Chapter 6

Infrared analysis of urinary calculi, applying a single reflection accessory and a neural network interpretation algorithm

Marcel Volmer¹, Jules C.M. de Vries,² Henk M.J. Goldschmidt²

¹ Department of Pathology and Laboratory Medicine, University Hospital Groningen (UHG), P.O. Box 30001, 9700 RB Groningen, The Netherlands,
² Department of Clinical Chemistry and Hematology (CKCHL), St. Elisabeth Hospital, P.O. Box 10111, 5000 JC Tilburg, The Netherlands.

ABSTRACT

Background: Preparation of KBr tablets, used for Fourier infrared (FT-IR) analysis of urinary calculus composition, is time-consuming and often hampered by pellet breakage. Therefore, we developed a new FT-IR method for urinary calculus analysis. This method makes use of a Golden Gate Single Reflection Diamond Attenuated Total Reflection sample holder, a computer library, and an artificial neural network (ANN) for spectral interpretation.

Methods: The library was prepared from 25 pure components and 236 binary and ternary mixtures of the 8 most commonly occurring components. The ANN was trained and validated with 248 similar mixtures and tested with 92 patient samples, respectively.

Results: The optimum ANN model yielded root mean square errors of 1.5% and 2.3% for the training and validation sets, respectively. Fourteen simple expert rules were added to correct systematic network inaccuracies. Results of 92 consecutive patient samples were compared with those of a FT-IR method with KBr tablets based on an initial computerized library search followed by visual inspection. The bias was significantly different from zero for brushite (~0.8%) and the concomitantly occurring whewellite (~2.8%) and weddellite (3.8%), but not for ammonium hydrogen urate (~0.1%), carbonate apatite (0.5%), cystine (0.0%), struvite (0.4%), and uric acid (~0.1%). The 95% level of agreement of all results amounted to 9%.

Conclusions: The new Golden Gate method is superior because of its smaller sample size, user-friendliness, robustness and speed. Expert knowledge for spectral interpretation is minimized by the combination of library search and ANN prediction, but visual inspection remains necessary.

INTRODUCTION

Therapy to prevent urinary calculi recurrence requires quantitative estimates of the composition of urinary calculi. Extracorporal shock wave lithotripsy, now widely used for removal of urinary calculi, necessitates the use of laboratory techniques that allow component identification in minute amounts of material. Traditional wet chemistry techniques, X-ray diffraction and infrared (IR) spectroscopy are the current analytical methods. Of these, wet chemical analysis is rather inaccurate and imprecise (1) and requires relatively large amounts of sample. X-ray diffraction is suitable for quantification of mineral containing samples, such as urinary calculi (2), but it cannot adequately detect amorphous substances, such as carbonate apatite or dahlite (3). Infrared (IR) spectroscopy has been applied in clinical chemistry for analyses of biofluids and solid biosamples (4). This technique often produces complex spectra with contributions from a sizeable number of unknown interfering substances when applied to authentic biological material. Analyses of these complex spectra is facilitated by use of chemometrics (5), which is a generic term for the application of expert systems, neural networks, and other mathematical and statistical methods.

Fourier transform infrared (FT-IR) spectroscopy has become a standard technique for urinary calculus analysis. FT-IR makes use of a diversity of sample holders, such as photoacoustic detection (6), diffuse reflectance FT-IR (DRIFT) (7), and KBr tablet transmission (8;9). For the routine visual interpretation of urinary calculus IR spectra, Hesse et al. (10) have issued an atlas with IR spectra from pure urinary calculus...
components and their mixtures, all embedded in KBr tablets. Another, less time-consuming option is computerized analyses based on library search [e.g. SEARCH (11) and LITHOS (2)], expert rules [CIRCOM (12), STONES (9)] or other chemometrical techniques like partial least-squares (PLS) regression (13) and artificial neural networks (ANNs) (14).

IR spectroscopy using KBr tablets is the current method for analysis of urinary calculus compositions in our laboratory. The preparation of KBr tablets is time-consuming and often hampered by pellet breakage. To overcome these drawbacks, we developed a new IR method, using a Golden Gate Single Reflection Diamond Attenuated Total Reflection (ATR) device. This method makes use of authentic sample material without any sample pretreatment. The results of the Golden Gate assay were quantified by a program dedicated for the prediction of the outcome of urinary calculus composition analyses. The new method was validated by comparing the results with those obtained with the IR assay with KBr tablets. The quantitative results from this KBr method were estimated from the IR spectra by the use of an initial computerized library search and followed by visual inspection of the spectra.

MATERIALS AND METHODS

Samples

For the construction of a library for the GGN method, we prepared IR spectra of 25, mostly commercially available, components (Table 1) and 236 mixtures. Usually, no more than three components can be detected in one patient sample. The majority of all urinary calculi contain one or more (maximum, three) of the eight most commonly occurring components. Therefore, the 236 mixtures were prepared from these commonly occurring urinary calculus components. The components are as follows: ammonium hydrogen urate (AMUR), brushite (BRUSH), carbonate apatite (CARB), cystine (CYST), stuvite (STRUV), uric acid (URIC), weddellite (WEDD), and whewellite (WHEW; Table 1). The mixtures were restricted to binary and ternary mixture designs. These, so-called constrained and balanced, mixtures were prepared in linear ranges of 0–100% with step sizes of 10%. A more detailed description of the preparation of these mixtures can be found elsewhere (13). AMUR was synthesized according to the following brief instructions: 1.68 g of uric acid was suspended in 500 mL water of 37 °C. A 30 g/L ammonia solution was added drop by drop (2 drops/s) until the solution was completely clear. After the solution cooled, the water layer was aspirated. The crystals were subsequently rinsed with water and diethyl ether and dried at 100 °C. With this synthesis, ~0.84 g AMUR can be obtained, which remains stable for ~6 months.

