High throughput urinary neopterin-to-creatinine ratio monitoring of systemic inflammation

Charlotte M. Stuart[1], Elina Zotova[1], Grielof Koster[1], Aravinthan Varatharaj[1], Grace Richardson[1], Faye R. Cornick[1], Mark Weal[2], Tracey A. Newman[1], Anthony D. Postle[1], Ian Galea[1]

[1]Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton, Southampton, United Kingdom

[2]Electronics and Computer Science, Faculty of Engineering and Physical Sciences, University of Southampton, Southampton, United Kingdom

Corresponding Author:

Dr Ian Galea, Associate Professor in Experimental Neurology, Clinical Neurosciences, Clinical & Experimental Sciences, Faculty of Medicine, University of Southampton, Mailpoint 806, Level D, Southampton General Hospital, Southampton SO16 6YD, UK. E-mail: I.Galea@soton.ac.uk

Keywords: neopterin; urinary neopterin creatinine ratio; inflammation; liquid chromatography; mass spectrometry; urine; multiple sclerosis

Acknowledgements: Akosua Agyemang-Prempeh, Rachel Kidd, Hugh Perry (University of Southampton)

Funding: MS Society

List of abbreviations:

UNCR	Urinary neopterin to creatinine ratio
UPLC-MS	Ultra performance liquid chromatography - mass spectrometry
IFN-γ	Interferon-γ
GTP	Guanosine triphosphate
HPLC	High performance liquid chromatography
TQD	Triple quadrupole
PFP	Pentafluorophenyl
MRM	Multiple-reaction monitoring
LOD	Limit of detection
LOQ	Limit of quantification
SD	Standard deviation
AUC	Area-under-the-curve
SPoC	Sample pooling calculator
HC	Healthy controls
pMS	Progressive multiple sclerosis
TIC	Total ion current
RRMS	Relapsing remitting multiple sclerosis

Abstract

BACKGROUND: Systemic inflammation is a marker of ill health and has prognostic implications in multiple health settings. Urinary neopterin is an excellent candidate as a non-specific marker of systemic inflammation. Expression as urinary neopterin-to-creatinine ratio (UNCR) normalizes for urinary hydration status. Major attractions include: (1) urine *versus* blood sampling, (2) integration of inflammation over a longer period compared to serum sampling, (3) high stability of neopterin and creatinine.

METHODS: A high-throughput ultra performance liquid chromatography - mass spectrometry method was developed to measure neopterin and creatinine together from the same urine sample. The assay was applied in several clinical scenarios: healthy controls, symptomatic infections and multiple sclerosis. Area-under-the-curve was compared between weekly and monthly sampling scenarios. Analysis of a single pooled sample was compared with averaging results from analysis of individual samples.

RESULTS: The assay has excellent intra-assay and inter-assay precision, linearity of dilution and spike-and-recovery. Higher UNCR was demonstrated in females *versus* males, older age, inflammatory disease (multiple sclerosis) and symptomatic infections. In healthy controls, fluctuations in inflammatory state also occurred in the absence of symptomatic infection or other inflammatory triggers. Analysis of a single pooled sample facilitates weekly urine sampling to integrate inflammatory activity over time.

CONCLUSIONS: UNCR is a useful biomarker of systemic inflammation. The method presented offer simplicity, speed, robustness, reproducibility, efficiency and proven utility in clinical scenarios. UNCR fluctuations underline the importance of

longitudinal monitoring, versus a single time point, to capture a more representative estimate of an individual's inflammatory state over time.

Impact statement

Longitudinal monitoring of systemic inflammatory status is important to follow activity of chronic inflammatory disease in an individual over time, or response to treatment. A high-throughput ultra performance liquid chromatography - mass spectrometry method was developed to measure urinary neopterin-to-creatinine ratio (UNCR), with excellent intra-assay and inter-assay precision, linearity of dilution and spike-and-recovery. The assay was applied in several clinical scenarios, including healthy young and aged individuals, symptomatic infections and multiple sclerosis. We demonstrate fluctuations in UNCR which underline the importance of longitudinal monitoring, versus a single time point, to capture a more representative estimate of an individual's inflammatory state.

Introduction

The role of systemic inflammation in determining outcome in a variety of health settings is increasingly recognized. These range from frailty in normal community-dwelling adults *(1)* to common conditions such as hearing loss *(2)*, cardiovascular disease *(3)*, dementia *(4, 5)* , postoperative delirium *(6)*, and cancer *(7)*. Hence, there is a need for a practical and robust way to measure systemic inflammation serially and non-invasively.

The cytokine cascade is central to systemic inflammation and a pivotal cytokine appears to be interferon-γ (IFN-γ), which can cross-talk with other cytokines in a variety of ways *(8)* and act as a master checkpoint regulator for many cytokines *(9)*. While for these reasons IFN-γ is a very good candidate as a marker of systemic inflammation, it has several disadvantages from a technical point of view. It is rapidly degraded, requires blood sampling, and receptor-binding makes its measurement in soluble phase difficult to relate to its biological effect.

Specific aromatic pteridines called neopterins are produced by monocytes and macrophages upon stimulation with interferon-γ (IFN-γ), with reasonable specificity for IFN-γ versus other cytokines *(10)*. Myeloid cells can convert guanosine triphosphate (GTP) to 7,8-dihydroneopterin triphosphate, but the conversion of the latter to tetrahydrobiopterin is inefficient, due to a high ratio of GTP cyclohydrolase I to 6-pyruvol-tetrahydropterin synthetase, which occurs after IFN-γ stimulation. This results in production of a high amount of 7,8-dihydroneopterin after IFN-γ stimulation which is not seen in any other cell type examined, conferring cellular specificity *(11)*. Most 7,8-dihydroneopterin is oxidized to D-erythro-neopterin (2-amino-4-hydroxy-6-(D-erythro-1´, 2´, 3´- trihydroxypropyl)-pteridine) *(12)*; a variable small amount undergoes epimerization to L-threo-neopterin (L-threo-6-(l',2',3'-trihydroxypropyl)

pterin, **Supplemental Figure 1**) *(13)*. Both neopterin isomers (D-erythro-neopterin and L-threo-neopterin), collectively referred to as neopterin from here on, are excreted into urine. When L-threo-neopterin is detectable in urine, the ratio of D-erythro-neopterin is roughly 1:10, but this ratio varies *(14)*. Being stereoisomers, D-erythro-neopterin and L-threo-neopterin peaks are difficult to completely resolve *(15)*.

Neopterin presents several advantages over IFN-γ as an inflammatory biomarker, since it is more chemically stable, does not bind to receptors, and undergoes rapid renal clearance, meaning urine concentrations closely reflect neopterin biosynthesis. Urinalysis is particularly attractive since: it is non-invasive, integrates inflammatory activity over a longer period of time compared to a blood sample, and as a less protein-rich solvent, is less likely to suffer from matrix effects. Neopterin is expressed as urinary neopterin to creatinine ratio (UNCR, measured in µmol/mol), to normalize for urinary hydration status.

UNCR is measurable at a basal level and increases non-specifically in response to a number of immunological stimuli such as bacterial, viral and parasitic infections *(16)* as well as cancer and autoimmune diseases *(17)*. It is an excellent candidate for two applications: (1) to follow activity of an existing chronic inflammatory disease in an individual over time, or in response to treatment; (2) an integrator of systemic inflammation over a period of time.

Techniques used for the analysis of pteridines include enzyme-linked immunosorbent assay (ELISA) *(18)* and capillary electrophoresis−laser induced fluorescence *(19, 20)*. It would be advantageous, however, if creatinine and neopterin are measured in the same run on the same platform, rather than using different kits at different times, to minimize variability. Methods that overcome this

include high performance liquid chromatography (HPLC) *(21)* and ion-pair HPLC *(22)* using fluorometric detection for neopterin and UV detection for creatinine; and HPLC-mass spectrometry *(23)*. Ultra performance liquid chromatography – mass spectrometry (UPLC-MS) presents an appealing alternative to these other methods due to its high throughput and sensitivity. All published MS-based neopterin measurement methods use D-erythro-neopterin as standard, but none have clarified whether L-threo-neopterin is also measured along with D-erythro-neopterin. Hence, we set out to develop and validate a robust high-throughput assay using UPLC-MS/MS, and apply it to several clinical settings as exemplars.

Methods

UPLC-MS/MS optimization

An ACQUITY UPLC was interfaced with a Waters Xevo triple quadrupole (TQD) mass spectrometer equipped with an electrospray ionization probe, column oven and autosampler (Waters, Elstree, Herts, UK). Reagents were purchased as described in Supplemental Methods. During sample preparation and analysis, urine samples were kept in the dark at 5°C. UPLC and MS/MS were optimized as described in Supplemental Methods.

