Diagnosis was confirmed for COM, COD, UA, CHP–MAP association, and CHPD stones.¹¹ In contrast to our study, this was a

single-surgeon study, and they did not evaluate the intra-observer agreement.

Estrade’s group also evaluated whether the recognition of stone compositions could be learned. Fifteen urologists-in-training were asked to recognize and describe the

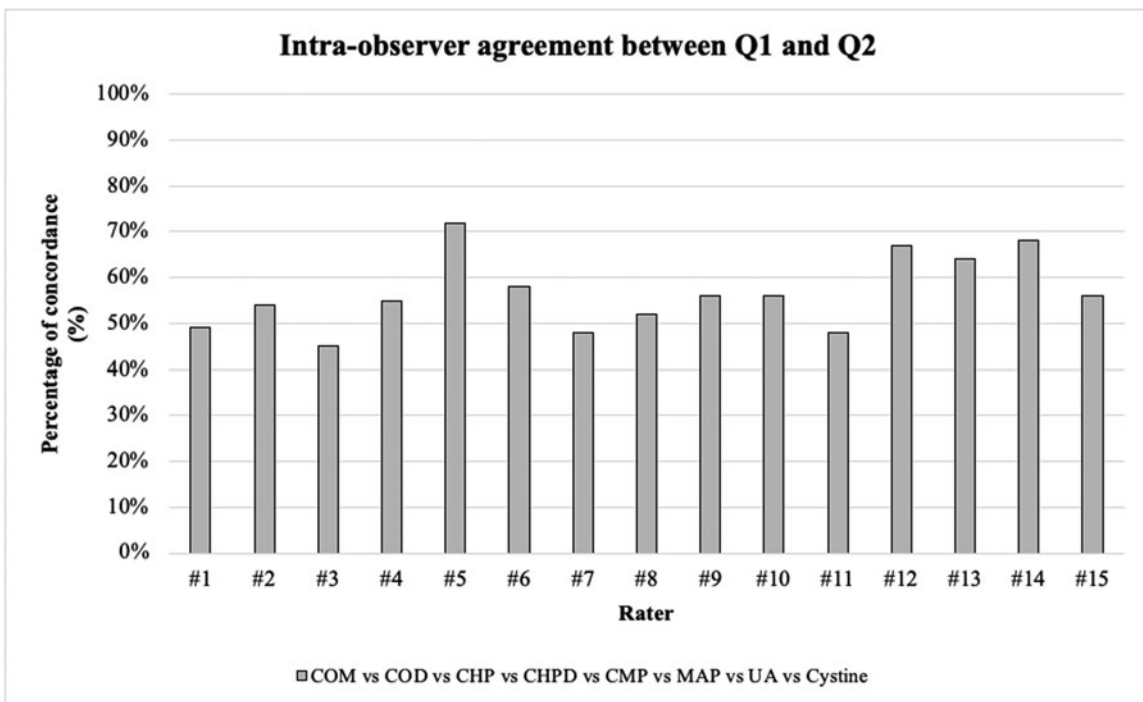


FIG. 5. Intra-observer agreement for stone composition.

morphology of nine stones before and after expert coaching in their study and they concluded that it was possible to train urologists in stone recognition based on their visual appearance.²¹ Our study tried to bypass the said learning curve by having experienced endourologists perform the assessment. Further, we did not find that experience was correlated with a higher accuracy in the prediction of stone composition.

This might be due to the high level of experience of the participating endourologists, and the involvement of less experienced raters may lead to another conclusion. Further, it is still possible that dedicated training on visual identification of stone composition might lead to better results. Further research is needed to confirm these hypotheses.