CARB and WEDD were obtained from patient samples by a selection based on purity. Purity was established by comparison of IR spectra with those in the Hesse atlas (10) and by standard wet chemical analysis. All mixtures were carefully mixed using a pestle and mortar.

For the development of the neural network, the previously mentioned 236 library mixtures and pure samples of the 8 commonly occurring components were used. Two additional mixtures were added to the set, giving a total number of 248 mixtures. The two extra mixtures were added to make the validation set more representative.
Patient samples for comparison of the KBR and GGN methods

One hundred consecutively collected urinary calculus samples from 70 males (median age, 56.5 years; range 5–75 years) and 30 females (median age, 49 years; range 21–74 years) served for testing the predictive performance of the new GGN method. The majority of them (>95%) were derived from patients treated with extracorporeal shock wave lithotripsy. Before analysis, the whole patient sample was carefully ground using a pestle and mortar. The quantitative compositions each sample was also obtained from the KBr method, using computerized library search followed by visual inspection of the spectrum. The samples were considered a representative selection of urinary calculi in our routine practice. The composition of the calculi will be described in greater detail in the Results section.

Table 1. Compounds used for the development of the GGN method

<table>
<thead>
<tr>
<th>Component name</th>
<th>Abbreviation a,b</th>
<th>Source c</th>
</tr>
</thead>
<tbody>
<tr>
<td>2,8-dihydroxy adenine</td>
<td></td>
<td>Sigma, Sigma-Aldrich Chemie (Zwijndrecht)</td>
</tr>
<tr>
<td>Albumin (human)</td>
<td></td>
<td>Patient</td>
</tr>
<tr>
<td>Aammonium hydrogen urate</td>
<td>AMUR</td>
<td>Synthesized</td>
</tr>
<tr>
<td>Amorphous calcium phosphate</td>
<td></td>
<td>Patient</td>
</tr>
<tr>
<td>Bilirubin</td>
<td></td>
<td>Merck, Fisher Scientific (Den Bosch)</td>
</tr>
<tr>
<td>Blood clot</td>
<td></td>
<td>Patient</td>
</tr>
<tr>
<td>Bovine albumin</td>
<td></td>
<td>Sigma, Sigma-Aldrich Chemie (Zwijndrecht)</td>
</tr>
<tr>
<td>Brushite</td>
<td>BRUS</td>
<td>Fluka Chemika, Sigma-Aldrich Chemie (Zwijndrecht)</td>
</tr>
<tr>
<td>Calcite</td>
<td></td>
<td>Fluka Chemika, Sigma-Aldrich Chemie (Zwijndrecht)</td>
</tr>
<tr>
<td>Carbonate apatite</td>
<td>CARB</td>
<td>Patient</td>
</tr>
<tr>
<td>Cholesterol</td>
<td></td>
<td>Merck, Fisher Scientific (Den Bosch)</td>
</tr>
<tr>
<td>Cystine</td>
<td>CYST</td>
<td>Sigma, Sigma-Aldrich Chemie (Zwijndrecht)</td>
</tr>
<tr>
<td>Fatty substance</td>
<td></td>
<td>Patient</td>
</tr>
<tr>
<td>Gypsum</td>
<td></td>
<td>Merck, Fisher Scientific (Den Bosch)</td>
</tr>
<tr>
<td>Hydroxyl apatite</td>
<td></td>
<td>Fluka Chemika, Sigma-Aldrich Chemie (Zwijndrecht)</td>
</tr>
<tr>
<td>Monetite</td>
<td></td>
<td>Merck, Fisher Scientific (Den Bosch)</td>
</tr>
<tr>
<td>Newberyrite</td>
<td></td>
<td>Fluka Chemika, Sigma-Aldrich Chemie (Zwijndrecht)</td>
</tr>
<tr>
<td>Palmitic acid</td>
<td></td>
<td>Fisher Scientific (Den Bosch)</td>
</tr>
<tr>
<td>Quartz</td>
<td></td>
<td>Fluka Chemika, Sigma-Aldrich Chemie (Zwijndrecht)</td>
</tr>
<tr>
<td>Sodium-hydrogen urate</td>
<td>STRUV</td>
<td>Sigma, Sigma-Aldrich Chemie (Zwijndrecht)</td>
</tr>
<tr>
<td>Struvite</td>
<td></td>
<td>Riedel de Haën, Sigma-Aldrich Chemie (Zwijndrecht)</td>
</tr>
<tr>
<td>Uric acid</td>
<td>URIC</td>
<td>Fluka Chemika, Sigma-Aldrich Chemie (Zwijndrecht)</td>
</tr>
<tr>
<td>Weddellite</td>
<td>WEDD</td>
<td>Patient</td>
</tr>
<tr>
<td>Whewellite</td>
<td>WHEW</td>
<td>Fluka Chemika, Sigma-Aldrich Chemie (Zwijndrecht)</td>
</tr>
<tr>
<td>Xantine</td>
<td></td>
<td>Fluka Chemika, Sigma-Aldrich Chemie (Zwijndrecht)</td>
</tr>
</tbody>
</table>

a Components with an abbreviation are used for neural network processing.
b The eight most commonly occurring components
c All source located in the Netherlands
**Analytical Methods**