In the final optimized protocol, separation was achieved using a UPLC penta-fluoro-phenyl (PFP, 1.7 µm, 2.1 × 100 mm, Waters, Elstree Herts, UK) column kept at 24°C and fitted with its equivalent guard column (1.7 µm, 2.1 × 5mm, VanGuard, Waters, Elstree Herts, UK). The primary mobile phase (A) was 0.2% formic acid and the co-solvent (B) consisted of 0.2% formic acid in acetonitrile. Total run time was eight minutes and included a 2.5 minute gradient from 99% (A) to 93.5% (A) followed by a 0.5 minute gradient to 100% (B) which was kept for 2 minutes before a 0.5 minute gradient back to starting conditions (99% A) for 2.5 minutes, as shown in

Supplemental Table 1. Injection volume was set at 10µl for neopterin analysis and 5µl for creatinine analysis. Using this method, creatinine showed very little retention, eluting just after the solvent front at 0.9 minutes, and D-erythro-neopterin eluted at 1.8 minutes (**Figure 1A**). Similar retention times were seen in urine (**Figure 1B**).

After separation, the compounds were monitored in scheduled multiple-reaction monitoring (MRM) mode with positive electrospray ionization. Injection parameters consisted of cone energy = 30V and collision energy = 20V. The MRM transitions were 114.1 > 44.1 m/z for creatinine, 117.1 > 47.1 m/z for creatinine-D_3, 254.1 > 206.1 m/z for D-erythro-neopterin and L-threo-neopterin and 259.1 > 210.1 m/z for D-erythro-neopterin-$^{13}C_5$. Concentration of analyte was calculated by integrating the area under the peaks using MassLynx software (v4.1, Waters, UK).

Calibration curve preparation

Stock standard solutions of creatinine and creatinine-D_3 were made in milli-Q water. D-erythro-neopterin and D-erythro-neopterin-$^{13}C_5$ solutions were made in 33 mM ammonium hydroxide. Aliquots were stored at -80°C.

For the calibration curve, creatinine (5 mg/dL) and D-erythro-neopterin (100 µg/L) stock solutions were serially diluted in starting conditions, i.e 99% mobile phase (A) and 1% mobile phase (B). Solutions of creatinine-D_3 (2.5 µg) and D-erythro-neopterin-$^{13}C_5$ (50 ng) were added to make up the standard solutions to 1ml. Standards were measured in duplicate at the beginning of each run and every subsequent 24 hours. The calibration curves were constructed by plotting average ratio of the peak areas (creatinine: creatinine-D_3 or D-erythro-neopterin: D-erythro-neopterin-$^{13}C_5$) against concentration of analyte. A typical calibration curve for D-erythro-neopterin and creatinine is shown in **Figure 1C** and **D**.

Urine samples and analysis

Midstream urine was collected with informed consent from adults (National Research Ethics Approvals 12 SC 0176 and 13/SC/0507, and institutional research ethics approvals ERGO 5562 and ERGO 7923). Participants were given kits and trained to shield urine from light. Urine was stored at -20˚C, then thawed and centrifuged at 10°C, 2000 rcf for five minutes. Aliquots were stored at -80°C.

For analysis, individual urine aliquots were diluted 1:100 for neopterin and 1:500 for creatinine. Isotopically-labelled internal standard was added to a total of 1ml solution. Samples were assayed in duplicate or triplicate. Urinary creatinine and neopterin concentrations were determined based on the calibration curve, expressed as µg/L and mg/dL respectively. UNCR was expressed in µmol/mol. A reference urine sample pooled from samples taken from ten different individuals was used during each experiment to assess inter-assay precision.

Method validation

Primary validation was employed, since this is more robust than comparative validation. Validation included linearity, spike-and-recovery, precision, limit of detection (LOD) and limit of quantification (LOQ). Linearity was evaluated by plotting actual *versus* theoretical concentration of matrix-matched calibration points over a clinically relevant range *(23)*.

Spike-and-recovery experiments were used to monitor matrix effects in healthy and diseased states *(24)*. Urine samples from both healthy controls and progressive multiple sclerosis patients were spiked, in triplicate, with three concentrations of D-erythro-neopterin (100-5000 µg/L) or three concentrations of creatinine (5.0–300.0

mg/dL). Recovery was calculated using the following equation, using spike as denominator as per guidelines *(26)* :

$$\% \text{ Recovery} = \frac{\text{measured concentration with spike} - \text{measured concentration without spike}}{\text{Spike concentration}} \times 100$$

Method precision was assessed by inter- and intra-assay coefficients of variation (CV). Intra-assay variation was established by replicate analyses (n=10) of a pooled urine sample. Inter-assay variation was established by replicate analysis of the same pooled sample measured on separate days (n=5).

Determination of LOD and LOQ is only necessary for analytes with concentrations close to zero *(25)*, therefore only the LOD and LOQ for neopterin was studied as creatinine is always present in urine at high concentrations. LOD and LOQ were calculated according to Eurachem guidelines (2014) *(26)* in which the standard deviation (SD) of replicate measurements of samples with low levels of analyte was calculated and multiplied by 3 for LOD or 10 for LOQ (see Supplemental Methods for details).

Statistics

Analyses were performed in SPSS Statistics v25 (IBM, Armonk, USA) or Prism v7 (GraphPad Software, San Diego, USA). Descriptive statistics are shown as mean ± SD. Non-parametric data were log-transformed. To compare groups where age was a scalar covariate, ANCOVA was used, otherwise two-way ANOVA or Student's t-tests were used. A p-value of <0.05 was considered statistically significant.

Results

Method validation

Matrix effects were tested in spike-and-recovery experiments. Ion suppression, often caused by matrix effects, can be a significant issue when using electrospray ionisation *(27)*. In order to assess a varied range of matrices, urine samples from both healthy controls and progressive multiple sclerosis patients were used. The influence of matrix effects is generally more pronounced at lower retention times due to the co-elution of polar compounds commonly found in urine. Consequently, it was expected that creatinine, which elutes just after the solvent front, may show signs of ion suppression. Indeed ion suppression accompanied by poor spike-and-recovery was seen in up to a third of samples when using a 1:100 dilution. This phenomenon was abolished at a 1:500 dilution. As a result, for all further analysis, two sets of dilutions were used during the same run, namely 1:100 for neopterin and 1:500 for creatinine. The average percentage recoveries for D-erythro-neopterin and creatinine in healthy control and progressive multiple sclerosis urine samples are shown in **Table 1**.

Both D-erythro-neopterin and creatinine exhibited excellent linearity ($R^2 > 0.99$) over a range of 0.5-140 µg/L for D-erythro-neopterin and 0.001-0.88 mg/dL for creatinine, corresponding to urinary concentrations of 50-14000 µg/L and 0.5-450 mg/dL respectively (see **Supplemental Figure 2**). These ranges were selected as they span the ranges encountered by us and others *(23)*. Intra- and inter-assay CVs (**Table 2**) indicate that the method is reliable and precise. The calculated values for the LOD and LOQ of D-erythro-neopterin were 0.3 µg/L and 1µg/L respectively.

L-threo-neopterin

The D-erythro-neopterin peak exhibited a small shoulder in some urine samples (**Figure 1B**). We confirmed this as L-threo-neopterin (**Figure 1E and F**). A spike-and-recovery experiment in urine was performed to show that total neopterin (D-erythro-neopterin + L-threo-neopterin) was reliably measured within the same peak, using the D-erythro-neopterin standard curve, in a more biologically relevant matrix. The spike (100 µg/L) was made up of 10% L-threo-neopterin and 90% D-erythro-neopterin so as to recapitulate a realistic ratio when L-threo-neopterin was present. Recovery was 106.48 ± 7.26 for healthy control urine samples (n=4), 99.48 ± 13.64 for progressive multiple sclerosis urine samples (n=4), and 102.98 ± 11.48 for all urine samples combined (n=8).

Clinical application

We assessed the use of UNCR in several clinical scenarios. Our aim was not to address the hypothesis that UNCR rises with inflammation since this has been tested before, but to demonstrate the usability of our assay. Two different sampling techniques were tested:

1. Single time-point urine samples in which UNCR was measured once (spot UNCR).
2. Urine samples taken serially over an extended time period to determine UNCR area-under-the-curve (AUC).

<u>Spot UNCR</u>

UNCR is reported to be higher with increasing age and in females compared to males *(28, 29)*. To confirm this using our assay, spot urine samples were collected from older (n=59, average age = 69.4 ± 3.2, male/female ratio = 17/42) and younger

(n=22, average age = 25.1 ± 5.8 years, male/female ratio = 11/11) healthy control individuals. UNCR was higher in the older age group and in females (**Figure 2A**). In a two-way analysis of variance of UNCR, both age and sex significantly influenced UNCR (age: $F_{1,78}$=54.69, p<0.0001, partial eta squared = 0.41; sex: $F_{1,78}$=12.97, p=0.002, partial eta squared = 0.14).

