2023

Non-linear dilution adjustment of exemplary urine arsenic

PART I: CURVING THE LINES

THOMAS CLEMENS CARMINE MD

Etzelstrasse 21 • 8808 Pfäffikon SZ • Switzerland • Tel +41 55 415 50 30 • consult68@me.com

Non-linear dilution adjustment of exemplary urine arsenic
Part I: Curving the lines

Contents

Abstract **2**

1. Introduction **3**
1.1. Urinary dilution correction in biomonitoring 3
1.2. Pitfalls in conventional creatinine correction (CCRC) 3
1.3. Systemic calculation errors due to misperceived linearity 5
1.4. Arsenic in urine 6
1.6. Objectives of study 7

2. Study population **7**

3. Algebraic standardization of uAs_{UC} to 1g/L uCR **8**
3.1. Power functional relation between uCR and TWuAs 8
3.2. Dependence of exponent b on uAs_{UC} 8
3.3. Algebraic standardization of uAs_{UC} to 1g/L uCR 10
3.4. Validating standardization 11

4. Discussion **13**
4.1. Power-functional creatinine correction (PFCRC) 13
4.2. Standardization of other biomarkers and dilution adjustments 14
4.3. Benefits and limitations of standardization by V-PFCRC 14

5. Abbreviations **16**

6. Acknowledgements **16**

7. Author contribution **16**

8. Conflict of Interest **16**

9. Ethics **17**

10. Funding **17**

11. References **18**

11. Supplementary Material/Appendices **23**
Appendix 1: Age structure of study population by sex. 23
Appendix 2: Urinary creatinine by sex. 23
Appendix 3: Uncorrected urinary Arsenic (uAs_{UC}) by sex. 24
Appendix 4: Conventionally CR-corrected urinary Arsenic (uAs) by sex 24
Appendix 5: Distribution of absolute and percentual deviations of uAs_N from uAs_C. 24
Appendix 6: Multilinear regression analysis of contributions of uCR and uAsUC to %CRCE. 25
Appendix 7: Sex- and Age-differentiated efficacy of standardization. 25
Appendix 8a: Enlarged illustration of uCR/uAs-relations in sextiles 1-5 26
Appendix 8b: Compare quartiles of uAs_{UC}, uAs_C, and uAs_N for S-PFCRC (c) and V-PFCRC (d) 26
Appendix 9a: The uCR-differential linear relationship between coefficient *a* and uAs_{UC} 27
Appendix 9b: The uCR-differential logarithmical relationship between exponent b and uAs_{UC}. . 27
Appendix 10a: Effect of logarithmization on data distribution of uCR. 28
Appendix 10b-d: Effect of logarithmization on data distribution of uAs. 28

Non-linear dilution adjustment of exemplary urine arsenic
Part I: Curving the lines

Thomas Clemens Carmine M.D

Abstract

Background
Conventional dilution adjustment of spot-urinary biomarkers by correction for creatinine (uCR, CCRC), osmolality, or specific gravity remains controversial. Apart from unaccounted confounders like age, sex, muscle mass, or diet, the misperception of constant mass ratios between analyte and corrector over a wide hydration range entails deceptive correction errors foremost at both ends of the hydration spectrum. Mitigating creatinine-correction errors (CRCE) by restricting uCR-ranges to 0.3(0.4)-3 g/L generates high numbers of sample rejects and allows for misleading fluctuations within presumably tolerable uCR-ranges undermining the validity of clinical, forensic, and epidemiological investigations.

Methods
The CRCE for exemplary total weight Arsenic (TWuAs) was analyzed in a large set of n= 5599 unselected spot urine samples. After confining data to 14 – 80 years, uncorrected arsenic (uAs_{UC}) < 500 µg/l, and uCR < 4.5g/L, the remaining 5400 samples were partitioned, and a calculation method to standardize uAs_{UC} to 1 g/L uCR developed based on uAs_{UC}-stratified power functional regression analysis (PFRA).

Findings
The obtained standardizing method of variable power-functional creatinine correction (V-PFCRC) represents a progression of reportedly proven simple power-functional modifications of CCRC (S-PFCRC). In contrast to CCRC and S-PFCRC, standardization to 1g/L uCR by V-PFCRC yielded constant uAs_N-values over the entire uCR-range in all sextiles. Residual bias was largely neutralized in all sextiles of uAs_N (R^2 9E-06 – 0.04), compared to uAs_{UC} (R^2 0.93 – 0.99) and CCRC (R^2 0.42 – 0.89). The efficacy of CRCE compensation was confirmed by simple PFRA in both sexes and seven age bands. The strongest %CRCEs were found in samples at low concentrations of uCR (SBC -0.96) and/or uAs_{UC} (SBC -0.11).

Interpretation
Standardization to 1g/L uCR by V-PFRA allows for more reliable dilution adjustment of spot urinary results than CCRC and S-PFCRC. As an easily adoptable, novel approach, the calculation method developed in this study can neutralize residual dilution bias in urinary biomarkers adjusted by uCR, SG, and osmolality. Corrective V-PFCRC equations can be readily applied to laboratory software for both conventional reports and multiple linear regression models and. While minimizing sample rejects, V-PFCRC will improve reliability and comparability of urinary biomarkers.

Funding
The study was not funded.

Non-linear dilution adjustment by the example of urinary arsenic
Part I: Curving the lines

Thomas C. Carmine M.D.

1. Introduction

1.1. Urinary dilution correction in biomonitoring

Urine is the most used specimen in occupational and environmental biomonitoring of toxic metals.[1] For better practicability and compliance, spot urine is commonly favored to 24h urine samples which, unlike spot urine, account for diurnal variation and, to a degree, for hydration. However, spot urinary concentration of a biomarker expressed as mass per urine volume (e.g., µg/L) is an unreliable parameter. It primarily depends on hydration, and variable fluid intakes produce significant differences in urinary concentrations, with diluted urines usually yielding lower biomarker readings than concentrated ones (Fig.1a).[2] In the case of urinary arsenic (uAs), more than a third of the variation in spot urinary concentrations could be attributed to dilution.[3] To standardize intra- and interindividual concentrations over time, fluctuations in spot urine dilution are commonly adjusted for either by urinary creatinine (uCR) or less frequently by specific gravity (SG) and osmolality as proxies for urinary flow rate (UFR).[3]

All these conventional methods, however, only partially adjust for variation in diuresis.[1,3,4,5,6,7,8] Creatinine (CR), a nitrogenous waste product from muscular energy generation by breaking down creatine-phosphate, has been reported to be non-negligibly influenced by age, sex, ethnicity, lean muscle mass, meat-/creatin intake, and collection time, thus complicating inter- and intraindividual comparisons.[6,9,10,11] Furthermore, relations between analyte concentrations and UFR have been determined log-linear instead of linear, as wrongly assumed in dilution adjustments by CCRC and specific gravity/osmolality.[3,5] Despite these obvious fallacies, CCRC, specific gravity, and osmolality remain the most common default approaches to urinary dilution adjustment. All reasonable suggestions to mathematically account for excretion differences have yet to find their way into common clinical practice.[3,4]

1.2. Pitfalls in conventional creatinine correction (CCRC)

In CCRC, the analyte concentration is divided by uCR, thus converting the scale of results to the weight of analyte per weight of creatinine (analyte µg/uCR mg or g). This tends to standardize results because adults excrete CR at a relatively constant daily rate expressed as mg/kg body mass.[12] The justification of CCRC is essentially based on three implicit assumptions:

1) Individual and interindividual uCR-excretion rates were constant over time and vary only by UFR, while the influences of muscularity, age, ethnicity, exercise level, and diet (vegetarian vs. omnivorous) were negligible compared to variability in hydration.[3,12]
2) There was no confounding of CR and analyte by common causal factors nor direct causal relationships between the two.[2,13]
3) The relative ratio between the mass of the analyte and the mass of uCR was preserved with variations in hydration and UFR.

All three assumptions are inadequate and gave rise to criticism and even requirements for complete renunciation of CCRC or its substitution by multiple linear regression models (MLRM), including uCR as a covariate.[14,15] Most objections are directed to the unaccounted factors mentioned in assumption 1, entailing uncontrolled

fluctuations of uCR and analyte/uCR-ratios.[6,8,9,10,13,15] In contrast, the significant flaws in assumptions 2 and 3 appear to be less of a topic in the literature.[3] Dependent biological variables standardized by a denominator (here, uCR) generally tend to present unusual statistical properties, potentially impairing the sensitivity and precision of a marker.[11,16]

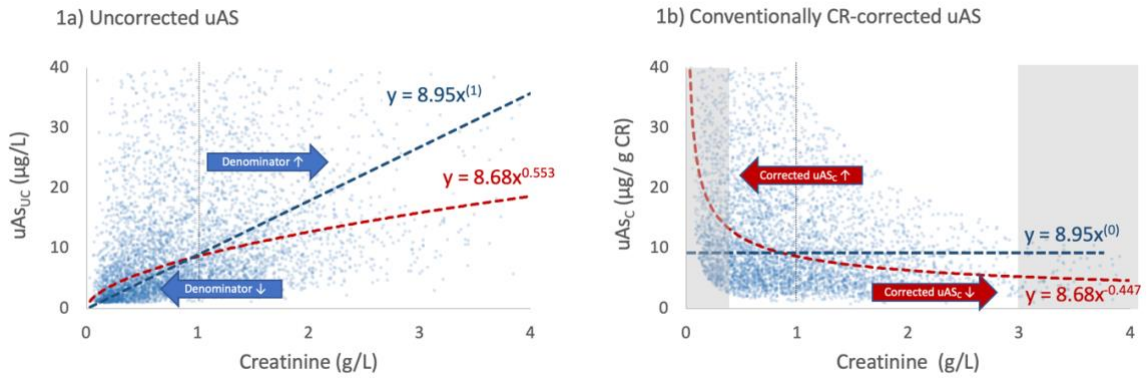


Figure 1a/b: Relationship between uCR and TWuAs in linear and power functional regression analysis (PFRA). Subset of all samples < 40 µg/L TWuAs (n=4854) for uncorrected (1a) and conventionally CR corrected (1b) TWuAs. Compare simple linear (blue line) and simple power functional regression (red line) of unstratified data. CCRC assumes linearity, which can be understood as a particular power function form with the exponent b (uAs_{UC})=1. If the assumption of linearity were adequate, the simple division of uncorrected TWuAs by uCR would eliminate the dependency of corrected TWuAs from uCR. Consequently, regression analysis in Fig.1b would produce the parallel blue line to x-Axis (here y = 8.95x/x = 8.95), yielding identical TWuAs results over the full range of uCR (= in all hydration states). Conventional CR-correction (CCRC), however, produces an inverse dependency of TWuAs from uCR (red line, Fig.2b). Instead of the hypothetical parallel to X-axis (blue line, Fig.1b), another power function can be seen after regression analysis now with a negative exponent b (here -0.447, Fig.1b). The inverse relationship results from the corrective denominator (uCR $^{(1)}$) used in CR-correction being too small for samples <1 g/L uCR and too high for urines > 1g/L uCR. The exponent b in uAs_C reflects the difference between the adequate exponent b (here 0.553, Fig.1a) and the erroneous exponent 1 used in conventional CR correction: $b_{uAsC} \approx b_{AsUC} - 1$. The more minor b in uAs_C, the more negligible the dependence of uAS_C on uCR, and the more parallel the regression line to the X-axis, the more adequate the corrective function. In this case, the division of uncorrected TWuAs by $uCR^{0.553}$ (instead of uCR ^{1}) would have yielded a regression function in Fig.1b of $y \approx 8.68\ x^{0.553 - 0.553} = 8.68\ x^0 = 8.68$, in other words, standardization of uAS_C to 1g/L uCR. The shaded areas in Fig.1b represent the uCR ranges considered ineligible for biomarker quantification, and samples within these areas will be considered invalid. Even in the restricted ranges of 0.3(0.4) - 3 g/L uCR, a considerable bias remains, with the regression curve deviating from the desired parallel to the abscissa.