**Standardization**

Before each series, we validated the FT-IR instrument by measurement of a polystyrene transmission standard. Validation comprised wave number positions and absorbancies of known IR bands. The linearity of IR analyses of the KBr and Golden Gate assays was tested using a dilution series of URIC at concentrations of 0–100% with step sizes of 10%. URIC was diluted with WHEW, and the area of the URIC band at 1120 cm\(^{-1}\) served for establishment of test linearity. The “runs-test” was used for establishment of significant deviations from a straight-line (15). \(P \leq 0.05\) was considered statistically different.

**KBr method**

Pulverized urinary calculus (1.5 mg) was mixed with 180 mg KBr with a pestle and mortar. From this mixture, 100 mg was taken for the preparation of a urinary calculus KBR-tablet at 10\(^9\) Pa pressure under vacuum for 2 minutes. A more extensive description can be found in the Hesse atlas (10). The spectra were scanned in the mid-infrared region from 4000–400 cm\(^{-1}\) at 4 cm\(^{-1}\) wave number intervals in a Bio-Rad FTS 135 spectrometer equipped with a cooled DTGS detector and Win-IR (Ver. 3.04) software (both from Bio-Rad Laboratories Inc., Spectroscopic Division). A 100- mg KBr-tablet was used as a blank for background subtraction. Samples producing weak spectra (absolute difference between absorbance maximum and minimum less than A=0.25) were reanalyzed using tablets with higher sample:KBr ratios.

The quantitative composition of each sample was estimated by comparison of the recorded spectra with KBr reference spectra that were stored in a computer library (LITHOS; Bio-Rad). This library contains data of pure components of urinary calculi and 227 mixtures. Its content was similar to a LITHOS library that is used for x-ray diffraction. Win-IR search (Ver. 1.03; Bio-Rad Laboratories Inc., Sadler Division) served as search engine. This search engine applies the Euclidean distance-matching algorithm to the fingerprint area (2000–400 cm\(^{-1}\)) of the absorbance spectra to obtain a spectral hit list. Additional evaluation and interpolation led to an estimate of the quantitative composition of a sample, since the first hit is not necessarily the correct one and because even large libraries cannot contain full detail. After the library search, the final composition was obtained by visual inspection of spectral band intensities by two experienced technicians blinded to the results of the Golden Gate method.

**General outline of the GGN method**

The so-called Golden Gate is a sample-holding device equipped with a Single Reflection Diamond ATR crystal (Graseby Specac) for measurement of micro samples. The standard ZnSe lens was replaced with a KSR5 lens to enable measurements between 600–250 cm\(^{-1}\). Carefully pulverized material (1-2 mg) was applied to the flat surface of the diamond crystal and pressurized at 3 x 10\(^8\) Pa. The reproducibility of this pressure was guaranteed by using the build-in pressure restraint of the pressure applicator of the Golden Gate device. The active sampling area of the crystal was 1.13 mm\(^2\) (diameter 0.6 mm). The uniformity of the crystal spreading on the sensing area was controlled by viewing through the looking glass of the pressure applicator of the Golden Gate device. The samples were always measured at room temperature. A Bio-Rad FTS 135 spectrometer, equipped with a cooled DTGS detector and Win-IR software, was used for scanning in the mid-infrared region from 4000–400 cm\(^{-1}\) at 4 cm\(^{-1}\) wave number intervals. An empty crystal served for background measurement and blank subtraction. The background spectra were always...
collected before a series of 10 sample spectra. All training, validation, and test samples were measured in a more or less random order over ~6 months. Each spectrum was acquired by coaddition and averaging of 16 interferograms. The Golden Gate crystal was cleaned with water and 960 mL/L alcohol after each measurement.

The NEURANET program (Ver. 3.0; Bio-Rad Laboratories Inc., Spectroscopic Division) was used for quantification of urinary calculi, whose compositions are expressed as mass percentage. The selection of this program was based on earlier studies (13;14). This program contains two supplementary quantification methods and was particularly developed for interpretation of IR spectra of urinary calculi in the range of 4000–400 cm⁻¹. The first method is computerized library search, and the second is based on artificial neural network prediction. Library search can be used for quantification of any composition of a calculus, assuming that the components are available at the library. The neural network may be used for more accurate predictions of compositions of urinary calculi, but is restricted to process the outcomes of a maximum of 10 components simultaneously. Therefore, the neural network can only be used for quantification of the most commonly occurring components of urinary calculi. For calculi composed of these commonly occurring components, both quantification methods should provide almost the same outcome. In this case, the library search serves as a verification method of the ANN because it can depict the unknown spectrum graphically together with a number of the library spectra (stacked, or overlaid). For rarely occurring components, the results obtained with library search must be used. The availability of both quantification methods facilitates the interpretation process. Additionally, the program offers the possibility of adding some simple expert rules to the network-predicted results. These rules may be added to solve problems caused by small but systematic inaccuracies in the network outcome of patient samples. Furthermore, they are used to give an indication that the results from library search should be used in case of rare components unavailable to the network model and round the network outcome of each component to the nearest 5%. The expert rules aim a generalization of the quantitative results of future patient samples. The rules may be defined as simple “Basic” like “IF ... THEN ... ELSE ...” statements.