To confirm that UNCR as measured with our assay is sensitive to systemic inflammation, we compared spot urine samples from patients with relapsing remitting multiple sclerosis (n=11) and healthy controls (n=12) (Figure **2B** see **Supplemental Table 2** for demographics). In an analysis of covariance of UNCR controlling for sex and age, multiple sclerosis patients had significantly higher UNCR compared to healthy controls ($F_{1,19}$=24.45, p<0.0001, partial Eta squared = 0.563).

We then applied our method to study longitudinal fluctuations in UNCR, in relation to symptoms suggestive of an infectious or inflammatory event, in healthy individuals (n=11, average age = 21.5 years, sex ratio, M/F = 7/4) who collected weekly urine samples during an eight week period in autumn. Participants kept a daily record of any symptoms and medical events, such as infections, dental procedures, vaccinations, surgery or injuries. Periods of raised UNCR (UNCR peaks) were identified using automated peak analysis in GraphPad Prism, with each individual's mean UNCR as baseline. Peak identification was also conducted visually, blinded to symptomatic events. Once peaks were identified, the baseline was adjusted to exclude peak UNCR values. When peaks were defined as a rise of 30% above the individual's baseline UNCR, there was 100% concordance between visual and automated methods of peak identification. A UNCR peak was considered to be associated with symptoms of an infection if it occurred within a week of reported symptoms. Six volunteers developed a total of nine events with symptoms

suggestive of infection which were associated with a UNCR peak (**Figure 2C, Figure 3A**, t (6) =3.542, p=0.0061, one-tailed paired t test). In two individuals, there were two symptomatic episodes close to each other, which were associated with the same broad peak. Hence, there were seven UNCR peaks associated with symptomatic episodes. A toe nail removal and a laceration in the leg were not associated with a UNCR peak. Additionally, ten UNCR peaks in six individuals occurred in the absence of symptoms (**Figure 3B**, t (9) =6.948, p< 0.0001, one-tailed paired t test) as exemplified in **Figure 2D**. One of these individuals had received an influenza vaccine three days before the recorded UNCR peak. There was no difference in the % UNCR rise above baseline between symptomatic and asymptomatic peaks (t (9.106) =1.516, p=0.1633, two tailed t test with Welch's correction, **Supplemental Figure 3**). Only two individuals in this cohort had troughs, defined in the same way as peaks, but below the baseline. One was not associated with any symptom or event. The other was associated with the influenza vaccination, and followed the peak.

AUC UNCR

Serial sampling and calculation of the UNCR AUC normalized for follow-up time might provide a more accurate estimate of an individual's overall inflammatory status over a period of time, since it is likely to be less affected by day-to-day fluctuations in UNCR. To investigate whether frequency of sampling makes a difference, we compared the UNCR AUC from monthly urine samples *versus* UNCR AUC from weekly urine samples in 20 progressive multiple sclerosis patients (demographics in **Supplemental Table 3**) who underwent weekly urine sampling over a 2.6 year period. The monthly samples were selected using a random number generator. A paired t-test showed a significant difference between the UNCR AUC calculated from weekly and monthly values (t (19) =2.805, p=0.0113). Although this may indicate that

weekly urine sampling delivers more information than monthly sampling, the increased sampling frequency has cost implications. We therefore investigated whether analysis of a single sample derived from pooling weekly samples could replace AUC calculated from individual analysis of weekly samples, to reduce running costs. Sample sets (n=20) consisting of weekly urine samples taken over an average of 2.6 years were: (i) measured individually and (ii) pooled into one sample, which was then measured as a single sample. In clinical practice the time intervals between serial samples may not be exactly the same, and there may be missing samples. Hence the volume of serial samples cannot be pooled in equal volumes. To help with the preparation of the pooled sample, we developed a Sample Pooling Calculator (SPoC) (available on https://eprints.soton.ac.uk/431818/) which calculates the volume of each sample to add to the pooled sample, based on the time interval between consecutive samples (see SPoC manual at https://eprints.soton.ac.uk/431818/ for methodology). When samples were prepared at 1:100 dilution, results from the two methods to measure UNCR AUC correlated well (R^2 = 0.903), but less well when a 1:500 dilution was used (R^2 = 0.669) (**Figure 4**).

Discussion

In chronic inflammatory diseases, systemic infections may increase the severity of symptoms or contribute to the overall progression of the disease *(30)*. Hence tracking systemic inflammation is important for research studies in this area. It also has significant potential utility in clinical management, when assessing response to treatment in inflammatory or infectious diseases. Urine sampling is non-invasive so UNCR measurement provides a practical and simple means for clinical practitioners and researchers to track inflammation or response to treatment in inflammatory

disorders. Patient cooperation is required to collect urine samples on certain days of the week using simple consumables supplied by the laboratory, date and/or time them and store them in their home freezer. In our experience, even disabled patients with progressive multiple sclerosis find this to be easy and acceptable.

Here we have demonstrated a robust and sensitive analytic technique for systemic inflammation monitoring using UPLC-MS/MS, taking advantage of the neopterin biochemical pathway in activated myeloid cells. Similar methods have been previously reported *(23, 31, 32)*. The method described here requires minimal sample preparation, enables high sample throughput, is cost-effective, possesses a low limit of quantification, and has excellent intra-assay and inter-assay coefficient of variation, linearity of dilution and spike-and-recovery while including both neopterin stereoisomers. Although two dilutions are used, neopterin and creatinine are measured in the same run. The LC-MS method is advantageous over a dual kit/platform approach since neopterin and creatinine are measured in the same run, using the same machine, in the same laboratory and by the same technician. This avoids common errors associated with performing assays on different days (batch effects) or in different in labs using different techniques. In all the clinical scenarios we tested, the changes in UNCR observed were well above the assay's LOQ and the variations observed in experiments assessing assay precision and recovery after spiking.

L-threo-neopterin is a stereoisomer of D-erythro-neopterin which has so far received little attention. L-threo-neopterin is not always present, and the factors determining its production alongside D-erythro-neopterin need further study. In humans, L-threo-neopterin is synthesized by monocytes and macrophages in response to IFN-γ via the same biosynthetic pathway as D-erythro-neopterin *(33)*. Hence it is important to

be able to include both isomers in peak area measurements, since they are measured as a surrogate marker for the process (inflammation) that has generated them. We have shown in spike-and recovery-experiments that when L-threo-neopterin is present, it can be measured reliably using the D-erythro-neopterin standard curve.

Although measurement of 7,8-dihydroneopterin in urine, in addition to neopterin, is useful *(36)* it is very unstable and difficult to reliably measure, making the quantification of downstream neopterin more practical in clinical practice or research studies embedded within busy clinical services. Indeed, it has been shown that measurement of neopterin has similar potential to the combined measurement of neopterin and 7,8-dihydroneopterin in clinical practice *(37)*. The ratio of 7,8-dihydroneopterin to neopterin in bodily fluids is stable (3:1) *(38)*.

In healthy controls, the absolute UNCR values using our assay are in agreement with those found in the literature. We also observed higher UNCR in the majority of older individuals and in females *versus* males, as described before *(28, 29)*. Hence UNCR needs age and sex-specific reference ranges. As we, and others, accumulate more UNCR data over the next few years, these ranges will become available.

To demonstrate the clinical utility of the assay as a marker of systemic inflammation, we applied it to two settings. First, we found a higher UNCR in multiple sclerosis patients compared to healthy controls, as expected in an inflammatory condition *(14, 39, 40)*. Second, weekly urine sampling in healthy controls identified UNCR peaks in association with periods of symptoms suggestive of infection. An interesting observation was that UNCR peaks also occurred in the absence of such symptoms. One such peak was associated with influenza vaccination, which suggests that

asymptomatic UNCR peaks may represent episodes of immune activation caused by, for instance, subclinical infections or stochastic fluctuations in immune responses. This sub-study has several limitations; the infectious nature of the symptoms was not established serologically, and the characteristics of the subjects (young and healthy) precludes generalizability. Further study is required to dissect the precise causes of asymptomatic UNCR peaks.

Healthy individuals had different baselines (coefficient of variation of 32.6%). This underlines the importance of establishing a baseline for the individual when interpreting serial UNCR measurements. Hence if one is asking whether in a particular individual, there has been a UNCR rise in a longitudinal study, during which UNCR is measured serially over time, establishing a baseline for the individual would be necessary, due to variability in the individual's baseline. How often should one sample? We have demonstrated how weekly and monthly urine sampling give different results; more frequent sampling is intuitively closer to the ground truth. However a higher frequency of sampling is costly and time consuming. We have shown that analysis of a single sample pooled from serial samples may suffice, provided that care is taken to ensure that the volume of serial sample added is proportional to the time interval between that serial sample and the next; we provide software for automated calculation of these volumes requiring minimum user input. Pooling of urine is done by the researcher, not by the patient, if analysis of a single sample is selected to be the method of choice to integrate the longitudinal UNCR profile over time. The pooling approach has limitations. High dilution during pooled sample preparation was less accurate probably due to the increased risk of pipetting errors. Also, pooling does not allow kinetic correlation between clinical events and UNCR.