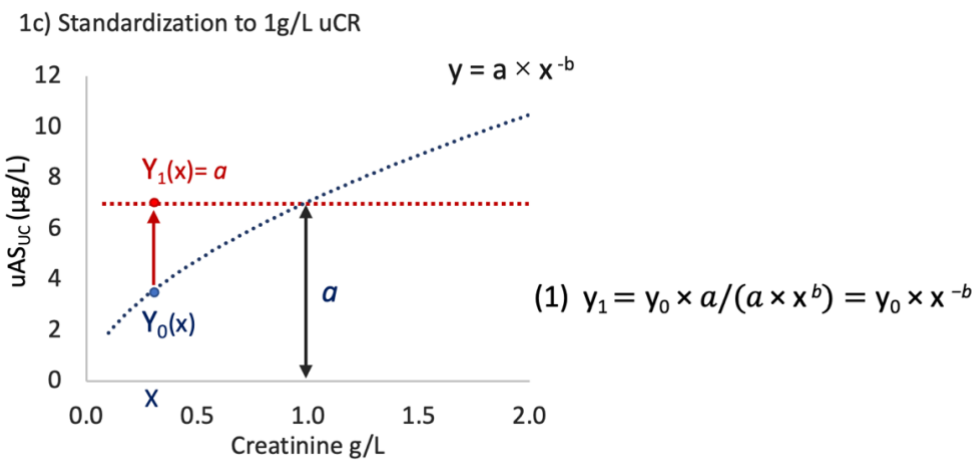


Figure 1c: Power functional dependency of uAs_{UC} (Y_0) from uCR (x) as the algebraic base for standardization of uAs_{UC} to 1g/L uCR. Normalized uAS_N (=Y_1) is yielded by simple division of uAs_{UC} by $(uCR)^b$. This operation generates identical uAs_N (=Y_1) across the whole range of uCR (dotted red parallel to the abscissa), whereby the uAs_N equals coefficient *a* at uCR=1 g/L.

To avoid ratio forms of data, it was proposed to take the denominator variable as a covariate (amongst age, sex, or ethnicity) in MLRM.[1,2,16,17] The covariate solution, however, is complicated by the potential role of uCR as a so-called 'collider', potentially inducing noncausal associations and further confounding.[2,13] Moreover, it was pointed out that adjusting for uCR as a covariate may not adequately control for the measurement error resulting from inter-individual variations in urinary dilution.[2] While more sophisticated MLRM reportedly minimized bias in a range of simulated theoretical scenarios, their broad implementation remains complicated by demographic heterogeneity, potentially necessitating regional adjustments of MLRM, classification issues for patients of minority- or mixed ethnical backgrounds, elaborate pre-analytical questionings, and an incomplete understanding of the number, nature, and complex interrelations of covariates to be included in MLRM.[2,13]

1.3. Systemic calculation errors due to misperceived linearity

While the controversy over the most accurate dilution adjustment in urinary biomonitoring continues, there seems to be only consensus on the non-validity of uCR-corrected results (uAs_C) in excessively dilute or concentrated urines.[1,2,3,17,18]
In fact, the net excretion rate ratio between uCR (or SG/osmolality determinants) and a biomarker does neither seem to be independent of biomarker concentrations nor of UFR, as erroneously assumed in conventional dilution adjustments by CCRC and SG/osmolality. Analyte and uCR (likewise SG) usually do change their proportionality in response to UFR, and the over-simplified division of analyte by uCR generates creatinine correction errors (CRCE) either with false elevations in diluted and declines in concentrated urines, like in total weight uAs (TWuAS, Fig.1a/b), or vice versa.[3] Despite its mathematically tangible and accountable nature, this systemic methodological error inherent in all common dilution adjustments (uCR > SG > osmolality) is mitigated mainly symptomatically by restricting dilution ranges.[3] Limiting valid ranges of uCR to 0.3 (0.4) – 3 g/L uCR, however, does render a good portion of samples (in the present study, a total of 22.3%) formally uninterpretable (Fig.1b) and will eminently affect lower uCR samples of females (29% rejection rate in this study), smaller infants (30% rejection rate in < 7-year-olds), and highly diluted samples after excessive intake or recent IV administration of fluids.[10] Since it does not address the root cause of misperceived linearity, uCR-restriction neither provides hydration-independent results within the presumed acceptable range (Figs.1a/b,2c). As the remaining hydration bias will exceed the magnitude of most other uncontrolled covariates, it involves a considerable potential for misjudgment of actual exposures, underlining the need for more accurate dilution adjustment.

Several investigators have already proposed mathematical modifications of CCRC to account for excretion differences in corrective equations.[3,5,18,19,20] A more recent study by Middleton et.al (2019) compared the efficacy of various dilution corrections for spot-urinary arsenic. The authors reported significant mitigation of CRCE by mathematically accounting for the analyte-specific discordance between the respective slopes of uCR and the analyte in relation to UFR.[3] The major improvements of the simple potency-functional approach over CCRC were demonstrated in terms of both the minimal residual dependence of As_N on uCR (R) and the degree of correlation of As_N with actual drinking water exposure. The authors modified CCRC by a simple power functional uCR correction (S-PFCRC) to standardize results to 1g/l uCR.[3,5] Standardized arsenic (uAs_N) was expressed as a function of uncorrected uAs (uAs_{UC}) and uCR corresponding to Eq.1 of the present study, being fundamental to the method developed herein (Fig.1):

$uAs_N = uAs_{UC} \times uCR^{-z}$ (z is equivalent to *b* in the present study, see Fig.1).

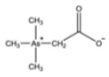
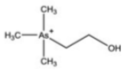
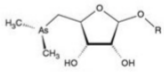
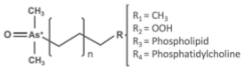
Theoretically, z is empirically derived by the ratio of the regression slopes of uAs (b_{uAs}) against UFR by that of uCR (b_{uCR}) against UFR ($z = b_{As}/b_{uCR}$).[3] To remove dilution bias from the sample set, the authors adopted a numeric approach to derive the prior unknown exponent z.[20] After computing uAs_N for a range of z values, z = 0.8 was considered most adequate based on the weakest linear correlation of uAs_N and uCR, with the R-value closest to zero understood as an indicator of the lowest residual dilution variation in uAs_N.[3] The authors resorted to this auxiliary method of z determination due to the absence of comprehensive empirically derived data while acknowledging that forcing a correlation of absolute zero this way may not be completely robust. Indeed, in the present study, the corrections achieved by S-PFCRC, though vastly improved compared to CCRC, appeared to be still skewed in the edges due to neglecting the dependence of corrective efficacy on the analyte concentration (Fig.3, App.8b). Another limitation of their study was the modest sample size of 202 spot urines, which precluded systematic analyses of demographic influences like sex or age on correction performance. Nevertheless, their results clearly demonstrated substantial improvements in CCRC results by S-PFCRC, which achieved equivalent minimal residual dilution dependencies for uAs_N as for analogously modified dilution adjustments by SG and osmolality.

In the present retrospective study, the comparatively large set of 5599 samples permitted the development of a more sophisticated progression of S-PFCRC. To additionally account for variations of z (here exponent b), as a function of uAs_{UC} levels, uAs_{UC}-stratified analyses of data sets (Fig.2a) proved to be a necessary step, as S-PFCRC of the entire, non-stratified data set (Fig.1a) preferably mitigates CRCEs in uAs_{UC} levels closer to the mean, with considerable residual dependencies between uCR and uAs_N in the marginal uAs_{UC}-quartiles 1 and 4, as revealed by stratified PFRA of full datasets being corrected with S-PFCRC (App.8b). This apparent weakness of S-PFCRC was addressed by the variable PFCFC (V-PFCRC) described in Sec.3.[3] Due to the comparatively large number of n=5599 samples in the present study, the effects of age and sex on the corrective functions could be investigated in more detail. They will be explained in Part II.

1.4. Arsenic in urine

According to EPA and ATSDR, the carcinogenic metalloid Arsenic (As) is the most ubiquitous and human-health-relevant environmental toxin.[21] Regulatory guidelines exist only for inorganic As (iAs), which is highly toxic in contrast to most organic forms of As (oAs).[26-28] In most studied populations, the methylated metabolites of iAs (DMA, MMA) predominate in urine besides Arsenobetaine (AsB, so-called 'fish-arsenic'), with however considerable interindividual variations in TWuAs composition (Tbl.1).[29,30] In clinical settings, TWuAs, which comprises all ICP-MS-detectable forms of uAs, is usually measured as a first-line test with speciation reserved for high TWuAs samples. Besides TWuAs, many studies measure the sum of speciated iAs and their metabolites (DMA, MMA) as they are pragmatically considered a parameter for chronic exposure to drinking water, albeit neglecting the potential contribution of marine oAs to DMA (Tbl.1).[31-36] This sum is often misleadingly referred to as 'total arsenic' although it does, except MMA and DMA, not contain other oAs like AsB. In the present retrospective study, TWuAs was chosen as a paradigmatic marker because of its abundant presence in human urine in a broad concentration range. [21-25]

Table 1: Compilation of the most abundant As species found in humans.