Similar to our study, Sampogna and colleagues described a limited diagnostic accuracy for ESR. Their study, where 32 urologists predicted the composition of 20 stones based on videos on YouTube, showed a diagnostic accuracy of 69.8% for COD and 78.1% for cystine, respectively 75% (50%) and 60% (80%) in our study. However, their results for COM were worse with a diagnostic accuracy of 41.8% vs a diagnostic accuracy of 54% and 59% in our study.²²

Black and colleagues were the first to describe a deep learning computer vision algorithm to recognize the composition of pure urolithiasis in 2020. They applied a neural network on digital photographs of 63 *ex vivo* urolithiasis (surface and section) and concluded that deep learning can be used to detect the composition of the most frequent urolithiasis. The neural network correctly predicted 94% of the UA, 90% of the COM, 86% of the MAP, 71% of the CHPD, and 75% of the cystine stones.²³ Similarly, following up on their previous studies, Estrade and colleagues also trained a neural network to predict stone composition.^{11,17}

In contrast to the study performed by Black and colleagues, *in vivo* endoscopic images of urolithiasis were used in this study. Their neural network analyzed pure COM, COD, and UA stones, as well as mixed COM–COD and COM–UA stones. They achieved high diagnostic accuracy for the surface of pure UA (98%) and pure COM (91% surface and 94% section). Mixed COM–COD and COM–UA had a lower, but still high sensitivity.¹⁷

Although different techniques, such as convolutional neural networks, dual energy CTs and ESR, have been studied to determine stone composition, none are as good as the recommended techniques for reliable formal stone analysis, XRD and IRS. Nonetheless, the results achieved with artificial intelligence are promising. More research and big data are needed to further improve this technique.

The present study holds limitations. A major limitation is the limited number of some stone types, which limits the power of this study. Moreover, the fact that only a fraction of the stone was analyzed could lead to sample bias. This is especially true in mixed stones and is unfortunately inherent to the current technique of formal stone analysis in stones that were not extracted and sent for analysis *in toto*. It has been shown that the diagnostic accuracy of IRS on stone fragments (80%) is lower than the diagnostic accuracy of the combination of microscopic assessment and IRS of entire stones (95%).²⁴

Further, Krambeck and colleagues²⁵ showed that there is a significant variability in the reporting of mixed stones when analyzed with IRS.²¹ Sutor had similar results in regards to

XRD.²⁶ As many stones these days are treated with a combination of fragmenting and dusting, the fragments sent for analysis may not be representative of the entire stone. One could send dust for formal stone analysis; however, the diagnostic accuracy of IRS for stone dust is even lower than for stone fragments (60% vs 80%).²⁴ A possible solution would be to start with visual identification to define what portion of the stone should be sent for formal stone analysis.¹⁶

Nonetheless, for this to work, urologists should be trained to visually identify different stone compositions and send the respective fragments for formal stone analysis. A way to bypass the difficulties of visual identification by urologists is the use of micro-CT. This technique was first described in 2001 by Cleveland and colleagues and has been shown to adequately recognize different stone compositions or regions of interest.^{27–33} As micro-CT has used formal stone analysis (IRS) as a benchmark to evaluate its specificity and sensitivity, additional formal analysis, such as IRS or XRD, is still necessary. Nonetheless, this technique can provide information about which regions of a stone should be further investigated. Further, Williams and colleagues stated that micro-CT is much easier than visual identification in the recognition of regions of interest.³⁴

In mixed stones, the difference in percentage composition between primary, secondary, and, if applicable, tertiary components, can vary. The lower the percentage of the primary component, the more difficult it might be to correctly predict the primary composition. However, it is unknown what the minimal percentage should be for ideal ESR. This study included stones where the primary component minimally made up 40% of the stone.

Further, even though the videos were carefully selected to optimally visualize the stones, more extensive fragments with different angles as well as images of both surface and sections of the stone may lead to a higher diagnostic accuracy.¹² Nevertheless, our dataset included mostly surface videos and prior studies have shown that images of stone surfaces are easier to correctly recognize.^{11,21}

The strength of this study lies in the international, multi-center design with videos rated of intermediate to high quality in a controlled questionnaire and a second assessment after 30 days to assess intra-observer agreement. Further, the raters were blinded to additional clinical information to evaluate a baseline diagnostic accuracy of ESR and avoid bias of the raters. Clinical information might increase the diagnostic accuracy even more; however, further research is needed to confirm this hypothesis.¹²

Conclusion

The visual appearance of urolithiasis during endoscopy did not allow accurate prediction of stone composition. The diagnostic accuracy was limited, and the intra-observer agreement was low.