The neural network engine of the NEURANET program is based on a back-propagation neural network (16). This program contains a three-layer network, consisting of an input layer with a number of nodes (neurons) equal to the number of input variables (absorbances at different wave numbers), a single hidden layer with a variable number of nodes, and an output layer with a number of nodes equal to the number of components (maximum, 10). The input nodes are connected to the output nodes via the nodes of the hidden layer. All nodes of the hidden layer have every possible connection with the input and the output nodes. Each connection carries the signal and an individual weight. The final weights in all connections reflect the knowledge of the underlying spectral patterns. The complex of weights may be interpreted as the regression coefficient in a regression analysis. The multiple inputs of a spectrum are converted to a single concentration of a single component. The final set of weights is found by learning by back-propagation. With this method, the neural network is provided with a set of training spectra (samples) with known concentrations and iterates around a loop in which it predicts for each sample the analyte concentrations and compares these to the known concentrations (forward step). Depending on the differences between the calculated and the known outcome concentrations, the weight values will be readjusted (backward step). This happens for each sample of the training set in turn and is repeated many times over the complete data set. The number of
iterations (epochs) is usually very large. Before training, the starting weights are randomized between –0.1 and 0.1.

The performance of the network is monitored by looking at the root mean square error (RMSE). The RMSE is calculated by first taking the sum of squared differences between the desired and obtained output values of the training set. The square root is taken from the average of the sum of squares, which are averaged by the number of outcomes (maximum of 10) and the number of training samples. NEURANET contains several parameters for data preprocessing (e.g., selection of wave number ranges and scaling) and network design (topology). The neural network parameters have to be tuned by means of an independent validation set. This validation set is a representative set of spectra from samples of known composition and is used for testing the performance of the network but not for training. The training behavior of the network is monitored by looking at a graph depicting the decrease and convergences of the RMSE of both the training and the validation sets against the number of epochs. In worse cases, both RMSE curves will diverge instead of converge. To prevent overfitting of the neural network model, the training process is stopped when the RMSE of the validation set is at its minimum value (early stopping rule). In addition to training and validation (used for tuning), the performance of the network should be tested with a test set. More information about neural network processing can be found elsewhere (17;18).

The NEURANET program enables building of one or more named methods, based on spectroscopic absorbance data. Each method contains a combination of standard information (e.g., description of the components), a spectral library, a trained neural network model, and some expert rules. After selection of a method and of a spectrum from the file list, the program automatically performs a library search, network predictions, and expert rule filtering for prediction of the outcome of a sample with an unknown composition. The whole combination of the analysis with the Golden Gate sampling device and the final NEURANET model is called the GGN method.

**Development of the library and neural network of the GGN method**

The 261 samples, composed of 25 pure components and 236 mixtures, were analyzed with the Golden Gate sampling device, and their spectra were added to the library of the NEURANET program. They were stored at 16 cm\(^{-1}\) resolution intervals of the 4000–400 cm\(^{-1}\) analysis range. A spectral range was defined for searching in the 3700–450 cm\(^{-1}\) range with the correlation-matching algorithm. The spectra of 248 pure components and mixtures of AMUR, BRUSH, CARB, CYST, STRUV, URIC, WEDD, and WHEW were recorded with the Golden Gate device. Of these, 199 were used as a training set for the neural network. This training set served for the construction of a network model that has a mapping (topology) suitable for the analysis these eight components (eight output neurons) in the unknown samples. The remaining 49 spectra were used as a validation set. A spectral range from 1840 to 448 cm\(^{-1}\) with 16 cm\(^{-1}\) resolution intervals was selected for neural network processing, giving rise to 88 input neurons. The network topology parameters for training the neural network were optimized by monitoring the RMSE of the validation set. The RMSE values for both the training and the validation sets were graphically depicted to check potential overfitting of the neural network model. A more extensive description of network processing in relation to urinary calculus analysis can be found elsewhere (14). A few expert rules were added to the network-predicted data for further optimization of the results of unknown samples. These rules were added as a result of small but structural differences in composition found between visual inspection of the Golden Gate spectra by
two technicians and the network predicted-results of the patient samples (test set). The rules we added without any foreknowledge of the results from the KBr method.

*Data processing and statistics*

The results of the KBr and GGN methods were compared using Altman-Bland agreement plots (19) for the 8 commonly occurring components and a combination of these. With these agreement plots, the individual differences (AMUR%[KBr]_n - AMUR%[GGN]_n, CYST%[KBr]_n - CYST%[GGN]_n of sample n, and so forth) of both methods are calculated and plotted against the individual mean results (e.g., mean of AMUR%[KBr]_n and AMUR%[GGN]_n of sample n) of both methods. The bias (mean of the individual differences between the GGN and KBr methods) and the 95% agreement limits (1.96 SD of the differences between both methods) are summarized in Table 3. The bias and 95% agreement level were calculated for the eight components separately and for a combination of them by taking the individual calculated differences of the eight components together. The bias and 95% agreement limits of the combination of the eight components are shown in an agreement plot in Fig.3.