Name	Abbr.	Chemical Structure	Occurence/Origin[23,41,42]	Toxicity[43,44]	% of TWuAs [45-47]	% of total blood As[46]	pKa[48]
Arsenite (inorganic)	As(III) iAs(III)		Drinking water, Rice, Flour, various foods, Pesticides	+++	1-3 %	20.2 %	pKa_1=9.23 pKa_2=12.13 pKa_3=13.4
Arsenate (inorganic)	As(V) iAs(V)		Drinking water, Rice, Flour, various foods, Pesticides	++	0.1-0.3%		pKa_1=2.19 pKa_2=6.98 pKa_3=11.53
Dimethylarsinic Acid	DMA DMA(V)		Seafood (small amounts) Primary metabolite of iAs, Arsenosugars, and -lipids	DMA(V) + DMA (III) ++++	30-41%	11.7 %	pKa=6.2
Monomethylarsonic Acid	MMA MMA(V)		Seafood (small amounts) Metabolite of iAs, Arsenosugars, and -lipids	MMA(V) + MMA(III) +++++	2-5%	2.4 %	pKa_1=4.1 pKa_2=8.7
Arsenobetaine (75-100% of 'Fish-Arsenic')	AsB		Mollusks, Crustaceans, Cephalopods, Finfish, Shellfish, Demersal Fish	(+)?	55-65%	51.1 %	pKa=2.18
Arsenocholine	AsC		Seafood, Fish, Fish oils	(+)?	<1%	4.5 %	
Arsenosugars	AsS		Major As compound in seaweed (brown kelp). Present in mussels, oysters, clamps	(+)?	Metabolized by humans mainly into DMA[69]		
Arsenolipids 1. As-containing Hydrocarbons 2. As-containing Fatty Acids 3. As-containing Phospholipids 4. As-containing Phosphatidylcholines	1. AsHCs 2. AsFAs 3. AsPLs 4. AsPCs		1. Fish oils 2. Fish oils 3. Seaweed 4. Roe	(+++)? (+)? (+)? (+)?	Metabolized by humans mainly into DMA[70]		

1.6. Objectives of study

The aims of the present study were:

1) assessing the dependence of TWuAs on uCR

2) developing a reliable and easily implementable mathematical correction method for standardization to uCR of 1 g/L, also applicable to other urinary biomarkers.

3) establishing the validity of the standardization function found by comparing residual dependencies in aggregate and sex-/age disaggregated data applying PFRA.

2. Study population

Anonymized data (sex, age, uCR, and TWuAs determined by GC-MS) of n = 5767 spot urine samples obtained between January 2014 and August 2022 were kindly provided by IMD Institute for Medical Diagnostics Berlin-Potsdam, Germany. The 155 samples (2.7%) presenting uAs_{UC} levels below the detection limit of 1 µg/L and nine outliers with uAs_{UC} > 500 µg/L were precluded from the study, which left a total of 5599 samples (2971 females and 2628 males) for analysis. For determining the standardization formula, a further six samples with extremely high creatinine values > 4.5 g/L and samples from donors outside the age band 14 – 82 years were not considered. Detailed data concerning muscle mass, dietary and exercise habits, clinical status, kidney function, exposure to toxins, race, and daytime of sample collection could not be provided by IMD. The age structure by sex is illustrated in App.1. Mean age (± SD) of females (51.5 ± 14.7) and males (50.7 ± 16.7) did not differ significantly (Mann-Whitney-U-Test). Conversely, significant male-dominant sex differences ($p < 0.001$) existed about uCR (F 0.82 ± 0.01, M 1.2 ± 0.014 g/L, means ± SE) and uAs_{UC} (F 17.8 ± 0.62, M 26.6 ± 0.9 µg/L), as illustrated in Appendices 2 and 3. In 82% of samples, uAs_{UC} was < 30 µg/L, in 95.7% < 100 µg/L and nine outliers were found with uAs_{UC} > 500 µg/L (548 – 5,763 µg/L). In contrast to uAs_{UC}, uAs_C was significantly higher ($p = 0.014$) in females (F 25.4 ± 0.83, M 23.9 ± 0.72 µg/g CR, App.4).

3. Algebraic standardization of uAs_{UC} to 1g/L uCR

3.1. Power functional relation between uCR and TWuAs

Figures 1 and 2 illustrate the rather power functional than linear relations between uCR and uAs_{UC}. The power function between uAs_{UC} and uCR can be formally described as:

(1) $uAs_{UC} = a \times uCR^{b}$ (Fig.1c)

The coefficient *a* represents the expected uAs_{UC} value for uCR = 1g/L:

(2) $uAs_{UC}\ (\mu g/L) = a \times 1^{b}\ (g/L) = a \times 1 = a$ *(for uCR =1 g/L)*

Any adequate corrective function for uAs_{UC} is expected to generate identical corrected TWuAs values over a broad uCR concentration range, reflected by a parallel line to the abscissa (Fig.1b). This can be achieved by algebraic standardization to uAs_{UC} at 1g/L creatinine (Fig.1c):

(3) $uAs_{N} = uAs_{UC} \times a / a \times uCR^{b} = uAs_{UC} \times uCR^{-b}$

Exponent *b* describes the slope of the power function between TWuAs and uCR. In the case of uAs_{UC}, *b* usually takes on values between 0 and 1 (here b =0.558, red line, Figure 1a). The resulting curve is ascending but flattening with increasing uCR, which means the uAs_{UC}/uCR ratio is incrementally diminished in higher concentrated urines, probably due to disproportional renal adjustment of TWuAs- and uCR- net excretions to urinary flow.[31,49] A particular case exists in *b,* taking on the value 1, which is the implicit flawed assumption of CCRC. Given the free glomerular filtration of both uCR and TWuAs, *b* = 1 would require the ratio of net secretions of both substances to be constant over a wide urinary dilution range, meaning both hydration and renal arsenic load would have to influence the respective tubular secretions of TWuAs and uCR to the same extent.

In contrast, when *b* approximates zero, a parallel to the x-axis results with no correlation between uCR and TWuAs. In the case of uAs_{C}, *b* = 0 indicates optimal correction for uCR with no remaining bias. In uAs_{UC}, an unrealistic *b* of 0 would render any adjustment for uCR superfluous.

3.2. Dependence of exponent b on uAs_{UC}

To adequately determine the dependence of exponent b on uAs_{UC}, the misleading effects of extreme outliers and potential age biases were mitigated by restricting data to samples of the age band 14 - 82 years with $uAs_{UC} \leq 500$ µg/L and uCR < 4.5 g/L. The 5400 samples fulfilling all three criteria were stratified into 15 uCR ranges and 6 uAs_{UC}-sextiles (Fig.2a). Power-functional regression curves were determined for each uAs_{UC} sextile separately (Fig.2b). The prior partitioning and uAs_{UC}-sextile-stratified PFRA of data (Fig.2a) proved to be a necessary step, as S-PFCRC preferably corrects uAs_{UC} values closer to the mean (App.8b). Despite of a decent overall computational correction of CRCE, there remain substantial residual dependencies between uCR and uAs_{N} in the marginal quartiles 1 and 4, as revealed by PFRA (App.8b).

Alternatively, to the power function used, the relationship between uCR and uAs_{UC} could be described by a logarithmic function of the form $uAs_{UC} = a + b \times \ln(uCR)$. However, the logarithmic regression curves do not intersect the abscissa at the zero-point (as PFRA and CCRC curves do) but in the positive range leading to nonsensical negative uAs_{N} correction values at very high urine dilutions (here uCR < 0.11 g/L, data not shown).

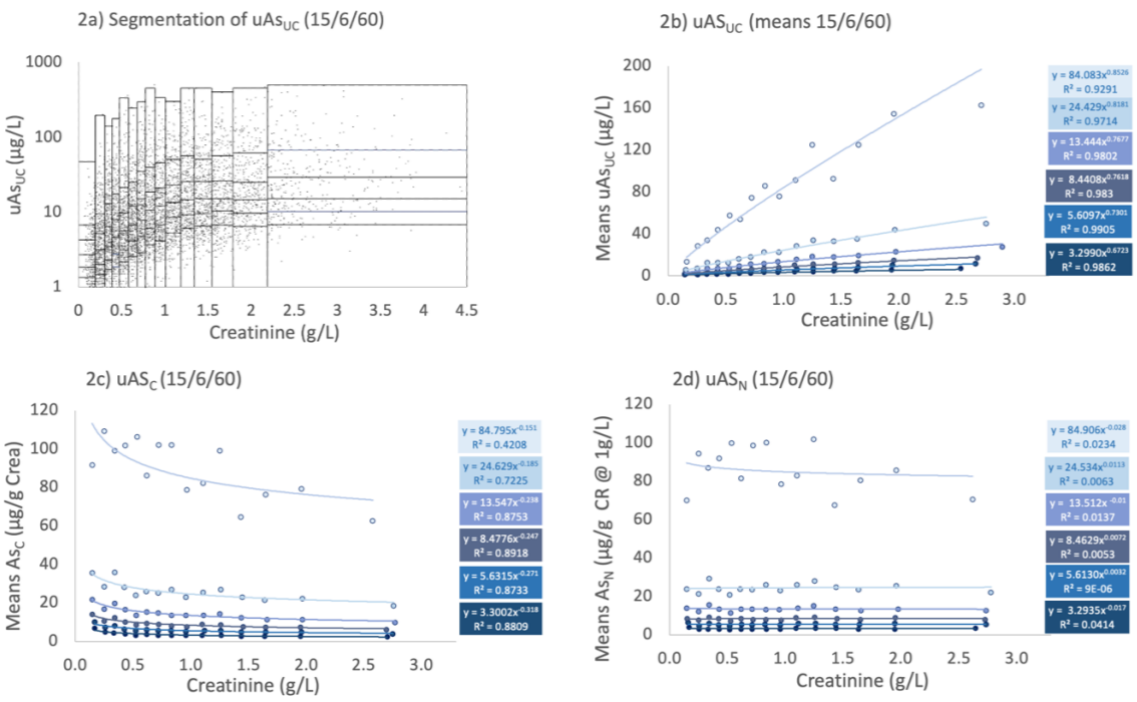


Figure 2a: Apportionment of samples into 90 equinumerous rectangles (60 values/rectangle).
All 5400 samples of the age band 14 – 82 years with uAs$_{UC}$ ≤ 500 µg/L and uCR < 4.5 g/L were partitioned into 15 ascending uCR ranges (quindeciles) of 360 values each, and all uCR-quindecil were further divided into uAs$_{UC}$-sextiles of 60 samples each. This way, the area of the graph was subdivided into 90 equinumerous rectangles of variable size.
Figure 2b: Power functional regression analysis of uAs$_{UC}$ sextile based on means of rectangles (Figure 2a). Dots represent the arithmetical means of uAs$_{UC}$ and uCR for each rectangle. Separate power-functional regression analyses of all sextiles yielded 6 equations taking the general form $y = a \times x^b$ with R^2 indicating the dependence of uAs$_{UC}$ on uCR.
Figure 2c: **Power functional regression analysis of uAs$_C$.** Apportionment of the 5400 samples into 90 equinumerous rectangles was performed for uAs$_C$ analog to the analysis of uAs$_{UC}$ in Figs.2a/b. Regression functions, including R^2, were calculated for uAs$_c$-sextiles
Figure 2d: **Power functional regression analysis of normalized uAs (uAs$_N$).** The correction for TWuAs-standardization to 1g/L uCR was determined $uAs_N = uAs_{UC} \times uCR^{-[Ln(uAsUC) + 11.82)]/[Ln(uCR) + 18.73]}$ (Sec.3.3, Eq.8). The 5400 uAs$_N$ values were subdivided and analyzed by analogy with uAs$_{UC}$ (Fig.2b) and uAs$_C$ (Fig.2c).
An enlarged visualization of sextiles 1-5 is given in App.8a.