Author's Contributions

M.M.E.L.H.: Protocol/project development, data collection or management, data analysis, and article writing/editing.

S.J.M.S.: Article writing/editing.

D.M.D.B.: Protocol/project development, data analysis, and article writing/editing.

H.W.: Protocol/project development, data collection or management, and article writing/editing.

J.E.F.: Protocol/project development, article writing/editing.

O.J.W.: Data collection or management, article writing/editing.

A.P.: Data collection or management, article writing/editing.

A.S.: Data collection or management, article writing/editing.

B.K.S.: Data collection or management, article writing/editing.

T.E.Ş.: Data collection or management, article writing/editing.

E.E.: Data collection or management, article writing/editing.

L.B.D.: Data collection or management, article writing/editing.

L.V.: Data collection or management, article writing/editing.

M.T.: Data collection or management, article writing/editing.

M.D.: Article writing/editing.

O.T.: Article writing/editing.

P.K.: Data collection or management, article writing/editing.

S.D.: Data collection or management, article writing/editing.

T.T.: Data collection or management, article writing/editing.

T.T.: Data collection or management, article writing/editing.

N.H.: Article writing/editing.

H.P.B.: Article writing/editing.

J.B.: Data collection or management, article writing/editing.

G.M.K.: Protocol/project development, data collection or management, article writing/editing.

Informed Consent

Informed consent was not obtained from individual participants included in the study as consent is not required as long as information is anonymized, and the submission does not include images that may identify the person. Therefore, our institution waived the need for informed consent.

Author Disclosure Statement

M.M.E.L.H. is a consultant for Coloplast; S.J.M.S. declares no conflicts of interest; D.M.D.B. has no conflicts of interest; H.W. declares no conflicts of interest; and J.E.F. declares no conflicts of interest. O.J.W. is a consultant for Boston Scientific, Coloplast, EMS, and Ambu, has received research funding from Coloplast and EMS, and has undertaken education for Boston Scientific, Coloplast, EMS, Ambu, and Olympus. A.P. declares no conflicts of interest; A.S. declares no conflicts of interest; B.K.S. is a consultant for Lumenis, Coloplast, Boston Scientific, and Pusen; T.E.Ş. declares no conflicts of interest; E.E. declares no conflicts of interest; L.B.D. declares no conflicts of interest; L.V. declares no conflicts of interest; M.T. declares no conflicts of interest; M.D. declares no conflicts of interest; O.T. declares no conflicts of interest; P.K. declares no conflicts of interest; S.D. declares no conflicts of interest; T.T. declares no conflicts of interest; T.T. declares no conflicts of interest; N.H. declares no conflicts of interest, H.P.B. declares no conflicts of interest; J.B. is a consultant for Coloplast; and G.M.K. is a consultant for Boston Scientific, EMS, Coloplast, and Olympus.

Research involved human participants and/or animals: This study was reviewed by and approved by the ethical commission of our hospital under reference W20_212 #20.245 on May 14, 2020.

Funding Information

No funding was received for this article.