**RESULTS**

A KBr transmission spectrum of a patient sample consisting of 60% BRUS and 40% WEDD is shown in Fig 1A. Fig. 1B shows a transmission spectrum of the same sample obtained with the Golden Gate device. Analysis of the URIC dilution series with the runs-test showed that the IR intensities of both the KBr and GGN assays were linear at 0–100%.

![Figure 1. IR transmission spectra of a urinary calculus containing 60% BRUS and 40% WEDD obtained by the KBr method (A) and the Golden Gate device (B).](image-url)
Compositions of urinary calculi with the KBr method

Two of the 100 analyzed patient samples produced weak spectra. They were removed from the data set because of insufficient sample material for reanalysis. The quantitative compositions of the remaining 98 patient samples, as analyzed with the KBr method, revealed that 92 of them contained at least one of the 8 commonly occurring components. All 92 patient samples were single components or binary or ternary mixtures of one of these 8 components. The majority of them contained calcium oxalate (WHEW and WEDD) and/or CARB. The detection frequency of each of the eight components in the 92 samples was: AMUR, 1.1%; CYST, 1.1%; URIC, 3.3%; STRUV, 4.3%; BRUS, 13%; CARB, 48%; WEDD, 70.7%; and WHEW, 75%. The percentage urinary calculi that contained one, two, or three of these components were 14%, 54%, and 32%, respectively. Six of the 98 urinary calculi contained less frequently occurring components. One consisted of quartz, two of uric acid dihydrate and three of a fatty substance. Two of the latter were highly similar to feces, while the other was similar to palmitic acid.

Development of the GGN method

Neural network

After repeated, batch-automated training of the neural network with different topologies, a final topology was found. Each training session took ~20 min for each topology. The final topology had eight hidden neurons. With this topology the RMSE steadily decreased to a minimum value of 2.3% for the validation-set (Fig. 2). This minimum value was reached after a training of 54000 epochs (cycles). At this number of epochs, the error of the training set was 1.5%.

Intermediate analysis of composition of urinary calculi

The 98 patient samples, which were also analyzed by the KBr method, were analyzed with the Golden Gate device. The composition of each calculus was estimated with library searches and neural network prediction of the intermediate GGN method. Computerized estimation of the composition of a single sample with the GGN method was obtained within ~1 s by means of simultaneous library search and network prediction followed by expert-rule filtering. The compositions of six of the 98 patient samples could not be estimated by the ANN because these samples did not contain any of the eight commonly occurring components available in the network model. Four of these samples could be detected with library searches of the GGN method in the first hit (searched from 3700 to 450 cm⁻¹). As with the KBr method, one was found to contain quartz, two contained feces and one contained a component similar to palmitic acid. The composition of the remaining two, which contained uric acid dihydrate according to the KBr method, could not resolved by library search with the GGN method, since the spectrum of uric acid dihydrate was not available in this library.

Addition of expert rules

As a result of visual inspection of the Golden Gate spectra and the outcome of the neural network predictions of the 92 patient samples (test set), 14 simple expert rules were added to the GGN method (Table 2). After addition of all expert rules to the GGN method, the composition of the 92 patient samples, as analyzed with the Golden Gate device, was reestimated with the final GGN method. The results of the final GGN method were used for comparison with those of the KBr method.
Comparison of the KBr method and the GGN methods

The agreement between the KBr and GGN methods, as obtained from the Altman-Bland plots of the results of 92 patient samples is shown in Table 3. The Altman-Bland plot of all patient results is shown in Fig. 3. The plot compares the results of all eight components obtained from the 92 patient samples, analyzed with the KBr and the GGN methods. The dashed lines express the 95% confidence interval of the differences between both methods. Of the 92 samples, 2 consisting of WHEW + WEDD and 1 consisting of CARB + STRU showed 20% difference between both methods (see Fig. 3). Because each of the three samples were mixed stones composed of two concomitantly occurring components, an increased amount (percentage) of one component relative to the other sample produced an equal decrease of the amount of the other component. Therefore, the Altman-Bland plot shows six data points at 20% difference between the methods. For example, the sample containing CARB and STRU was composed of 60% CARB and 40% STRU measured with the KBr method, whereas it contained 80% CARB and 20% STRU measured with the Golden Gate method. The resulting data points (x,y) of the Altman-Bland plot are (70,20) and (30,20).

DISCUSSION

We describe a new IR method for the analysis of the composition of urinary calculi. This method makes use of a Golden Gate Single Reflection Diamond ATR device and a newly developed library and neural network model for quantification. The outcome was compared with that of an IR method with KBr tablets.

Visual inspection of the IR spectra made clear that KBr spectra (Fig. 1A) have more definite bands than those recorded with the Golden Gate ATR device (Fig. 1B). The Golden
Golden Gate method

Gate ATR device also yields spectra with less absorbance intensities at higher wave numbers when compared to the traditional transmission spectra (Fig. 1A). The underlying cause is different sample radiation penetration depths at different wave numbers. It did not influence the interpretation of the spectra obtained with the Golden Gate assay, probably because of the predominant use of the bands with sufficient spectral definition in the 2000–400 cm$^{-1}$ region (fingerprint area).

Table 2. Expert rules added to the GGN method.