The coefficients *a* and exponents *b* of the 6 curves of Figure 2b are related to each other in a positive logarithmic way (Fig.3), which forms the basis for the variable power function correction derived in Eqs.1-8. The coefficient *a* of a regression curve equals its uAs$_N$ value, which is assigned to all uAs$_{UC}$ values of that curve along the total range of uCR (Figs.1c,2b).

The log relation between exponent *b* and coefficient *a* takes the general form of:

(4) $b_{(a)} = c \times Ln(a) + d$ *(here $b = 0.053 \times ln(a) + 0.631$; $c = 0.0534$, $d = 0.631$)*

Solving (4) for *a* yields:

(5) $a_{(b)} = exp\ (b/c - d/c)$ *(here $a = exp\ (18.73 \times b - 11.82)$)*

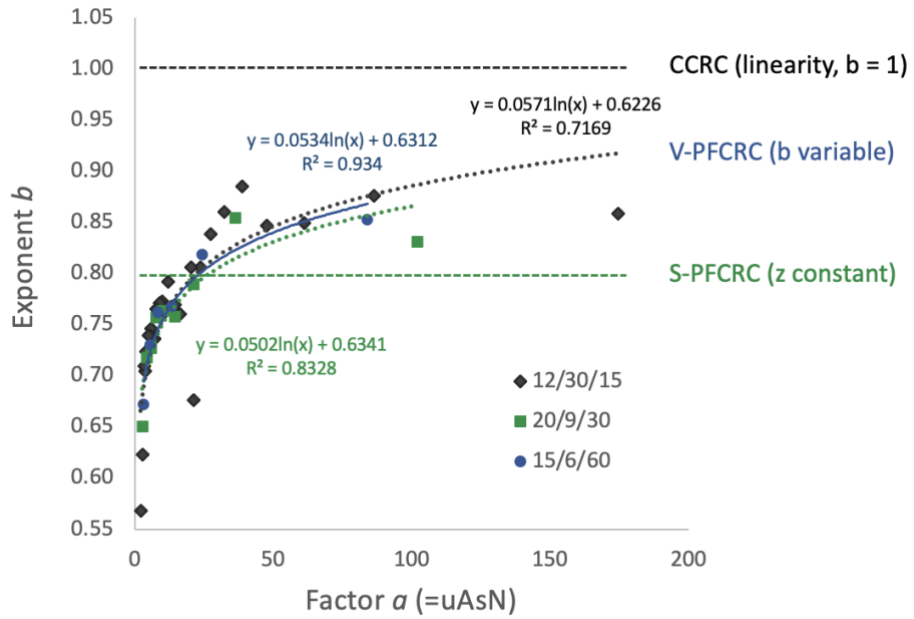


Figure 3: Log-Relationship between *a* (= uAs_N) and exponent *b*. Log regression curves were calculated for three different sample apportionments using 12, 20, or 15 uCR ranges subdivided into 30-, 9- or 6-tiles of uAs_{UC} yielding 360, 180, or 90 rectangles containing 15, 30, or 60 samples each. For validation, the log relation between *a* and *b* was investigated by two alternative calculations based on 12 or 20 uCR ranges and 30 or 9 uAs_{UC} levels comprising 15 or 30 samples per rectangle. These alternative regression curves (Figure 3, dotted green and black lines; squares, diamonds) widely overlapped with the original one (blue circles, straight line), indicating systematic fluctuations due to arbitrary sample apportionment to be insignificant. Alternatively, the coefficient *a* could be exponentially expressed by b in the form $a = 0.0001 \times \exp(14.3 \times b)$. The stability of this curve ($R^2 = 0.63$), though, was far lower than the log expression of *b* by *a* ($R^2 = 0.93$). In contrast to the uAs-dependent exponent b in V-PFCRC, both the z of S-PFCRC (green dashed line) and b=1 in CCRC (black dashed line) are invariable. The improvement of CCRC by S-PFCRC is due to z = 0.8 approximating the mean of variable b in V-PFCRC. The greater the deviations of uAs from mean values, the more z in S-PFCRC will differ from *b* in V-PFCRC. The resulting asymmetrical correction is illustrated in App.8b.

Appendix 9a illustrates the direct proportionality of coefficient *a* to uAs_{UC} for 8 different uCR concentration ranges. Congruent with both the linearity between uAs_N and *a* and the positive log relationship between *a* (uAs_N) and *b* (Fig.3), also the dependence of exponent *b* on uAs_{UC} was found to be of positive logarithmic nature (App.9b). With the apparent increase of exponent *b* as a function of uAs_{UC}, the relationship between uCR and uAs_{UC} will approximate direct proportionality (*b* = 1) in high uAs_{UC} samples.

In summary, higher TWuAs loads, represented either by *a* (Fig.3, App.9a) or by uAs_{UC} (App.9b), appear to increase the fractional excretion of uAs and thus shift exponent b closer to the value of 1, implying more constant uAs_{UC}/uCR-ratios and lower magnitudes of CCRE (Figs.4a-c).

3.3. Algebraic standardization of uAs_{UC} to 1g/L uCR

The challenge was finding a corrective equation that eliminated any bias of uAs_C from uCR by compensating for the influences of both uCR and the magnitude of renal TWuAs-load, incurring variations of TWuAs/uCR ratios (reflected by $b \neq 1$). The first resolution step was to express *b* by a function of uCR and uAs_{UC} while eliminating *a*:

Solving (1) for *a*:

(1') $a = uAs_{UC}/uCR^{b}$

Elimination of *a* by the junction of (1') and (5):

(6) $uAs_{UC}/uCR^{b} = exp\ (b/c - d/c)\ (= a)$

Solving (6) for *b*:

(7) $b = [Ln(uAs_{UC}) + d/c]/[Ln(uCR) + 1/c]$

The final corrective equation is generated by plugging (7) in (3):

(8) $uAs_{N} = uAs_{UC} \times uCR^{-[LN\ (uAsUC) + d/c]/[Ln(uCR) + 1/c]}$

3.4. Validating standardization

The TWuAs-specific form of Equation 8, $uAs_N = uAs_{UC} \times uCR^{-[Ln(uAsUC) + 11.82]/[Ln(uCR) + 18.73]}$, was applied to the 5400 uAs_{UC} samples forming the basis of Equation 8 (Fig.2b/d, App.8). For validation of CRCE elimination, uAs_N-stratified PFRA was performed for the uAs_N sextiles analog to uAs_{UC} (Fig.2b) and uAs_C (Fig.2c). All sextile curves of uAs_N show approximate parallelism to the x-axis with *b* in uAs_N (0.003 – 0.028) far closer to the desired value of zero than in uAs_C (-0.151 – -0.318). Residual dependencies on uCR were found negligible for uAs_N (R^2 9E-06 – 0.04) compared to uAs_C (R^2 0.42 – 0.89). The extensive mitigating effect of standardization on CRCE in aggregate data was confirmed in separate analyses of both sexes and seven different age groups (App.7, Part II).

Albeit significant, the absolute differences in means of the three result modes are modest (Tbl.2). In contrast, much broader variations exist in the modal values, adding evidence for higher corrective disparities in lower TWuAs samples (Tbl.2, Figs.4b/c).

The quantitative effects of standardization are illustrated by deviations of uAs_C from uAs_N (App.5). After restriction to 0.3 – 3 g/L uCR, there remains an average % deviation of uAs_C from uAs_N by 11.9 % ± 10.1 (0 – 272%), compared to 16.8% ± 16.4 in the complete, non-uCR restricted dataset. After uCR restriction to 0.3 – 3 g/L, 17% of uAS_C values (unrestricted 28.6%) still deviate from uAs_N by more than 20% and 4.7% (unrestricted 15.6%) by more than 30%. In accordance with the derivation of Eq.8, the percentage deviations of uAs_C from uAs_N strongly depend on both uAs_{UC} and uCR (Figs.4a-c). Since uCR and uAs_{UC} are both confounded primarily by hydration status and possibly by methylation capacity or other potential metabolic factors, they cannot be considered independent variables influencing %CRCE. To better estimate the individual contributions of uCR and uAs_{UC}, multiple linear regression analysis was pragmatically performed with %CRCE as the dependent variable and uCR-adjusted uAs_{UC} percentile rank (Fig.4c) besides log uCR as independent variables (Apps.6,10a). Standardized β coefficients were determined -0.96 for uCR and -0.11 for uAs_{UC}, suggesting the overall influence of uCR on %CRCE to be substantially stronger than that of uAs_{UC}, even though uAs_{UC}, particularly at lower percentile ranks, seems to noticeably influence CRCE (Figs.4b/c).