Supplementary Material

Supplementary Appendix SA1

Supplementary Appendix SA2

Supplementary Appendix SA3

Supplementary Appendix SA4

References

1. Sorokin I, Mamoulakis C, Miyazawa K, et al. Epidemiology of stone disease across the world. *World J Urol* 2017; 35:1301–1320; doi: 10.1007/s00345-017-2008-6.
2. Türk C, Neisius A, Petřík A, et al. EAU Guidelines on Urolithiasis, 2021st ed. EAU Guidelines Office: Arnhem, the Netherlands; 2021.
3. Daudon M, Donsimoni R, Hennequin C, et al. Sex- and age-related composition of 10 617 calculi analyzed by infrared spectroscopy. *Urol Res* 1995;23:319–326; doi: 10.1007/BF00300021.
4. Zhang S, Huang Y, Wu W, et al. Trends in urinary stone composition in 23,182 stone analyses from 2011 to 2019: A high-volume center study in China. *World J Urol* 2021;39: 3599–3605; doi: 10.1007/s00345-021-03680-y.
5. Spivacow FR, Del Valle EE, Lores E, et al. Kidney stones: Composition, frequency and relation to metabolic diagnosis. *Medicina* 2016;76:343–348.
6. Katz JE, Soodana-Prakash N, Jain A, et al. Influence of age and geography on chemical composition of 98043 urinary stones from the USA. *Eur Urol open Sci* 2021;34:19–26; doi: 10.1016/j.euros.2021.09.011.
7. Kravdal G, Helgø D, Moe MK. Kidney stone compositions and frequencies in a Norwegian population. *Scand J Urol* 2019;53:139–144; doi: 10.1080/21681805.2019.1606031.
8. Corrales M, Doizi S, Barghouthy Y, et al. Classification of stones according to Michel Daudon: A narrative review. *Eur Urol Focus* 2021;7:13–21; doi: 10.1016/j.euf.2020.11.004.
9. Daudon M, Dessombz A, Frochot V, et al. Comprehensive morpho-constitutional analysis of urinary stones improves etiological diagnosis and therapeutic strategy of nephrolithiasis. *Comptes Rendus Chim* 2016;19:1470–1491; doi: 10.1016/j.crci.2016.05.008.
10. Daudon M, Bader CA, Jungers P. Urinary calculi: Review of classification methods and correlations with etiology. *Scanning Microsc* 1993;7:1081–1104; discussion 1104–1106.
11. Estrade V, Denis de Senneville B, Meria P, et al. Toward improved endoscopic examination of urinary stones: A concordance study between endoscopic digital pictures vs microscopy. *BJU Int* 2021;128:319–330; doi: 10.1111/bju.15312.
12. Freund JE, Legemate JD, Baard J, et al. Upper tract urothelial carcinoma grade prediction based on the ureteroscopic appearance: Caution should be taken. *Urology* 2019; 132:69–74; doi: 10.1016/j.urology.2019.05.046.
13. Singh VK, Rai PK. Kidney stone analysis techniques and the role of major and trace elements on their pathogenesis: A review. *Biophys Rev* 2014;6:291–310; doi: 10.1007/s12551-014-0144-4.
14. McHugh ML. Interrater reliability: The kappa statistic. *Biochem Med (Zagreb)* 2012;22:276–282.
15. Bichler K-H, Eipper E, Naber K, et al. Urinary infection stones. *Int J Antimicrob Agents* 2002;19:488–498; doi: 10.1016/s0924-8579(02)00088-2.