<table>
<thead>
<tr>
<th>Name</th>
<th>Expert rule</th>
</tr>
</thead>
<tbody>
<tr>
<td>cystcheck1</td>
<td>IF Cyst&lt;10 THEN Cyst=0;</td>
</tr>
<tr>
<td>uriccheck</td>
<td>IF [AmUr+ Uric]&lt;7.5 THEN AmUr=0 AND Uric=0 AND Normalize;</td>
</tr>
<tr>
<td>amurcheck1</td>
<td>IF (Uric &lt; 7.5) AND (AmUr &gt; 12.5) THEN AmUr=[Uric + AmUr] AND Uric=0 AND Normalize;</td>
</tr>
<tr>
<td>amurcheck2</td>
<td>IF (AmUr &lt; 7.5)AND (Uric &gt; 12.5) THEN Uric=[Uric + AmUr]AND AmUr=0 AND Normalize;</td>
</tr>
<tr>
<td>whewwed1</td>
<td>IF Wedd&lt;0 THEN WEDD=0 AND Whew=[Wedd+Whew];</td>
</tr>
<tr>
<td>whewwed2</td>
<td>IF Whew&lt;0 THEN Whew=0 AND Wedd=[Wedd+Whew];</td>
</tr>
<tr>
<td>whewwed3</td>
<td>IF ((Wedd&lt;3.5) AND (Whew&lt;3.5)) AND ((Wedd&gt;0) AND (Whew&gt;0)) THEN IF Wedd&gt;Whew THEN Wedd=[Wedd+Whew] AND Whew=0 ELSE ,Whew=[Whew+Wedd] AND Wedd=0 AND Normalize;;</td>
</tr>
<tr>
<td>carbcheck1</td>
<td>IF (Carb&gt;35) AND ((Brus&lt;10) AND (Stuv&lt;10)) THEN Carb=[Brus+Carb+Stuv] AND Brus=0 AND Stuv=0;</td>
</tr>
<tr>
<td>carbcheck2</td>
<td>IF (Carb&gt;35) AND ((Brus&gt;10) AND (Stuv&lt;10)) THEN Carb=[Carb+Stuv] AND Stuv=0 AND Normalize;</td>
</tr>
<tr>
<td>amurcheck3</td>
<td>IF (([\text{AmUr}+Uric])&gt;4) AND (([\text{AmUr}+Uric])&lt;30)) AND (([\text{Brus}+Carb+Stuv])&gt;50) THEN AmUr=0 AND Uric=0;</td>
</tr>
<tr>
<td>amurcheck4</td>
<td>IF (AmUr&gt;25) AND ((Brus+Carb+Stuv)&lt;30) THEN AmUr=[AmUr+Brus+Carb+Stuv] AND Brus=0 AND Carb=0 AND Stuv=0;</td>
</tr>
<tr>
<td>amurcheck5</td>
<td>IF (([\text{AmUr}+Uric+Stuv])&lt;10) AND (Carb&gt;10) THEN Carb=[Carb+AmUr+Uric+Stuv] AND AmUr=0 AND Uric=0 AND Stuv=0;</td>
</tr>
<tr>
<td>amurcheck6</td>
<td>IF (([\text{AmUr}+Uric+Stuv])&lt;10) AND (wedd&gt;10) THEN Wedd=[Wedd+AmUr+Uric+Stuv] AND AmUr=0 AND Uric=0 AND Stuv=0;</td>
</tr>
<tr>
<td>amurcheck7</td>
<td>IF (([\text{AmUr}+Uric+Stuv])&lt;10) AND (Whew&gt;10) THEN Whew=[Whew+AmUr+Uric+Stuv] AND AmUr=0 AND Uric=0 AND Stuv=0;</td>
</tr>
</tbody>
</table>
Table 3. Comparison of the results of the KBr and GGN methods for 92 patient samples.

<table>
<thead>
<tr>
<th></th>
<th>Number</th>
<th>Bias (95% confidence Interval), %</th>
<th>95% Level of agreement, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMUR</td>
<td>92</td>
<td>-0.1% (-0.3 to 0.1)</td>
<td>2.0%</td>
</tr>
<tr>
<td>BRUS</td>
<td>92</td>
<td>-0.8% (-1.5 to -0.2)</td>
<td>6.2%</td>
</tr>
<tr>
<td>CARB</td>
<td>92</td>
<td>0.5% (-0.7 to 1.6)</td>
<td>10.8%</td>
</tr>
<tr>
<td>CYST</td>
<td>92</td>
<td>0.0% (0.0 to 0.0)</td>
<td>0.0%</td>
</tr>
<tr>
<td>STRU</td>
<td>92</td>
<td>0.4% (-0.9 to 0.1)</td>
<td>4.7%</td>
</tr>
<tr>
<td>URIC</td>
<td>92</td>
<td>-0.1% (-0.6 to 0.4)</td>
<td>4.4%</td>
</tr>
<tr>
<td>WEDD</td>
<td>92</td>
<td>3.8% (2.5 to 5.2)</td>
<td>12.8%</td>
</tr>
<tr>
<td>WHEW</td>
<td>92</td>
<td>-2.8% (-4.3 to -1.3)</td>
<td>14.2%</td>
</tr>
<tr>
<td>ALL</td>
<td>736</td>
<td>0.0% (-0.3 to 0.3)</td>
<td>9.0%</td>
</tr>
</tbody>
</table>

*Significant bias

Figure 3. Bland-Altman method agreement plot of the results of eight components in 92 urinary calculi as analyzed with the KBr method and the GGN method. The dashed lines are the 95% confidence interval of the differences between both methods.