In conclusion, UNCR is a useful biomarker of systemic inflammation. The UNCR assay method presented here has been comprehensively assessed, and offers simplicity, speed, robustness, reproducibility, and proven utility in clinical scenarios. UNCR fluctuations underline the importance of longitudinal monitoring, versus a single time point, to capture a more representative estimate of an individual's inflammatory state.

References

1. Gale CR, Baylis D, Cooper C, Sayer AA. Inflammatory markers and incident frailty in men and women: The english longitudinal study of ageing. AGE 2013;35:2493-501.
2. Verschuur C, Agyemang-Prempeh A, Newman TA. Inflammation is associated with a worsening of presbycusis: Evidence from the mrc national study of hearing. International Journal of Audiology 2014;53:469-75.
3. Christodoulidis G, Vittorio TJ, Fudim M, Lerakis S, Kosmas CE. Inflammation in coronary artery disease. Cardiology in Review 2014;22:279-88.
4. Leung R, Proitsi P, Simmons A, Lunnon K, Güntert A, Kronenberg D, et al. Inflammatory proteins in plasma are associated with severity of alzheimer's disease. PLOS ONE 2013;8:e64971.
5. Tsui A, Richards M, Davis D. Systemic inflammation and modifiable risk factors for cognitive impairment in older persons: Findings from a british birth cohort. AGING MEDICINE 2018;1:243-8.
6. Hall RJ, Watne LO, Idland A-V, Raeder J, Frihagen F, MacLullich AMJ, et al. Cerebrospinal fluid levels of neopterin are elevated in delirium after hip fracture. Journal of Neuroinflammation 2016;13:170.
7. Volgger BM, Windbichler GH, Zeimet AG, Graf AH, Bogner G, Angleitner-Boubenizek L, et al. Long-term significance of urinary neopterin in ovarian cancer: A study by the austrian association for gynecologic oncology (ago). Annals of Oncology 2016;27:1740-6.
8. Bezbradica JS, Medzhitov R. Integration of cytokine and heterologous receptor signaling pathways. Nature Immunology 2009;10:333.
9. Zha Z, Bucher F, Nejatfard A, Zheng T, Zhang H, Yea K, Lerner RA. Interferon-γ is a master checkpoint regulator of cytokine-induced differentiation. Proceedings of the National Academy of Sciences 2017;114:E6867-E74.
10. Troppmair J, Nachbaur K, Herold M, Aulitzky W, Tilg H, Gastl G, et al. In-vitro and in-vivo studies on the induction of neopterin biosynthesis by cytokines, alloantigens and lipopolysaccharide (lps). Clin Exp Immunol 1988;74:392-7.
11. Werner ER, Werner-Felmayer G, Fuchs D, Hausen A, Reibnegger G, Yim JJ, et al. Tetrahydrobiopterin biosynthetic activities in human macrophages, fibroblasts, thp-1, and t 24 cells. Gtp-cyclohydrolase i is stimulated by interferon-gamma, and 6-pyruvoyl tetrahydropterin synthase and sepiapterin reductase are constitutively present. J Biol Chem 1990;265:3189-92.
12. Widner B, Mayr C, Wirleitner B, Fuchs D. Oxidation of 7,8-dihydroneopterin by hypochlorous acid yields neopterin. Biochemical and biophysical research communications 2000;275:307-11.
13. Eto I, Bandy MD, Butterworth CE, Jr. Plasma and urinary levels of biopterin, neopterin, and related pterins and plasma levels of folate in infantile autism. Journal of autism and developmental disorders 1992;22:295-308.
14. Wachter H, Fuchs D, Hausen A, Reibnegger G, Werner ER. Neopterin as marker for activation of cellular immunity: Immunologic basis and clinical application. Adv Clin Chem 1989;27:81-141.
15. Burton C, Shi H, Ma Y. Development of a high-performance liquid chromatography - tandem mass spectrometry urinary pterinomics workflow. Anal Chim Acta 2016;927:72-81.
16. Eisenhut M. Neopterin in diagnosis and monitoring of infectious diseases. Journal of Biomarkers 2013;2013:10.
17. Berdowska A, Zwirska-Korczala K. Neopterin measurement in clinical diagnosis. Journal of Clinical Pharmacy and Therapeutics 2001;26:319-29.
18. Inancli SS, Caner S, Balkan F, Tam AA, Guler G, Ersoy R, Cakir B. Urinary neopterin levels in patients with thyroid cancer. Indian journal of otolaryngology and head and neck surgery : official publication of the Association of Otolaryngologists of India 2014;66:302-8.

19. Gibbons SE, Stayton I, Ma Y. Optimization of urinary pteridine analysis conditions by ce-lif for clinical use in early cancer detection. Electrophoresis 2009;30:3591-7.
20. Han F, Huynh BH, Shi H, Lin B, Ma Y. Pteridine analysis in urine by capillary electrophoresis using laser-induced fluorescence detection. Analytical chemistry 1999;71:1265-9.
21. Werner ER, Bichler A, Daxenbichler G, Fuchs D, Fuith LC, Hausen A, et al. Determination of neopterin in serum and urine. Clinical Chemistry 1987;33:62-6.
22. Palmisano F, Rotunno T, La Sorsa M, Zambonin CG, Abbate I. Simultaneous determination of pseudouridine, neopterine and creatinine in urine by ion-pair high-performance liquid chromatography with in-series ultraviolet and fluorescence detection. The Analyst 1995;120:2185-9.
23. Burton C, Shi H, Ma Y. Simultaneous detection of six urinary pteridines and creatinine by high-performance liquid chromatography-tandem mass spectrometry for clinical breast cancer detection. Analytical chemistry 2013;85:11137-45.
24. Matuszewski BK, Constanzer ML, Chavez-Eng CM. Strategies for the assessment of matrix effect in quantitative bioanalytical methods based on hplc-ms/ms. Analytical chemistry 2003;75:3019-30.
25. NATA. Guidelines for the validation and verification of quantitative and qualitative test methods. https://www.nata.com.au/ (Accessed 03 Dec 2017).
26. Magnusson B, Örnemark U. Eurachem guide: The fitness for purpose of analytical methods – a laboratory guide to method validation and related topics. https://eurachem.org/ (Accessed 03 December 2017).
27. Annesley TM. Ion suppression in mass spectrometry. Clinical Chemistry 2003;49:1041-4.
28. Spencer ME, Jain A, Matteini A, Beamer BA, Wang N-Y, Leng SX, et al. Serum levels of the immune activation marker neopterin change with age and gender and are modified by race, bmi, and percentage of body fat. The journals of gerontology Series A, Biological sciences and medical sciences 2010;65:858-65.
29. Murr C, Hainz U, Asch E, Berger P, Jenewein B, Saurwein-Teissl M, et al. Increased neopterin concentration in older age coincides with decline of cd28+cd45ra+ t-cells. Pteridines, Vol. 15, 2004. p. 170.
30. Perry VH, Cunningham C, Holmes C. Systemic infections and inflammation affect chronic neurodegeneration. Nature reviews Immunology 2007;7:161-7.
31. Allegri G, Costa Netto HJ, Ferreira Gomes LN, Costa de Oliveira ML, Scalco FB, de Aquino Neto FR. Determination of six pterins in urine by lc-ms/ms. Bioanalysis 2012;4:1739-46.
32. Shao Z, Zhang R, Shrestha K, Borowski AG, Schuster A, Thakur A, et al. Usefulness of elevated urine neopterin levels in assessing cardiac dysfunction and exercise ventilation inefficiency in patients with chronic systolic heart failure. Am J Cardiol 2014;113:1839-43.
33. Werner-Felmayer G, Werner ER, Fuchs D, Hausen A, Reibnegger G, Wachter H. Neopterin formation and tryptophan degradation by a human myelomonocytic cell line (thp-1) upon cytokine treatment. Cancer research 1990;50:2863-7.
34. Blau N, Steinerstauch, P., Redweik U., Pfister K., Schoedon G., Kierat L., and Niederwieser A. Unconjugated pterins and related biogenic amines. Berlin, Germany: Walter de Gruyter & Co; 1987.
35. Hanneken L, Cremer-Bartels G, Gerding H, Krause K. Retinal monapterin and biopterin of different rat strains. In: Brüne B, editor Eighth Winter Workshop on Pteridines, Vol. 370: Biol. Chem. Hoppe-Seyler, 1989. p. 384.
36. Gieseg SP, Baxter-Parker G, Lindsay A. Neopterin, inflammation, and oxidative stress: What could we be missing? Antioxidants (Basel, Switzerland) 2018;7.
37. Fuchs D, Milstien S, Kramer A, Reibnegger G, Werner ER, Goedert JJ, et al. Urinary neopterin concentrations vs total neopterins for clinical utility. Clin Chem 1989;35:2305-7.
38. Murr C, Widner B, Wirleitner B, Fuchs D. Neopterin as a marker for immune system activation. Current drug metabolism 2002;3:175-87.