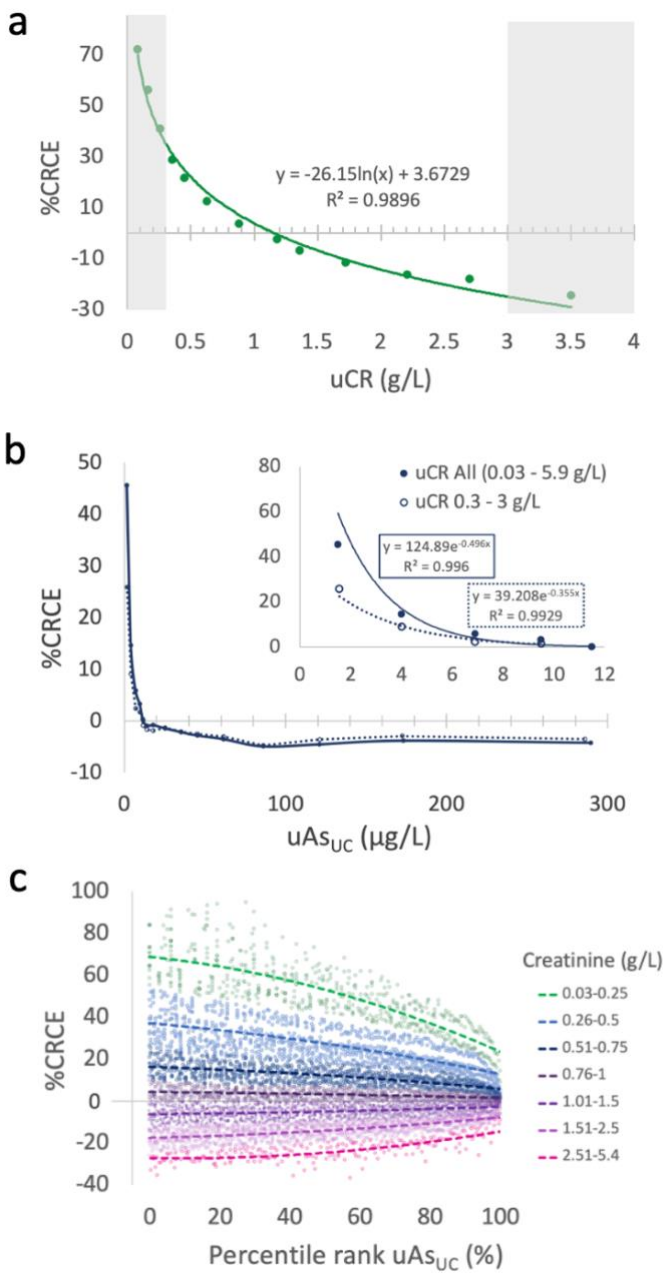


Figure 4a: Influence of uCR on %CRCE.
The percentual CRCE (%CRCE) of all (n=5599) samples < 500 µg/L uAs_{UC} was calculated using the formula: $\%CRCE = 100 \times (uAs_C - uAs_N)/uAs_N$. The dataset was subdivided into 13 uCR ranges, and the effects of uCR on %CRCE were analyzed by log regression of means of each uCR range. Grey-shaded areas represent conventionally restricted uCR ranges.

Figure 4b: Influence of absolute uAs_{UC} on %CRCE. All n = 5599 samples were subdivided into 15 uAs_{UC} ranges either without (dots, straight lines, n=5599) or with uCR restriction to 0.3 – 3 g/L (circles, dotted lines, n=4781). The mean values of each range are represented by dots (all uCR) or circles (0.3 – 3 g/L uCR). Stronger negative associations and potential contributions of uAs_{UC} to %CRCE are mainly expected in low uAs_{UC} < 5 µg/L. This effect, though, is strongly confounded by urinary dilution (uCR), as accounted for in Fig.4c. Above levels of 12 µg/L, the mean deviations changed the sign from positive to negative. The inlay depicts an enlarged section of the lower uAs_{UC}-range.

Figure 4c: Influence of relative uAs_{UC} on %CRCE. To accommodate the positive confounding effect of uCR on uAs_{UC}, all n=5599 values were partitioned into 20 equinumerous uCR ranges of 280 samples each (except for 279 values in the highest range). All subsets were percentile-ranked separately, and each uAs_{UC} value was assigned to the percentile of its subset and plotted against %CRCE. Percentile values more adequately represent renal As-loads as they are largely adjusted for hydration. The effect of uCR on the relative influence of uAs_{UC} is illustrated by 7 differently colored uCR areas and their corresponding polynomial regression lines. In all 7 uCR ranges, the magnitude of %CRCE decreases with increasing percentile rank. All 7 lines converge toward the %CRCE value of 0 with increasing percentile rank irrespective of the algebraic sign of %CRCE, which is positive in the case of uCR<1 and negative in the case of uCR>1.

	Correction mode	Mean ± SE	Median	Modal	Minimum	Maximum
Linear	uAs_{UC} (µg/L)	22.0 ± 0.54	8.9	1.5	1	484
	uAs_C (µg/g CR)	24.7 ± 0.56	11.3	10	0.8	741
	uAs_N (µg/1g CR)	23.3 ± 0.54	10.4	4.3	0.9	724
Log (LN)	uAs_{UC} (µg/L)	2.3 ± 0.016	2.19	1.25	-0.04	6.18
	uAs_C (µg/g CR)	2.6 ± 0.014	2.42	2.48	-0.21	6.61
	uAs_N (µg/1g CR)	2.5 ± 0.014	2.34	2.2	-0.06	6.59

Table 2: Quantitative differences in uAsUC, uAsC, and uAsN. Basic descriptive statistics for all n = 5400 samples of the age band 14 – 82 years, < 500 µg/L uAs_{UC}, and uCR < 4.5 g/L, constituting Eq.8. Arsenic values were logarithmized to approximate normal distribution. The means of each correction mode (uAsUC vs. uAsC, uAsUC vs. uAsN, and uAsC vs. uAsN) all differed significantly within the respective log and linear datasets ($p > 0.001$, paired 2-tailed student t-test).

4. Discussion

4.1. Power-functional creatinine correction (PFCRC)

Power-functional approaches to dilution adjustment generally partition the graphic relation between the uncorrected analyte (here, uAs_{UC}) and uCR into a fan of curved lines intersecting the zero point, in contrast to the erroneously assumed straight lines in CCRC. All values on a curve are attributed to the same standardized value. The curvature of the lines as a function of uCR and uAs_{UC} is determined by the disproportional dependence of the respective renal net excretion rates on UFR. Successful attempts to compensate CRCE by simple PFCRC have been well documented.[3,5,18] In S-PFCRC, the dependence of the correction on UFR is represented by the constant exponent z, which can be either empirically determined by resource-intense measurements of biomarker-specific analyte-UFR slopes or, alternatively, by adopting the somewhat asymmetric numerical approach potentially involving the observed unwanted residual dependencies at both edges of corrected data sets (App.8).[3,20] The herein-developed progression of S-PFCRC to an analyte-stratified variable PFCRC (V-PFCRC) additionally compensates for the independent influence of analyte level on exponent z, denominated '*b*' in V-PFCRC (Fig.3). This proofed necessary, as uAs_{UC}, being confounded with uCR by UFR, seems to make a minor independent contribution to %CRCE, especially at lower uCR-adjusted percentile-ranks of uAs_{UC} (Fig.4b/c, App.6). In contrast to S-PFCRC, V-PFCRC does not only attain overall minimization of residual correlations between uAs and uCR in uAs_N-data (R^2 and exponent b close to zero comparable to S-PFCRC) but additionally proved to reliably remove residual dependencies in the marginal concentration ranges as seen in unstratified data sets corrected by S-PFCRC (Fig.1d, App.8b). In V-PFCRC the degree of disproportion in UFR dependence is accounted for by two coefficients (*c* and *d*) modifying the standardizing function (Eq.8) in an analyte-specific manner (Figs.3,4). The coefficients *c* and *d* can be easily numerically identified by a larger dataset that includes uCR and analyte concentrations while avoiding more elaborate UFR measurements.

The accuracy of standardizing functions principally depends on the representativity and number of samples available for their establishment. In this study, 5400 samples proved to be enough to attain a robust log relationship between coefficient *a* and exponent *b*. The derived coefficients *c* and *d* proved stable, enabling adequate analyte-specific dilution adjustment (Fig.3). Approximate estimations on fewer samples (≥ 1,000)

will most likely already allow for considerable improvement in results, especially for less complex markers.

Standardization to 1g/L uCR by V-PFCRC effectively eliminated CRCE over broad urine dilution- and uAs_{UC}-ranges not only in aggregated (Fig.2d, App.8) but also in age- and sex-disaggregated analyses (App.7) as verified by simple PFRA. The quite uniform mitigation of CRCE in all subgroups can be seen as evidence of its principal induction by a common fundamental physiological misconception with only minor quantitative modifications by other biasing factors like sex, age, or diet, which may nevertheless induce considerable bias by other mechanisms than dilution (Part II).

Apart from the asymmetrical net excretions, the generally higher proportions of poorly reabsorbed, methylated As-species like AsB and DMA from predominant marine sources in high TWuAs-samples could further contribute to the effect (Sec.1.4).[31,42,45] Stronger similarity in tubular behavior of methylated As-species to uCR should shift TWuAs/uCR relations closer to linearity, thus minimizing CRCE. In the present study, the imprecision due to varying As-sources and the resulting differences in TWuAs composition and, thus, CRCE is understood to be largely compensated for by the large number of samples. Future studies should aim to determine corrective V-PFCRC formulas for all of the individual TWuAs components separately to further control any TWuAs-concentration-dependent composition bias.

4.2. Standardization of other biomarkers and dilution adjustments

Apart from the more complex exemplary parameter TWuAs used herein, the general form of Eq.8, $uAs_N = uAs_{UC} \times uCR^{-[LN\,(uAsUC) + d/c]\,/\,[Ln\,(uCR) + 1/c]}$, can be applied to all other spot urinary analytes by determining coefficients *c* and *d* for any urinary biomarker as the analyte-specific basis for adequate dilution adjustment.

The inverse correlation of %CRCE and TWuAs (Fig.4c) suggests CCRC incurring the strongest misestimates of real exposure in urinary analytes present in only minute amounts. Standardizing these low-concentrated analytes to 1 g/L uCR thus appears to be a prerequisite for their precision.

The reliabilities and validities of dilution adjustments by SG or osmolality, both of which are considered less susceptible to interference than uCR, are compromised in a similar way to CCRC.[3] The negligence of the asynchronous hydration-dependent tubular handling of a biomarker on the one hand and of the urinary electrolytes and osmolytes such as urea, uCR, and uric acid that define the SG of urine on the other will induce related correction errors in both methods, albeit quantitatively weaker than in CCRC.[3,11,18] Previous promising results for S-PFCRC in SG/adjustments and the linear relationship between CCRC and specific weight correction results support this assumption.[3,5,11,54] It appears consistent that V-PFCRC, like S-PFCRC, can provide the algebraic basis for adequately compensating dilution adjustment errors by uCR, SG, and osmolality.