16. Williams JC, Gambaro G, Rodgers A, et al. Urine and stone analysis for the investigation of the renal stone former: A consensus conference. *Urolithiasis* 2021;49:1–16; doi: 10.1007/s00240-020-01217-3.
17. Estrade V, Daudon M, Richard E, et al. Towards automatic recognition of pure & mixed stones using intraoperative endoscopic digital images. *BJU Int* 2022;129:234–242; doi: 10.1111/bju.15515.
18. Keller EX, De Coninck V, Doizi S, et al. Thulium fiber laser: Ready to dust all urinary stone composition types? *World J Urol* 2021;39:1693–1698; doi: 10.1007/s00345-020-03217-9.
19. Villavicencio H. AUA University. V11-12: Holmium laser settings during lithotripsy of different types of kidney stones. *Urolithiasis* (2017 SVL). Available from: <https://auau.auanet.org/content/v11-12-holmium-laser-settings-during-lithotripsy-different-types-kidney-stones> [Last accessed: 5/29/2022].
20. Beck EM, Riehle RA. The fate of residual fragments after extracorporeal shock wave lithotripsy monotherapy of infection stones. *J Urol* 1991;145:6–9; discussion 9–10; doi: 10.1016/s0022-5347(17)38230-7.
21. Bergot C, Robert G, Bernhard J-C, et al. The basis of endoscopic stones recognition, a prospective monocentric study [in French]. *Prog Urol* 2019;29:312–317; doi: 10.1016/j.purol.2019.04.002.
22. Sampogna G, Basic D, Geavlete P, et al. Endoscopic identification of urinary stone composition: A study of South Eastern Group for Urolithiasis Research (SEGUR 2). *Actas Urol Esp* 2021;45:154–159; doi: 10.1016/j.acuro.2020.01.015.
23. Black KM, Law H, Aldoukhi A, et al. Deep learning computer vision algorithm for detecting kidney stone composition. *BJU Int* 2020;125:920–924; doi: 10.1111/bju.15035.
24. Estrade V, Daudon M, Traxer O, et al. Why should urologist recognize urinary stone and how? The basis of endoscopic recognition. *Prog Urol* 2017;27:F26–F35; doi: 10.1016/j.fpurol.2017.03.002.
25. Krambeck AE, Khan NF, Jackson ME, et al. Inaccurate reporting of mineral composition by commercial stone analysis laboratories: Implications for infection and metabolic stones. *J Urol* 2010;184:1543–1549; doi:10.1016/j.juro.2010.05.089.
26. Sutor DJ. Difficulties in the identification of components of mixed urinary calculi using the x-ray powder method. *Br J Urol* 2008;40:29–32; doi: 10.1111/j.1464-410X.1968.tb11809.x.
27. Cleveland RO, McAteer JA, Müller R. Time-lapse nondestructive assessment of shock wave damage to kidney stones in vitro using micro-computed tomography. *J Acoust Soc Am* 2001;110:1733–1736; doi: 10.1121/1.1401742.
28. Williams JC, Lingeman JE, Daudon M, et al. Using micro computed tomographic imaging for analyzing kidney stones. *Comptes Rendus Chim* 2021;25:1–12; doi: 10.5802/crchim.89.
29. Pramanik R, Asplin JR, Jackson ME, et al. Protein content of human apatite and brushite kidney stones: Significant correlation with morphologic measures. *Urol Res* 2008;36:251–258; doi: 10.1007/s00240-008-0151-7.
30. Zarse CA, McAteer JA, Sommer AJ, et al. Nondestructive analysis of urinary calculi using micro computed tomography. *BMC Urol* 2004;4:15; doi: 10.1186/1471-2490-4-15.
31. Zarse CA, Hameed TA, Jackson ME, et al. CT visible internal stone structure, but not Hounsfield unit value, of calcium oxalate monohydrate (COM) calculi predicts lithotripsy fragility in vitro. *Urol Res* 2007;35:201–206; doi: 10.1007/s00240-007-0104-6.
32. Williams JC, Borofsky MS, Bledsoe SB, et al. Papillary ductal plugging is a mechanism for early stone retention in Brushite stone disease. *J Urol* 2018;199:186–192; doi: 10.1016/j.juro.2017.08.063.
33. Primak AN, Fletcher JG, Vrtiska TJ, et al. Noninvasive differentiation of uric acid versus non-uric acid kidney stones using dual-energy CT. *Acad Radiol* 2007;14:1441–1447; doi: 10.1016/j.acra.2007.09.016.
34. Williams JC, McAteer JA, Evan AP, et al. Micro-computed tomography for analysis of urinary calculi. *Urol Res* 2010;38:477–484; doi: 10.1007/s00240-010-0326-x.

Address correspondence to:

Michaël M.E.L. Henderickx, MD, FEBU

Department of Urology

Amsterdam UMC Location University of Amsterdam

Meibergdreef 9

Amsterdam 1105 AZ

the Netherlands

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

Abbreviations Used

CHP = calcium hydroxyphosphate (apatite)
 CHPD = calcium hydrogen phosphate dihydrate (brushite)
 CMP = calcium magnesium phosphate (whitlockite)
 COD = calcium oxalate dihydrate (whewellite)
 COM = calcium oxalate monohydrate (weddelite)
 CT = computed tomography
 ESR = endoscopic stone recognition
 IRS = infrared spectroscopy
 MAP = magnesium ammonium phosphate (struvite)
 Q1 = questionnaire 1
 Q2 = questionnaire 2
 UA = uric acid
 XRD = X-ray diffraction