39. Yuksekol I, Ozkan M, Akgul O, Tozkoparan E, Al-Rashed M, Balkan A, et al. Urinary neopterin measurement as a non-invasive diagnostic method in pulmonary tuberculosis. The International Journal of Tuberculosis and Lung Disease 2003;7:771-6.
40. Helmut Wachter DF, Arno Hausen, Gilbert Reibnegger and Ernst R. Werner. Neopterin as marker for activation of cellular immunity: Immunologic basis and clinical application. Advances in clinical chemistry 1989;21:81-125.

Tables

Table 1. Average spike-and-recovery experiments for neopterin and creatinine in urine samples from healthy controls (HC) and progressive multiple sclerosis (pMS) patients. Results are shown for HC, pMS and both groups combined.

Neopterin Recovery (%)						
Sample type	**Low 100 µg/L**		**Med 1000 µg/L**		**High 5000 µg/L**	
	% Recovery	**±SD**	**% Recovery**	**±SD**	**% Recovery**	**±SD**
HC (n=4)	102.4	5.5	104.7	2.8	101.4	2.0
pMS (n=4)	99.2	11.5	101.9	2.0	100.9	0.5
All (n=8)	100.8	9.2	103.3	2.8	101.1	1.5
Creatinine Recovery (%)						
Sample type	**Low 5 mg/dL**		**Med 50 mg/dL**		**High 300 mg/dL**	
	% Recovery	**±SD**	**% Recovery**	**±SD**	**% Recovery**	**±SD**
HC (n=3)	97.3	4.3	96.2	1.5	109	4.6
pMS (n=3)	99.7	12.2	103.1	3.6	101.7	5.5
All (n=6)	98.5	9.2	99.7	4.4	105.4	6.3

Table 2. Measurement precision (percentage coefficient of variation, CV). UNCR = urinary neopterin to creatinine ratio.

Precision		
Compound	**Inter-day %CV (n=5)**	**Intra-day %CV (n=11)**
Neopterin (µg/L)	6.34	2.77
Creatinine (mg/dL)	1.32	1.15
UNCR (µmol/mol)	5.55	1.97

Figures

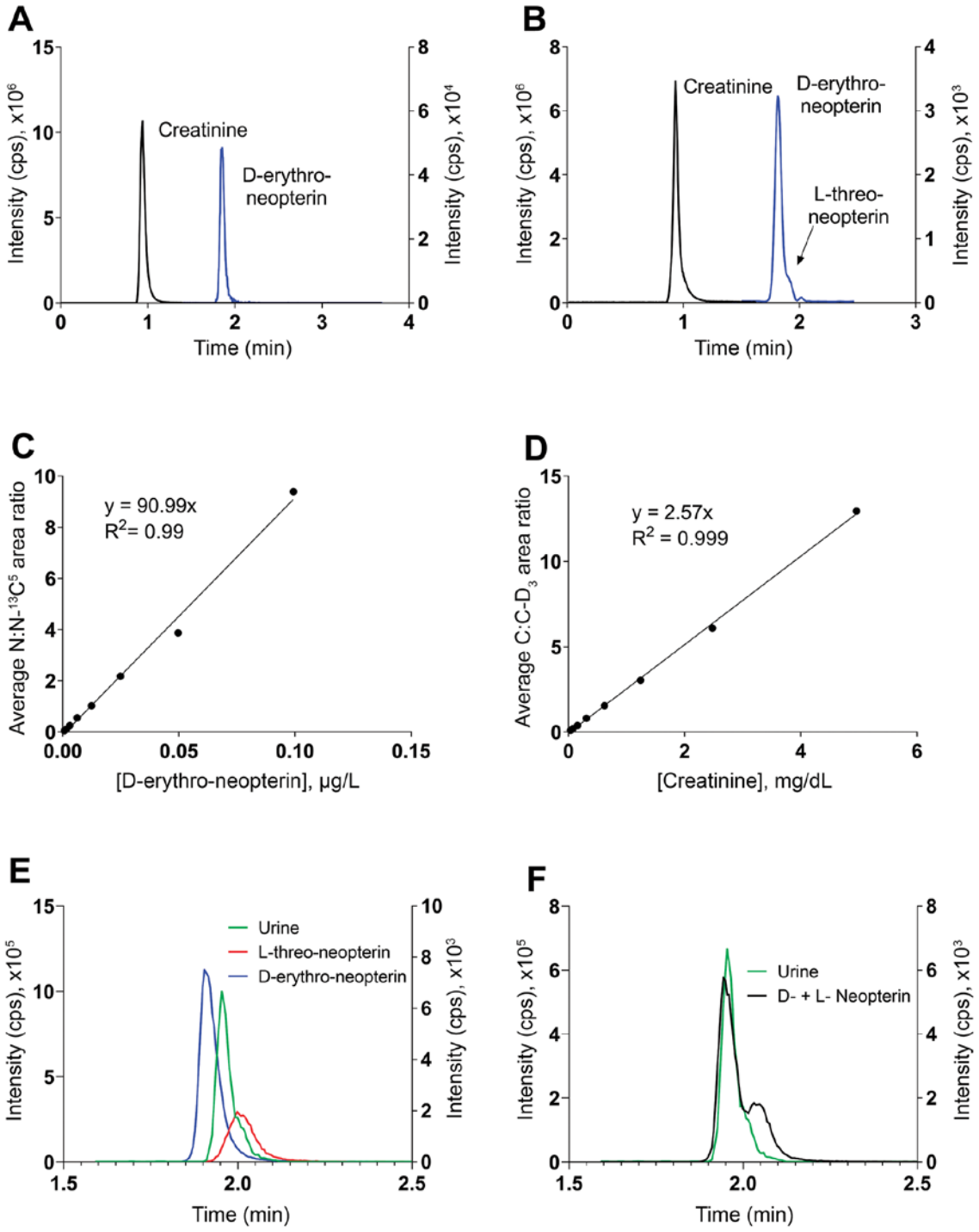


Figure 1. Total ion current (TIC) chromatogram for standards (A) and urine (B). Creatinine (left axis) elutes at 0.9 minutes and D-erythro-neopterin (right axis) at 1.8 minutes. Calibration curves for D-erythro-neopterin (C) and creatinine (D). E and F: Urine samples sometimes exhibited a “shoulder”, identified as L-threo neopterin by overlaying the TIC chromatograms of the two neopterin isomers (left axis) and urine (right axis).