4.3. Benefits and limitations of standardization by V-PFCRC

S-PFCRC and V-PFCRC account for the physiological hydration-dependent differences in analyte/uCR net excretion ratios. The stable results over a broad urinary dilution (or UFR) range will increase the reliability and comparability of spot urinary biomarkers and permit the omission of symptomatic uCR-range restrictions to 0.3 – 3 g/L as a source of considerable sample rejections rates. Avoiding cumbersome pre-analytical questioning and exam, as required for more elaborate MLRM, V-PFCRC and analog power-functional modifications of SG and osmolality, are comparably resource and cost-saving methods. They can be readily implemented in clinical and epidemiological practice by integrating the corrective functions in lab software. Its retrospective application to existing datasets

will enable a more adequate understanding of previously reported associations between urinary markers and proxies for UFR like uCR, osmolality, and SG.[31,32,45] That said, it should be noted that standardization primarily mitigates the systemic disproportional dependences of uCR- and analyte-excretions on UFR. It will thus strongly improve intra- and interindividual comparability and form a more rational basis for establishing normal ranges or toxic thresholds, which may still exhibit age-, sex-, diet-, ethnical-, cross-cultural or other biases due to hydration-independent toxicokinetic factors. These factors may additionally need to be compensated for, be it by differentiating reference ranges or in MLRM using PFCRC values as covariates (see Part II).[2,17]

A potential for compromise of V-PFCRC lies in the differential tubular reabsorption characteristics of various analyte components in more complex biomarkers. In TWuAs, this might become relevant in major deviations from common TWuAs composition, like in rare acute iAs-poisoning with increased tubular reabsorption rates seen in high TWuAs, usually featuring higher proportions of readily excreted oAs. High doses of iAs will entail lower exponents *b* and higher CCREs, being under-compensated by normalization and leading to overestimated As-loads in dilute and underestimations in concentrated urines (Fig.1). However, this case appears to be of only minor clinical relevance, since standardized results, due to its lower exponent *b,* will still be more precise than CCRC and excessively high TWuAs samples will be subjected to speciation anyway. Apart from acute poisoning, major regional differences in alimentation, particularly in seafood consumption or IV-fluid-administration shortly before urine sampling, may necessitate adjustments of standardization functions.

5. Abbreviations

As	Arsenic
AsB	Arsenobetaine
ATSDR	Agency for Toxic Substances and Disease Registry
CCRC	Conventional Creatinine Correction
CR	Creatinine
CRCE	Creatinine Correction Error
%CRCE	% deviation of uAs_C from uAs_N ($\%CRCE = 100 \times (uAs_C - uAs_N)/uAs_N$)
DMA	Dimethylarsinic Acid
EPA	Environmental Protection Agency (USA)
iAs	Inorganic Arsenic
ICP-MS	Inductively Coupled Plasma Mass Spectrometry
ICR	Interquartile Range
MLRM	Multiple Linear Regression Models
MMA	Monomethylarsonic Acid
oAs	Organic Arsenic
PO_4^-	Phosphate
PFCRC	Power Functional Creatinine Correction
PFRA	Power Functional Regression Analysis
SAME	S-Adenosyl-Methionine
S-PFCRC	Simple Power Functional Creatinine Correction
sP	Serum Phosphate
TWuAs	Total Weight Urinary Arsenic
uAs	Urinary Arsenic
uAs_{UC}	Uncorrected Urinary Arsenic
uAs_C	Conventionally Creatinine-corrected Urinary Arsenic
uAs_N	Standardized Urinary Arsenic, normalized to 1 g/L uCR
uCR	Urinary Creatinine
UFR	Urinary Flow Rate
V-PFCRC	Variable Power Functional Creatinine Correction

6. Acknowledgements

Dr. Katrin Huesker, department manager of Special Immunology, IMD Berlin, is thanked for providing anonymized spot urine data.
Mischa Katz GE is thanked for his timesaving help in data preparation and programming.

7. Author contribution

The idea, basic concept, procurement, preparation and evaluation of data, literature research, developing corrective functions, scrapping erroneous corrective functions, graphs, tables, labels, writing, and submitting the paper.

8. Conflict of Interest

No conflict of interest is known to me with persons, companies, institutions, myself, or any other entity known or unknown to me.

9. Ethics

The study was exempt from ethical approval by the Ethic Committee of Northwest- and Central Switzerland (EKNZ), 4056 Basel, on basis of retrospective analysis of anonymized data without clinical information other than urinary creatinine/TWuAs, age in years, sex, and date of exam. The purpose was methodological quality improvement only. No pharmacological or other medical intervention, no risk for physical or psychological harm, identifiability, or breach of confidentiality for any subject involved.

10. Funding

The study was not funded.

11. References

1) Santonen T, Schoeters G, Nordberg M. Chapter 9 - Biological monitoring of metals and biomarkers. Editor(s): Nordberg GF, Costa M. Handbook on the Toxicology of Metals (Fifth Edition), Vol 1, *Academic Press*, 2022: 217-35.
ISBN 9780128232927, https://doi.org/10.1016/B978-0-12-823292-7.00007-3.

2) O'Brien KM, Upson K, Cook NR, Weinberg CR. Environmental Chemicals in Urine and Blood: Improving Methods for Creatinine and Lipid Adjustment. *Environ Health Perspect.* 2016 Feb; **124(2):** 220-27. doi: 10.1289/ehp.1509693. Epub 2015 Jul 24. PMID: 26219104; PMCID: PMC4749084.

3) Middleton DRS, Watts MJ, Polya DA. A comparative assessment of dilution correction methods for spot urinary analyte concentrations in a UK population exposed to arsenic in drinking water. *Environ Int.* 2019 Sep; **130**: 104721. doi: 10.1016/j.envint.2019.03.069. Epub 2019 Jun 14. PMID: 31207477; PMCID: PMC6686075.

4) Akerstrom M, Barregard L, Lundh T, Sallsten G. The relationship between cadmium in kidney and cadmium in urine and blood in an environmentally exposed population. *Toxicol Appl Pharmacol.* 2013 May 1; **268(3):** 286-93. doi: 10.1016/j.taap.2013.02.009. Epub 2013 Feb 27. PMID: 23454399.

5) Araki S, Sata F, Murata K. Adjustment for urinary flow rate: an improved approach to biological monitoring. *Int. Arch. Occup. Environ. Health.* 1990; **62(6)**: 471–77.

6) Sauvé JF, Lévesque M, Huard M, et al. Creatinine and specific gravity standardization in biological monitoring of occupational exposures. *J Occup Environ Hyg*. 2015; **12(2)**: 123-29. doi: 10.1080/15459624.2014.955179. PMID: 25192246.

7) Boeniger MF, Lowry LK, Rosenberg J. Interpretation of urine results used to assess chemical exposure with emphasis on creatinine adjustments: a review. *Am Ind Hyg Assoc J.* 1993 Oct; **54(10)**: 615-27. doi: 10.1080/15298669391355134. PMID: 8237794.

8) Suwazono Y, Akesson A, Alfvén T, Järup L, Vahter M. Creatinine versus specific gravity-adjusted urinary cadmium concentrations. *Biomarkers*. 2005 Mar-Jun; **10(2-3)**: 117-26. doi: 10.1080/13547500500159001. PMID: 16076727.

9) Alessio L, Berlin A, Dell'Orto A, Toffoletto F, Ghezzi I. Reliability of urinary creatinine as a parameter used to adjust values of urinary biological indicators. *Int Arch Occup Environ Health*. 1985; **55(2)**: 99-106. doi: 10.1007/BF00378371. PMID: 3988361.

10) Cocker J, Mason HJ, Warren ND, Cotton RJ. Creatinine adjustment of biological monitoring results. *Occup Med (Lond).* 2011 Aug; **61(5)**: 349-53.
doi: 10.1093/occmed/kqr084. PMID: 21831823.

11) Muscat JE, Liu A, Richie JP Jr. A comparison of creatinine vs. specific gravity to correct for urinary dilution of cotinine. *Biomarkers*. 2011 May; **16(3)**: 206-11.
doi: 10.3109/1354750X.2010.538084. Epub 2011 Feb 3. PMID: 21288164; PMCID: PMC3631104.

12) Allen R, Mage D, Gondy G, Christensen C, Barr D, Needham L. The use of a creatinine correction for reporting children's urinary pesticide concentrations. *Epidemiology*. July 2004; **15(4)**: 69-70.

13) Cole SR, Platt RW, Schisterman EF, et al. Illustrating bias due to conditioning on a collider. *Int J Epidemiol*. 2010 Apr; **39(2)**: 417-20. doi: 10.1093/ije/dyp334. Epub 2009 Nov 19. PMID: 19926667; PMCID: PMC2846442.

14) Gamble MV, Hall MN. Relationship of creatinine and nutrition with arsenic metabolism. *Environ Health Perspect.* 2012 Apr; **120(4):** A145-46. doi: 10.1289/ehp.1104807. PMID: 22469551; PMCID: PMC3339472.

15) Basu A, Mitra S, Chung J, Guha Mazumder DN, Ghosh N, Kalman D, von Ehrenstein OS, Steinmaus C, Liaw J, Smith AH. Creatinine, diet, micronutrients, and arsenic methylation in West Bengal, India. *Environ Health Perspect*. 2011 Sep; **119(9)**: 1308-13. doi: 10.1289/ehp.1003393. Epub 2011 Jun 7. PMID: 21652291; PMCID: PMC3230402.

16) Anderson D, Lydic R. Ratio data and the quantification of drug effects. *Biobehavioral Reviews*. 1977; **1(1)**: 55-7. doi: 10.1016/0147-7552(77)90041-9. ISSN 0147-7552.

17) Barr DB, Wilder LC, Caudill SP, Gonzalez AJ, Needham LL, Pirkle JL. Urinary creatinine concentrations in the U.S. population: implications for urinary biologic monitoring measurements. *Environ Health Perspect.* 2005 Feb; **113(2)**: 192-200. doi: 10.1289/ehp.7337. PMID: 15687057; PMCID: PMC1277864.

18) Middleton DR, Watts MJ, Lark RM, Milne CJ, Polya DA. Assessing urinary flow rate, creatinine, osmolality and other hydration adjustment methods for urinary biomonitoring using NHANES arsenic, iodine, lead and cadmium data. *Environ Health*. 2016 Jun10; **15(1)**: 68. doi: 10.1186/s12940-016-0152-x. PMID: 27286873; PMCID: PMC4902931.

19) Middleton D. Arsenic Research and Global Sustainability: Proceedings of the Sixth International Congress on Arsenic in the Environment (As2016), June 19-23, 2016, Stockholm, Sweden. *CRC Press*; 2016. A modified creatinine adjustment method to improve urinary biomonitoring of exposure to arsenic in drinking water.