Development of the neural network model

From a previous study (14), we found that the compositions of urinary calculi using neural network prediction were similar to the results obtained with PLS regression. In both cases, the urinary calculi were analyzed with IR spectroscopy, using KBr tablets. The PLS regression model and the neural network model were developed for quantification of
mixtures of WHEW, WEDD and CARB, whose incidence rate in urinary stones is ~80% in Western countries (20). The RMSE values of the validation sets were 1.7% for PLS regression and 1.6% for network prediction, respectively. The previous study also described the development of a neural network model trained with the eight most commonly occurring components. This neural network model has been used successfully for several years in our laboratory (UHG). On the basis of these findings, we concluded that for quantification of the composition of urinary calculi, neural networks would be useful as or better than linear models such as PLS regression. We therefore developed a new neural network model with the NEURANET program after replacing the KBr sampling device by the Golden Gate accessory.

Training of the ANN went remarkably well, despite the previously mentioned relatively poor spectral band definition in the fingerprint area compared with the KBr method. Although the NEURANET program contains several facilities to make network training rather simple, there are several difficulties in applying ANN models. The ANN parameters (topology) are often difficult to estimate, and large training sets are often needed. The number of training samples should be at least more than the number of input units (88 in our case), but the required number of training samples also strongly depends on the noise level in the targets and the complexity of the adaptation of the network to the target function. The absorbance values of IR spectra are linearly related to the concentration (Lambert–Beer law). Therefore, a relatively small number of learning samples (n=199) was needed for training. We gave special attention to the risk of overtraining (overfitting). In this case the network looses generalization (robustness) and will adapt (learn) to unimportant spectral features, such as noise. Overfitting (and underfitting) was monitored by looking at the RMSE errors of the validation and training sets (Fig. 2). If the validation error became much higher than training error, the network was probably overfitted and another topology was applied. In addition, the robustness of the neural network model is important and can be tested by retraining the neural network with different initial weights each training session, providing that the other conditions are left unchanged (e.g., topology settings, training and validation sets). If the RMSE values at a certain number of epochs show large differences for the different training sessions, the network model can be considered to be unstable. The final criterion for assessing the network model was the comparison of the results from the independent test set with the results obtained with the reference method.

Several heuristics exist for the choice of the starting values of many of the topology parameters. However, since the optimal parameter settings strongly depend on the nature of the problem and on the chosen representation of the input and output objects, it is not safe to rely exclusively on heuristics. Therefore, an operator must have sufficient knowledge of network training to select the topology parameters (e.g., number of hidden neurons) and must interpret the numerical and graphical network outcome. Using NEURANET, consecutive unattended training with different topologies was possible when a number of topologies were set out in advance. After training, estimation of compositions of patient samples is very fast when the stored ANN method is used.

More detailed information about the theoretical background of ANNs is out of the scope of this report, but can be found elsewhere (17,18).

Development of the expert rules
Urinary calculi are always composed of pure components or binary or ternary mixtures. For those components absent in the sample (e.g., five absent in case of a three-component
calculus), small positive or negative numbers may occur in the network outcome of patient samples on a regular base (Table 4). This is a consequence of network training, which will always predict the outcome of the eight components simultaneously, whether or not they are present in the sample. Because the total outcome of any sample is always 100% (Table 4), a small outcome (percentage) may occur for those components absent in the sample. These inaccuracies may occur because no prediction is perfect, the samples may contain trace amounts of impurities caused by their passage through the urinary tract, and some samples produce rather noisy spectra. In this last case, the network assumes detection of small amounts of a component characterized by a great number of spectral bands (e.g., AMUR). Therefore, a few expert rules were defined (see Table 2). In essence, the expert rules can be considered as automated corrections of the network outcome, which otherwise would have been made manually by expert technicians after visual inspection of the IR spectra. Some of these rules are counterparts of each other, describing almost the same type of correction (e.g., amurcheck1 and amurcheck2). Another rule, named whewwedd3, seems to be rather complex, but only assigns the smallest oxalate outcome (Whew or Wedd) to the largest one, providing that both oxalate outcomes are positive and <3.5%. The rationale is not a physical/chemical one, but only a small correction. If this rule was not applied, both oxalate outcomes would be forced to zero by rounding and normalization, in spite of the fact that a small amount of oxalate is present in the sample.

Table 4. Network- and expert system-predicted outcomes for samples A and B.

<table>
<thead>
<tr>
<th></th>
<th>AMUR</th>
<th>BRUS</th>
<th>CARB</th>
<th>CYST</th>
<th>STRU</th>
<th>URIC</th>
<th>WEDD</th>
<th>WHEW</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Patient A⁴</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Network</td>
<td>-3%</td>
<td>46%</td>
<td>44%</td>
<td>-2%</td>
<td>0%</td>
<td>0%</td>
<td>13%</td>
<td>2%</td>
</tr>
<tr>
<td>Expert</td>
<td>45%</td>
<td>40%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Patient B⁵</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Network</td>
<td>0%</td>
<td>2%</td>
<td>94%</td>
<td>0%</td>
<td>1%</td>
<td>-1%</td>
<td>-5%</td>
<td>9%</td>
</tr>
<tr>
<td>Expert</td>
<td>95%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5%</td>
</tr>
</tbody>
</table>

⁴ Sample A contains BRUS, CARB, and WEDD.
⁵ Sample B contains WHEW and CARB.
⁶ Both samples were analyzed with the Golden Gate assay.
⁷ Library first hit: WEDD-BRUS 10%-90%; second hit: BRUS-CARB 50%-50%.
⁸ Library first hit: WHEW-CARB 10%-90%; second hit: WEDD-CARB 10%-90%.