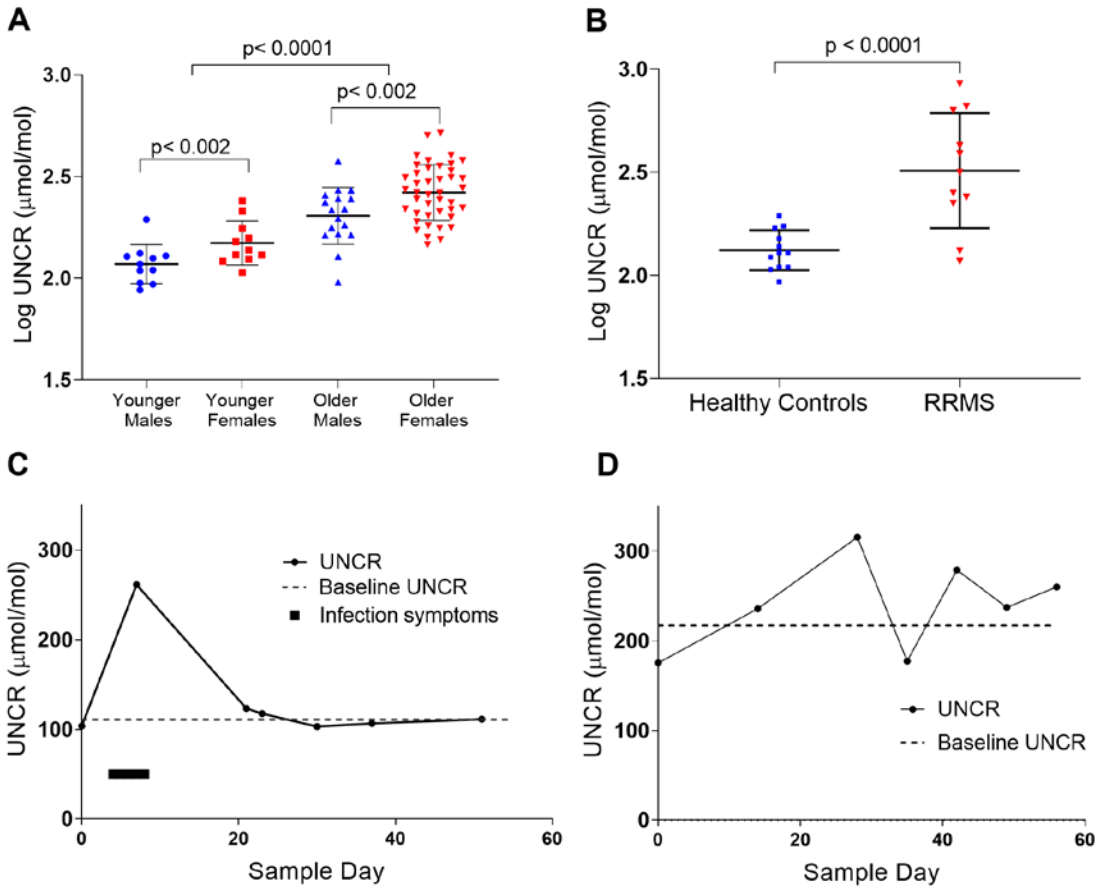


Figure 2. Urinary neopterin to creatinine ratio (UNCR) is significantly higher in females versus males and in an older population versus a younger one (A, *p* value from ANCOVA correcting for age and gender). UNCR is also significantly higher in relapsing remitting multiple sclerosis (RRMS) patients compared to healthy controls (B; *p* values are from two-way ANOVA correcting for age or gender respectively). Data is mean ± SD of log-transformed values. (C), UNCR increases with symptoms of infection. (D), UNCR can also increase in the absence of clinical symptoms suggestive of asymptomatic infection.

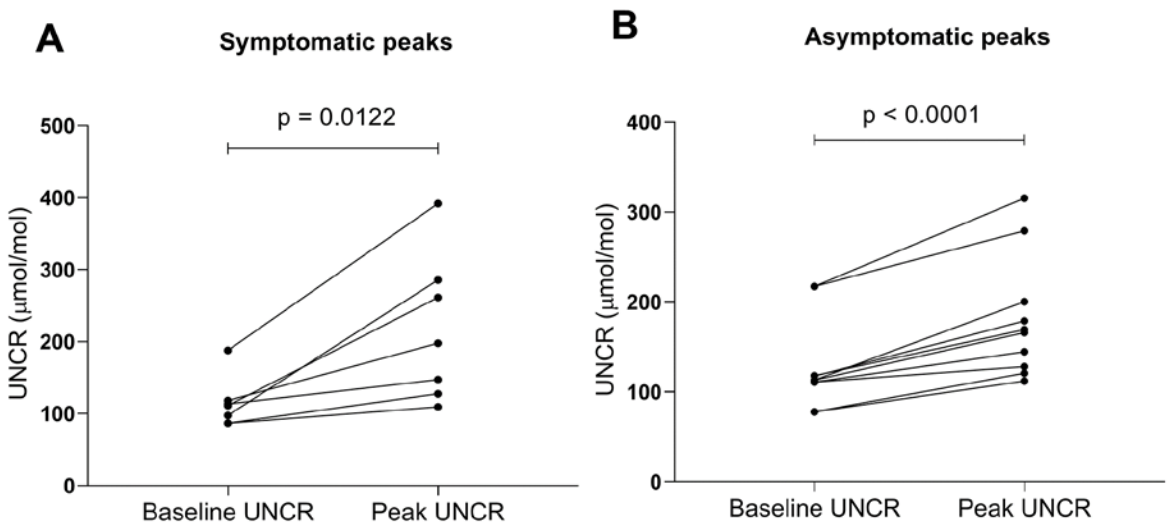


Figure 3. Each person's individual baseline urinary neopterin to creatinine ratio (UNCR) is paired with their respective peak UNCR during symptomatic (A) and asymptomatic (B) episodes. *p* Values are from one tailed paired t tests.

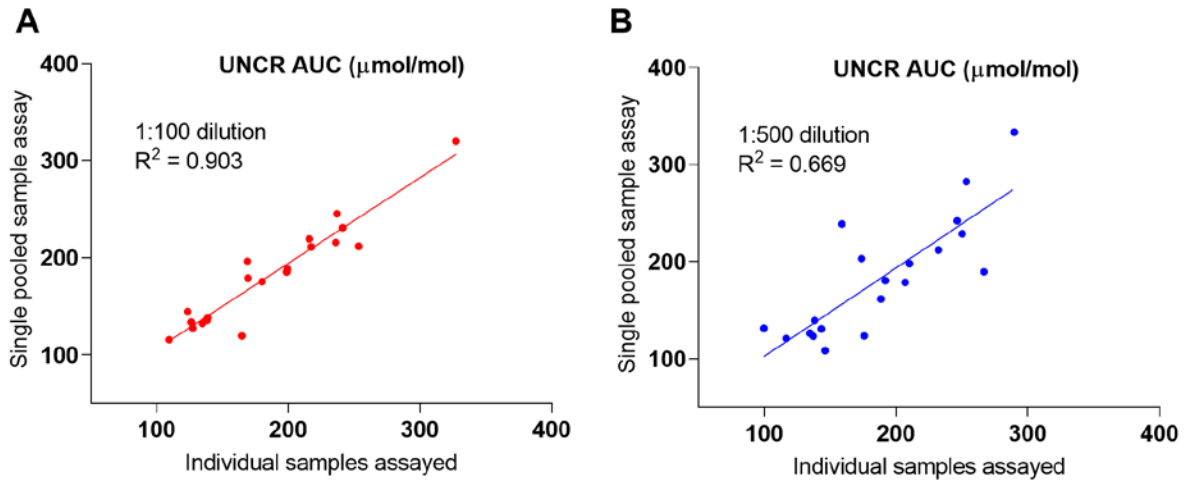


Figure 4. The urinary neopterin to creatinine ratio (UNCR) area under the curve (AUC) from serial sampling over a period of time can be measured using a single pooled sample (versus measurement of serial samples individually), particularly helpful when machine time or costs are a limitation. The two methods correlated well when samples for creatinine measurement were prepared at 1:100 dilution (A), but less well when a 1:500 dilution (B) was employed.

SUPPLEMENTARY MATERIALS FOR

High throughput UPLC-MS analysis of urinary neopterin-to-creatinine ratio

Supplemental methods

Chemicals and regents

Standards: D-erythro-neopterin (Enzo Life Sciences, Exeter, UK), L-threo-neopterin (Schirck's Laboratories, Bauma, Switzerland). Creatinine (Sigma-Aldrich, Gillingham, Dorset, UK). Creatinine-D_3 and D-erythro-neopterin-$^{13}C_5$ (Santa Cruz Biotechnology, Dallas Texas, US).

Solvents: Analytical-reagent grade acetonitrile, ammonium hydroxide and formic acid (Fisher, Loughborough, UK), HPLC-grade water obtained from a Milli-Q System (Merck, Watford, Hertfordshire, UK).

MS/MS optimisation

The detection parameters of MS were optimized using direct injection at a flow rate of 8 µL/min. Under electrospray ionization conditions, the protonated molecules ([M + H]+) of creatinine, creatinine-d_3, D-erythro-neopterin and D-erythro-neopterin-$^{13}C_5$ and their respective product ions were observed as peaks at the mass to charge ratios (m/z) described in the Methods. Both parent and daughter m/z values are the same, i.e. 254.1>206.1 m/z, as determined by direct infusion of both D- and L-neopterin standards.

LC optimization

During LC optimisation, acetonitrile and methanol were evaluated as organic mobile phase, as well as the volatile additives formic acid, acetic acid, ammonium formate and ammonium acetate. Initially two different HPLC columns, C18 (LiChrospher, 5

µm, 4 x 125mm) and Phenyl Hexyl (Luna, 3.0 µm, 3.0 x 150mm) (both from Phenomenex, Macclesfield, Cheshire, UK), were tested based on previous publications *(19) (21)* . The C18 column was found to be inappropriate on our platform due to poor resolution and peak shape for creatinine. The HPLC phenyl hexyl column was a viable candidate with good peak resolution, chromatographic peak shapes and signal-to-noise (S/N) ratios. A PFP column on a UPLC platform was found to be the best, however, with good peak shapes and separation, and short retention times, allowing for a higher throughput. The PFP column was selected since this was the column which delivered the best performance for both neopterin and creatinine combined. The PFP column was not selected to discriminate between D-erythro- and L-threo-neopterin.

Both isocratic and a range of gradient methods were explored using the PFP column, along with different flow rates, injection volumes and column temperatures. Starting at high aqueous (99%, solvent A) followed by a short gradient to between 90-95% aqueous (solvent A) gave the best results with optimum separation and peak shapes obtained with the gradient and conditions described in **Supplemental Table 1**.