20) Vij HS, Howell S. Improving the specific gravity adjustment method for assessing urinary concentrations of toxic substances. *Am Ind Hyg Assoc J*. 1998 Jun; **59(6)**: 375-80. doi: 10.1080/15428119891010622. PMID: 9670467.

21) *Agency for Toxic Substances and Disease Registry (ATSDR),* 2019 Substance Priority List, https://www.atsdr.cdc.gov/spl/index.html#2019spl.

22) Hall AH. Chronic arsenic poisoning. *Toxicol Lett*. 2002 Mar 10; **128(1-3)**: 69-72. doi: 10.1016/s0378-4274(01)00534-3. PMID: 11869818.

23) Fowler BA, Selene CH, Chou J, Jones RL, Costa M, Chen CJ. Chapter 3 - Arsenic. Editor(s): Gunnar F. Nordberg, Max Costa. Handbook on the Toxicology of Metals (Fifth Edition), Vol 2, *Academic Press*, 2022; 41-89, ISBN 9780128229460, https://doi.org/10.1016/B978-0-12-822946-0.00037-4.

24) Apostoli P, Bartoli D, Alessio L, Buchet JP. Biological monitoring of occupational exposure to inorganic arsenic. *Occup Environ Med.* 1999 Dec; **56(12)**: 825-32. doi: 10.1136/oem.56.12.825. PMID: 10658539; PMCID: PMC1757692.

25) Bertolero F, Marafante E, Rade JE, Pietra R, Sabbioni E. Biotransformation and intracellular binding of arsenic in tissues of rabbits after intraperitoneal administration of 74As labeled arsenite. *Toxicology*. 1981; **20(1)**: 35-44. doi: 10.1016/0300-483x(81)90103-7. PMID: 7268789.

26) *WHO.* Background document for development of WHO Guidelines for Drinking-water Quality (2003). Arsenic in Drinking-water. https://apps.who.int/iris/handle/10665/75375.

27) *ATSDR.* Arsenic Toxicity (2011). What Are the Standards and Regulations for Arsenic Exposure? https://www.atsdr.cdc.gov/csem/arsenic/docs/arsenic.pdf

28) *US Food and Drug Administration (FDA).* Arsenic in rice and rice products risk assessment report (2016). https://www.regulations.gov/docket/FDA-2016-D-1099.

29) Buchet JP, Lauwerys R, Roels H. Comparison of the urinary excretion of arsenic metabolites after a single oral dose of sodium arsenite, monomethylarsonate, or dimethylarsinate in man. *Int Arch Occup Environ Health*. 1981; **48(1)**: 71-9. doi: 10.1007/BF00405933. PMID: 6894292.

30) Vahter M. Genetic polymorphism in the biotransformation of inorganic arsenic and its role in toxicity. *Toxicol Lett*. 2000 Mar 15; **112-113**: 209-17. doi: 10.1016/s0378-4274(99)00271-4. PMID: 10720733.

31) Peters BA, Hall MN, Liu X, et al. Creatinine, arsenic metabolism, and renal function in an arsenic-exposed population in Bangladesh. *PLoS One*. 2014 Dec 1; **9(12)**: e113760. doi: 10.1371/journal.pone.0113760. PMID: 25438247; PMCID: PMC4249915.

32) Kile ML, Hoffman E, Hsueh YM, Afroz S, et al. Variability in biomarkers of arsenic exposure and metabolism in adults over time. *Environ Health Perspect*. 2009 Mar; **117(3)**: 455-60. doi: 10.1289/ehp.11251. Epub 2008 Nov 19. PMID: 19337522; PMCID: PMC2661917.

33) Ginsburg JM, Lotspeich WD. Interrelations of arsenate and phosphate transport in the dog kidney. *Am J Physiol.* 1963 Oct; **205**: 707-14. doi: 10.1152/ajplegacy.1963.205.4.707. PMID: 14060809.

34) Peters BA, Hall MN, Liu X, et al. Renal function is associated with indicators of arsenic methylation capacity in Bangladeshi adults. *Environ Res*. 2015 Nov; **143(Pt A)**: 123-30. doi: 10.1016/j.envres.2015.10.001. Epub 2015 Oct 19. PMID: 26476787; PMCID: PMC4740972.

35) Villa-Bellosta R, Sorribas V. Different effects of arsenate and phosphonoformate on P(i) transport adaptation in opossum kidney cells. *Am J Physiol Cell Physiol*. 2009 Sep; **297(3)**: C516-25. doi: 10.1152/ajpcell.00186.2009. Epub 2009 Jun 24. PMID: 19553564.

36) Hopenhayn-Rich C, Biggs ML, Smith AH, Kalman DA, Moore LE. Methylation study of a population environmentally exposed to arsenic in drinking water. *Environ Health Perspect*. 1996 Jun; **104(6)**: 620-28. doi: 10.1289/ehp.96104620. PMID: 8793350; PMCID: PMC1469390.

37) Orr SE, Bridges CC. Chronic Kidney Disease and Exposure to Nephrotoxic Metals. *International Journal of Molecular Sciences*. 2017; **18(5)**: 1039. https://doi.org/10.3390/ijms18051039

38) Tsuda A, Ishimura E, Machiba Y, et al. Increased Glomerular Hydrostatic Pressure is Associated with Tubular Creatinine Reabsorption in Healthy Subjects. *Kidney Blood Press Res* 2020; **45**: 996-1008. DOI: 10.1159/000510838

39) Shemesh O, Golbetz H, Kriss JP, Myers BD. Limitations of creatinine as a filtration marker in glomerulopathic patients. *Kidney Int*. 1985 Nov; **28(5)**: 830-38. doi: 10.1038/ki.1985.205. PMID: 2418254.

40) Ciarimboli G, Lancaster CS, Schlatter E, et al. Proximal tubular secretion of creatinine by organic cation transporter OCT2 in cancer patients. *Clin Cancer Res*. 2012 Feb 15; **18(4)**: 1101-08. doi: 10.1158/1078-0432.CCR-11-2503. Epub 2012 Jan 5. PMID: 22223530; PMCID: PMC3288323.

41) Molin M, Ulven SM, Meltzer HM, Alexander J. Arsenic in the human food chain, biotransformation, and toxicology - Review focusing on seafood arsenic. *J Trace Elem Med Biol*. 2015; **31**: 249-59. doi: 10.1016/j.jtemb.2015.01.010. Epub 2015 Jan 28. PMID: 25666158.

42) Taylor V, Goodale B, Raab B, et al. Human exposure to organic arsenic species from seafood. *Science of The Total Environment.* 2017; **580**: 266-82, ISSN 0048-9697, doi: org/10.1016/j.scitotenv.2016.12.113.

43) Petrick JS, Ayala-Fierro F, Cullen WR, Carter DE, Vasken AH. Monomethylarsonous acid (MMA(III)) is more toxic than arsenite in Chang human hepatocytes. *Toxicol Appl Pharmacol*. 2000 Mar 1; **163(2)**: 203-07. doi: 10.1006/taap.1999.8872. PMID: 10698679.

44) Vahter M, Concha G. Role of metabolism in arsenic toxicity. *Pharmacol Toxicol*. 2001 Jul; **89(1)**: 1-5. doi: 10.1034/j.1600-0773.2001.d01-128.x. PMID: 11484904.

45) Gilbert-Diamond D, Li Z, Perry AE, et al. A population-based case-control study of urinary arsenic species and squamous cell carcinoma in New Hampshire, USA. *Environmental Health Perspectives*. 2013 Oct; **121(10)**: 1154-60. doi: 10.1289/ehp.1206178. PMID: 23872349; PMCID: PMC3801199.

46) Choi JW, Song YC, Cheong NY, et al. Concentrations of blood and urinary arsenic species and their characteristics in general Korean population. *Environ Res*. 2022 Nov; **214(Pt 2)**: 113846. doi: 10.1016/j.envres.2022.113846. Epub 2022 Jul 9. PMID: 35820651.

47) Hata A, Yamanaka K, Habib MA, Endo Y, Fujitani N, Endo G. Arsenic speciation analysis of urine samples from individuals living in an arsenic-contaminated area in Bangladesh. *Environ Health Prev Med*. 2012 May; **17(3)**:235-45. doi: 10.1007/s12199-011-0247-5. Epub 2011 Nov 3. PMID: 22048870; PMCID: PMC3348244.

48) Reid MS, Hoy KS, Schofield JRM et al. Arsenic speciation analysis: A review with an emphasis on chromatographic separations. *TrAC Trends in Analytical Chemistry*. 2020; **123**: 115770; ISSN 0165-9936, https://doi.org/10.1016/j.trac.2019.115770.

49) Ginsburg JM. The renal mechanism for excretion and transformation of arsenic in the dog. *Am J Physiol*. 1965 May; **208**: 832-40. doi: 10.1152/ajplegacy.1965.208.5.832. PMID: 14286849.

50) Stockand JD. Vasopressin regulation of renal sodium excretion. *Kidney Int*. 2010 Nov; **78(9)**: 849-56. doi: 10.1038/ki.2010.276. Epub 2010 Aug 25. PMID: 20736986.

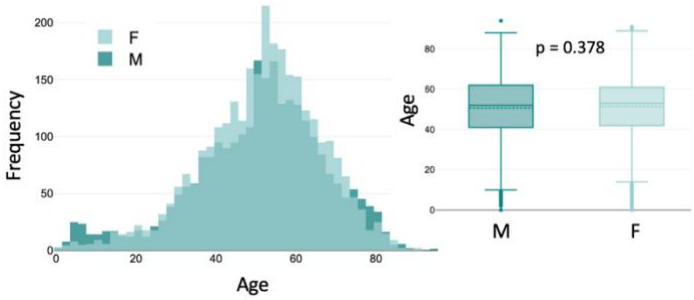
51) Perinpam M, Ware EB, Smith JA, Turner ST, Kardia SL, Lieske JC. Key influence of sex on urine volume and osmolality. *Biol Sex Differ*. 2016 Feb 9; **7**: 12. doi: 10.1186/s13293-016-0063-0. PMID: 26865949; PMCID: PMC4748596.

52) Sabolić I, Asif AR, Budach WE, Wanke C, Bahn A, Burckhardt G. Gender differences in kidney function. *Pflugers Arch.* 2007 Dec;**455(3)**:397-429. doi: 10.1007/s00424-007-0308-1. Epub 2007 Jul 19. PMID: 17638010.

53) Perucca J, Bouby N, Valeix P, Bankir L. Sex difference in urine concentration across differing ages, sodium intake, and level of kidney disease. *Am J Physiol Regul Integr Comp Physiol*. 2007 Feb;**292(2)**: R700-5. doi: 10.1152/ajpregu.00500.2006. Epub 2006 Sep 21. PMID: 16990487.