Except for normalization of the network outcome to the nearest 5%, no special expert rule was applied to patient sample A (Table 4). Sample B shows somewhat inaccurate results of both calcium oxalates (WHEW and WEDD). WHEW turned out to be predominant relative to WEDD. Because this has happened several times, an expert rule was defined by simply adding the values of both calcium oxalates and forcing the value of WEDD to zero. This rule was defined as follows:

\[
\text{IF wedd}<0 \text{ THEN whew=[wedd+whew] AND wedd=0;}
\]
This rule caused 0% WEDD and 4% WHEW. After normalization of sample B the final composition of the expert system was 5% WHEW and 95% CARB (Table 4). This outcome reflected the real composition of this sample, which was based on careful visual inspection of the band intensities of the KBr spectrum by a trained technician.

Method comparison

The 92 consecutive samples used for method comparison were regarded to be a representative selection of urinary calculi in our daily practice. They had similar frequency distribution of components and number of components per sample, compared with historical data (not shown). X-ray diffraction is occasionally recommended as a reference method for urinary calculus analyses. However, x-ray diffraction cannot adequately detect amorphous substances (3). CARB is, for example, sometimes overlooked, but can be detected by a simple CO₂ test following acidification with HCl. Quantitative analysis of CARB may, however, be difficult. We therefore decided to compare the GGN results in both an analytical and a managerial sense with those obtained by an IR method with KBr tablets. This method was routinely used at the time of the study and was to be substituted with a less time-consuming and more robust analytical method.

The bias of the outcome of the KBr and GGN methods of the 92 patients' samples was significantly different from zero for BRUS, WHEW, and WEDD (Table 3). The small bias of BRUS (−0.8%) seems irrelevant. The biases for WHEW (−2.8%) and WEDD (3.8%) are small and carry different signs, probably related to their concomitant occurrence in urinary calculi. The 95% levels of agreement of WHEW, WEDD, and CARB were >10% (Table 3). These components often occur concomitantly in a single sample, causing complex spectral patterns. The 95% level of agreement of all results was 9%. This value should be taken as an indication, since it is statistically not correct to base such calculations on mutually dependent variables (each sample occurs 8 times). Only 3 of 92 patient samples exhibited maximum differences of 20%. These differences occurred consistently in samples that contained two rather similar components (WHEW + WEDD, and CARB + STRU). It is not known what analytical precision and bias are relevant in terms of the prevention of urinary calculus recurrence. We nevertheless consider the encountered differences minor and possibly irrelevant with respect to the ultimate (dietary) advice.

Apart from adequate quantification of the eight commonly occurring components, the library search in the GGN method enabled detection and quantification of rarely occurring components in four samples. This feature may be further developed by the addition of other components to the library, like uric acid dihydrate, in the near future. On the other hand, library search may be used for verification of network results. However, it may sometimes be somewhat difficult to establish an accurate quantitative composition of a sample in this way. This is illustrated by the results from the first and second hits obtained with a library search of patient A in Table 4.

In conclusion, the Golden Gate assay seems superior to the KBr assay because of its smaller sample size, because there is no need for sample pretreatment except for grinding, the turnaround time is shorter, and no time is lost because of KBr tablet breakage (Table 5). The GGN method does, however, require higher initial investment because of the Golden Gate ATR sampling device.
Table 5. Managerial summary of the KBr and Golden Gate assays

<table>
<thead>
<tr>
<th></th>
<th>KBr assay</th>
<th>Golden Gate assay</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample size</td>
<td>1.5 mg</td>
<td>≤ 1 mg</td>
</tr>
<tr>
<td>Sample pretreatment</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>TAT</td>
<td>30 min.</td>
<td>15 min.</td>
</tr>
<tr>
<td>MTBF</td>
<td>1 day (pressing)</td>
<td></td>
</tr>
<tr>
<td>Price (US $)</td>
<td>27600</td>
<td>34800</td>
</tr>
<tr>
<td>Required Knowledge</td>
<td>experienced</td>
<td>experienced</td>
</tr>
</tbody>
</table>

TAT, turnaround time (analysis and interpretation); MTBF, mean time between failures.

Because no sample pretreatment is needed, different brands of FT-IR spectrometers give similar spectra under equal local conditions (e.g., temperature and sample pressure), and the chemical composition of urinary calculi is similar in most developed countries, it would be interesting to investigate whether the GGN method could be transferred to other laboratories, without retraining the neural network with local data. This, however, awaits confirmation. The required expert knowledge for spectral interpretation is minimized by use of the ANN and library, but visual inspection remains necessary.

Acknowledgements
We thank Wil Wennekes (CKCHL) and Theo de Haan (UHG) for excellent technical assistance. We also thank Bio-Rad for technical support.

REFERENCES