LOD and LOQ

One of the most common and well accepted methods of calculating the limit of detection (LOD) and limit of quantification (LOQ) is to multiply a standard deviation by a suitable coefficient. The value of this coefficient is based on the t-statistic at a significance level of 0.05 and is usually rounded to 3 for LOD and 10 for LOQ. If, however, a more accurate estimation is required, the exact t-statistic based on the appropriate degrees of freedom should be used (see **Supplemental Table 4**). For a

more in depth review of these statistics, see references *(2)* & *(3)*. The standard deviation is obtained from sufficient replicate measurements (n>10) of blank or low-concentration sample(s). When using the standard deviation it is important that it is representative of typical results i.e. it should be adjusted for any averaging or blank corrections that may occur during standard procedure.

Other methods for calculating LOD and LOQ exist and may be more suitable depending on the type of assay and user requirements. One option is to use the regression line obtained when serially diluting a sample. The standard deviations of the residuals of the regression line or the standard error of the calibration can be used in the calculation (**Supplemental Table 4**) *(4)*. For analytical methods in which there is significant baseline noise, such as in chromatography, a signal to noise ratio (S:N) is often utilised. A measurement for signal is taken as the maximum height of the signal (S) above a baseline (N) as shown in **Supplemental Figure 4**, and the baseline noise is estimated by measuring the distance between the minimum and maximum peaks for 30 seconds before and after the peak. A S:N of 3:1 for LOD and 5:1 or 10:1 for LOQ is usually used *(5)*. This approach often involves trial and error. It can be of some value to conduct a preliminary “visual” investigation to establish a minimum value at which the analyte can be reliably detected. Indeed for certain situations, performing a visual evaluation for LOD may be adequate. NATA guidelines recommend plotting a response curve of percentage positive results versus analyte concentration and from this, a threshold concentration at which the test becomes unreliable can be determined *(6)*.

Supplemental Table 1: Inlet method. The primary mobile phase (A) was 0.2% formic acid and the co-solvent (B) consisted of 0.2% formic acid in acetonitrile.

Time (min)	**Flow rate (ml/min)**	**%A**	**%B**
0.0	0.2	99.0	1.0
0.5	0.2	99.0	1.0
3.0	0.2	93.5	6.5
3.5	0.6	0.0	100.0
5.5	0.6	0.0	100.0
6.0	0.6	99.0	1.0
7.7	0.2	99.0	1.0
8.0	0.2	99.0	1.0

Supplemental Table 2: Characteristics of relapsing-remitting multiple sclerosis (RRMS) patients and healthy controls (HC). Values are mean (standard deviation). EDSS = Expanded Disability Status Score. Difference in means is by unpaired t-test, except in [(a)] Fisher's exact test.

	HC (n = 12)	**RRMS (n = 11)**	**p-value**
Age (years)	31.2 (10.8)	43.4 (10.7)	0.013
Sex (% female)	66.7	72.8	0.556 [(a)]
Disease duration (years)	-	11.4 (9.0)	-
EDSS	-	2.5 (1.9)	-
Treatment type			
No treatment	12	4	-
Interferon	-	4	-
Glatiramer	-	2	-
Fingolimod	-	1	-
Alemtuzumab	-		

Supplemental Table 3: Characteristics of progressive multiple sclerosis (MS) patients. None were on immunomodulatory treatment. Values are mean (standard deviation). EDSS = Expanded Disability Status Score.

Age (years)	55.8 (7.9)
Sex (% female)	30
Disease duration (years)	10.9 (8.6)
EDSS	4.625 (1.9)
Progressive MS type	
% Primary	90
% Secondary	10

Supplemental Table 4: A comparison of methods for calculating LOD and LOQ, using data obtained from 12 replicates of 2 low neopterin concentration urine samples. See footnotes below.

Method		**LOD µg/L**	**LOQ µg/L**	**Reference**
Using Student t-statistics*	cβ x SD (a)	0.305	1.018	Eurachem, 2014 *(7)*
	cβ x SD (b)	0.339	1.018	CLSI, 2004 *(2)*
	cβ x SD (c)	0.346	1.018	Currie, 1999 *(3)*
Using the regression line**	(cβ x SR)/M	0.269	0.898	Miller and Miller, 2010 *(4)*
	(cβ x SE)/M	0.271	0.902	NATA, 2013 *(6)*
Signal to noise ratio***	LOD = 3 x S:N LOQ = 10 x S:N	0.03 - 0.1	0.2 -0.58	Agilent, 2011 *(5)*
Visual evaluation	LOD = signal seen 100% of time	0.2 - 0.58	NA	NATA, 2013 *(6)*

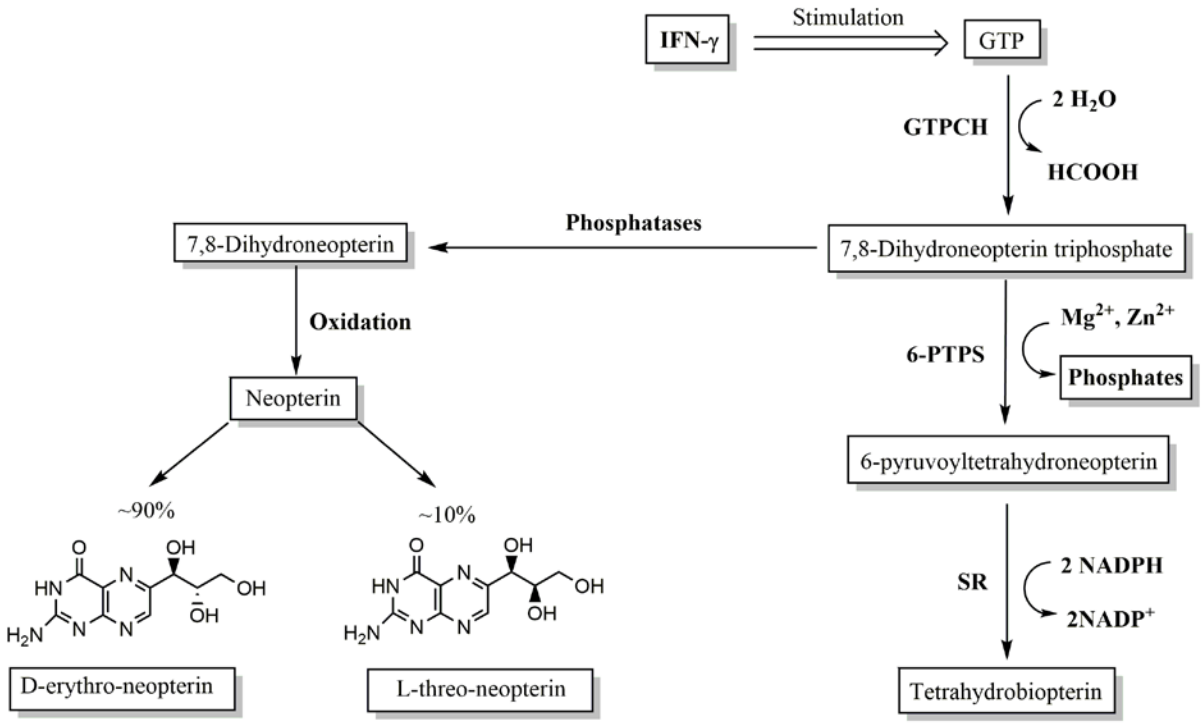
**SD = s/√n where s is the standard deviation obtained from replicates of low concentration urine samples divided by the square root of number (n) of samples averaged in standard procedure.*

cβ is the coefficient based on the t- statistic at a significance level of 0.05. For LOD, cβ can be rounded to 3 (a) or calculated more accurately using cβ = 2(1.645/(1-1/(4F)) where F is the relevant degrees of freedom (b) or even more accurately using cβ =(2t(4F/(4F+1)) where t = the t-statistic for the relevant F (c). For LOQ, cβ is always 10.

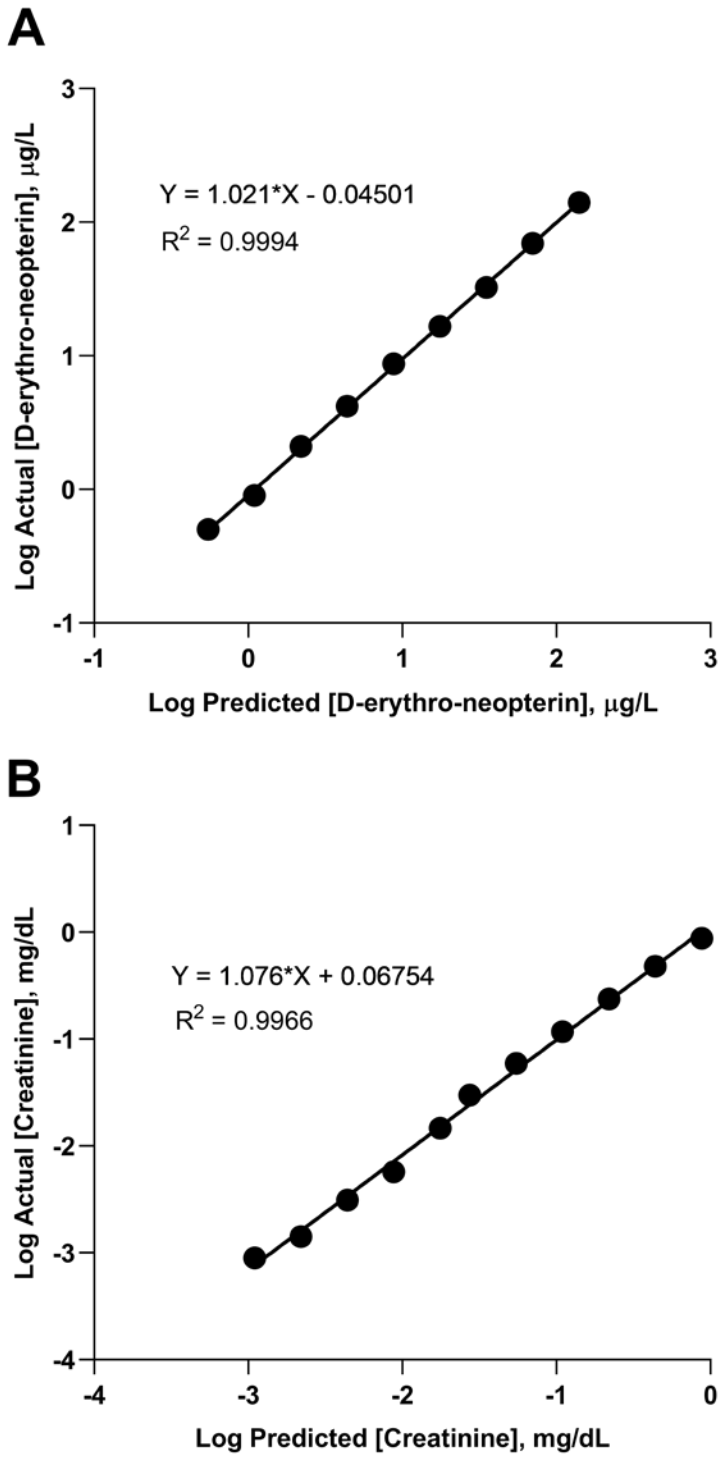
***cβ is set at 3 for LOD and 10 for LOQ, M is the slope of the regression line, SR is the root mean squared error or standard deviation of the residuals taken from the regression line and SE is the standard error of the calibration.*

**** LOQ was estimated using 10 x S:N. We found using the signal to noise ratio very unreliable as replicates of the same sample would produce widely varying amounts of signal.*

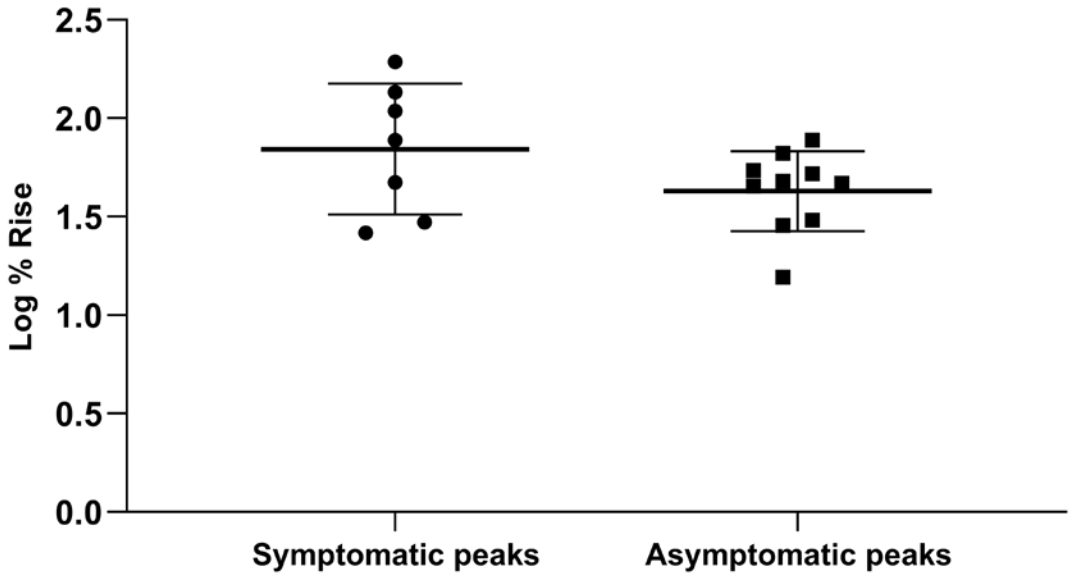
Supplemental Figure 1: Neopterin is synthesised as a by-product of the tetrahydrobiopterin *de novo* biosynthetic pathway. Upon interferon gamma (IFN-γ) stimulation, guanosine triphosphate (GTP) is converted to 7,8 dihydroneopterin triphosphate through the action of GTP cyclohydrolase I (GTPCH). Some of this is converted to tetrahydrobioperin through the actions of 6-pyruvoyl-tetrahydropterin synthase (6-PTPS) and sepiapterin reductase (SR) while some is converted to 7,8-dihydroneopterin through the action of phosphatases. 7,8-dihydroneopterin is oxidised into neopterin. Adapted from Ghisoni et al., 2015 *(1)*.



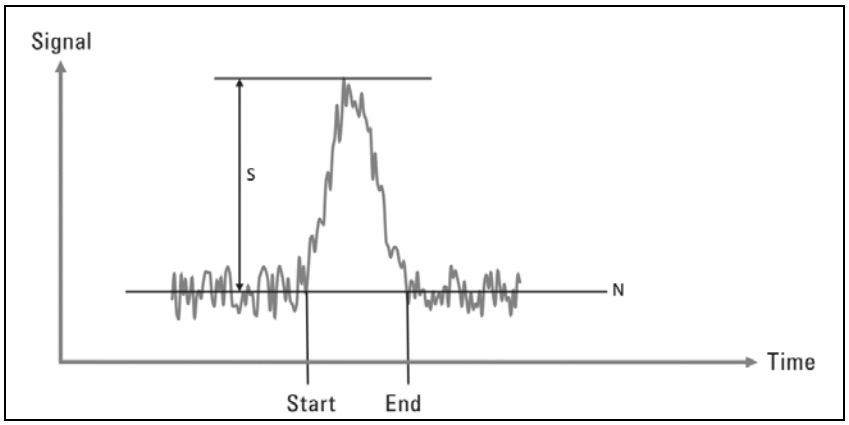
Supplemental Figure 2: D-erythro-neopterin (A) and creatinine (B) both show excellent linearity of dilution (range tested 0.5 - 140 µg/L and 0.001 - 0.88 mg/dL respectively). Values log transformed for better visualisation.



Supplemental Figure 3: For defined peaks, no significant difference in the percentage (%) rise from baseline urinary neopterin to creatinine ratio (UNCR) of symptomatic versus asymptomatic peaks was found (t (9.106) =1.516, p=0.1633, two tailed unpaired t test with Welch's correction).



Supplemental Figure 4: Example chromatogram demonstrating signal (S) to noise (N) ratio, modified from Agilent technologies, 2011 *(5)*.



References

1. Ghisoni K, Martins Rde P, Barbeito L, Latini A. Neopterin as a potential cytoprotective brain molecule. J Psychiatr Res 2015;71:134-9.
2. NCCLS. Protocols for determination of limits of detection and limits of quantitation; approved guideline. https://clsi.org/ (Accessed 13 December 2017).
3. Currie LA. Nomenclature in evaluation of analytical methods including detection and quantification capabilities. Analytica Chimica Acta 1999;391:105-26.
4. Miller JN, Miller JC. Statistics and chemometrics for analystical chemistry, sixth edition. Pearson Education Limited 2010.
5. Greg Wells HP, Charles William Russ IV. Signal, noise, and detection limits in mass spectrometry. https://www.agilent.com/cs/library/technicaloverviews/Public/5990-7651EN.pdf (Accessed.
6. NATA. Guidelines for the validation and verification of quantitative and qualitative test methods. https://www.nata.com.au/ (Accessed 03 Dec 2017).
7. Magnusson B, Örnemark U. Eurachem guide: The fitness for purpose of analytical methods – a laboratory guide to method validation and related topics. https://eurachem.org/ (Accessed 03 December 2017).