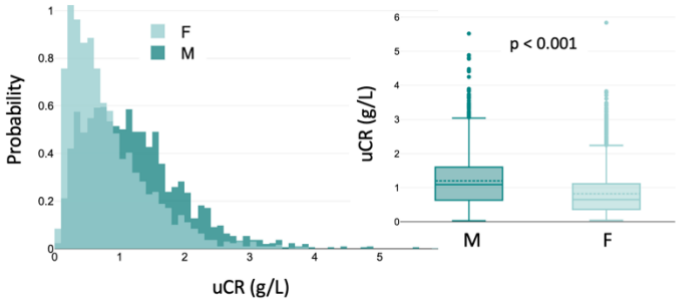
54) Ozdemir S, Sears CG, Harrington JM, et al. Relationship between Urine Creatinine and Urine Osmolality in Spot Samples among Men and Women in the Danish Diet Cancer and Health Cohort. *Toxics.* 2021 Nov 1;**9(11)**:282. doi: 10.3390/toxics9110282. PMID: 34822673; PMCID: PMC8625939.

11. Supplementary Material/Appendices



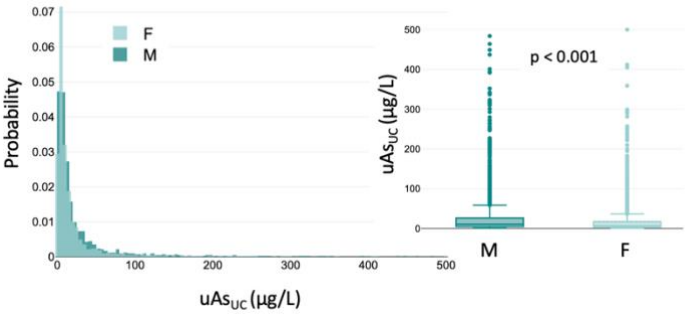
Age Structure		
Sex	F	M
Number of valid values	2971	2628
Mean ± Std.	51.48 ± 14.69	50.69 ± 16.65
Median	53	52
Modal	52	51
Minimum	0	0
Maximum	91	94

Appendix 1: Age structure of study population by sex. Female and male study groups were similar with respect to age structure. n=5599 samples ≤ 500 mcg/L uAs_{UC} (from a total of 5608 samples) were analyzed, and 9 outliers with uAs_{UC} >500 (548 – 5763 mcg/L) were precluded. A Mann-Whitney U Test did reveal no significant age difference between the female and male groups.



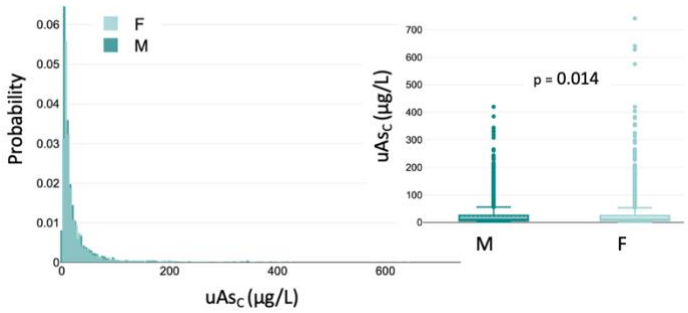
Urinary Creatinine (uCR)		
Sex	F	M
Number of valid values	2971	2628
Mean ± Std.	0.82 ± 0.62	1.2 ± 0.74
Median	0.65	1.09
Modal	0.32	1.17
Minimum	0.04	0.03
Maximum	5.84	5.52

Appendix 2: Urinary creatinine by sex. Female patients have significantly lower uCR than men ($p < 0.001$, Mann-Whitney U Test).



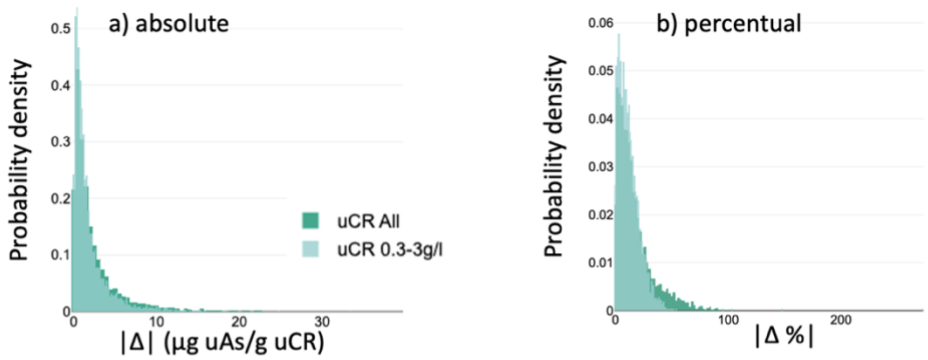
Uncorrected Urinary Arsenic (uAs_{UC})		
Sex	F	M
Number of valid values	2971	2628
Mean ± Std.	17.79 ± 33.76	26.64 ± 46.47
Median	7.3	10.9
Modal	1.5	3.5
Minimum	0.99	0.96
Maximum	500	484

Appendix 3: Uncorrected urinary Arsenic (uAs_{UC}) by sex. Female patients have significantly lower uAs than men ($p < 0.001$, Mann-Whitney U Test).



Conventionally uCR-corrected Urinary Arsenic (uAs_C)		
Sex	F	M
Number of valid values	2971	2628
Mean ± Std.	25.44 ± 45.47	23.89 ± 36.97
Median	11.5	10.88
Modal	10	11
Minimum	1.07	0.81
Maximum	741	420

Appendix 4: Conventionally CR-corrected urinary Arsenic (uAs) by sex. Female patients have significantly higher corrected uAs than men ($p = 0.014$, Mann-Whitney U Test).



a) Absolute deviation of uAs_C from uAs_N (µg uAs/g uCR)		
uCR range	0.3-3 g/L	All uCR
$\lvert\Delta\rvert$(µg/g $_{uCR}$); $\bar{x}$ ± SD	1.9 ± 2.5	2.7 ± 3.8
Minimum $\lvert\Delta\rvert$	0	0
Maximum $\lvert\Delta\rvert$	33.3	56.9

b) Percentual deviation of uAs_C from uAs_N (%)		
uCR range	0.3-3 g/L	All uCR
$\lvert\Delta\ \%\rvert$ ($\bar{x}$ ± SD)	11.9 ± 10.1	16.8 ± 16.4
Minimum $\lvert\Delta\ \%\rvert$	0	0
Maximum $\lvert\Delta\ \%\rvert$	271.8	271.8

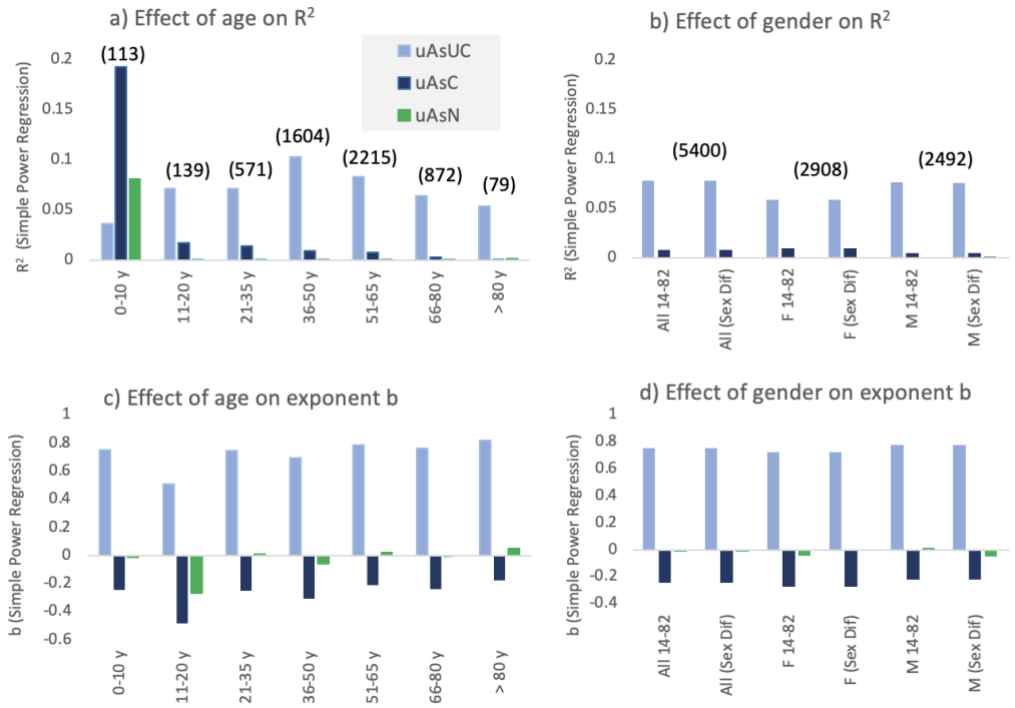
c) uAs_C values deviating from uAs_N by more than 20% or 30%				
uCR range	0.3-3 g/L uCR		All uCR	
n total	4781		5599	
Limit	20%	30%	20%	30%
n > Limit	814	223	1603	873
% n > Limit	17.0	4.7	28.6	15.6

Appendix 5: Distribution of absolute and percentual deviations of uAs_N from uAs_C. All samples with < 500 µg uAs_{UC}. a) Absolute differences were calculated for all by $\lvert\Delta\rvert = \lvert uAs_N - uAs_C\rvert$. Overlapping histograms are plotted for the full range of uCR (dark green, n = 5599) and after uCR-restriction adopting the usual exclusion limits, i.e., all samples within the range from 0.3 - 3 g/L uCR (light green, n = 4781). b) Percentual deviations were calculated by $\lvert\Delta\ \%\rvert = \lvert\Delta\rvert \times 100/uAs_{UC}$. They are either plotted as complete or uCR-restricted dataset analog to histogram a. c) Absolute numbers and percentages of samples with uAs_c deviating more than either 20% or 30% from uAs_N.

Multiple linear regression model	Unstandardized Coefficients	Standardized Coefficients				95% confidence interval for B	
	B	Beta	SE	t	p	Lower bound	Upper bound
Constant	5.36		0.15	35.78	<.001	5.07	5.66
LN uCR (g/L)	-26.23	-0.96	0.1	-275.17	<.001	-26.41	-26.04
Percentile uAs_{UC}	-0.08	-0.11	0	-31.41	<.001	-0.09	-0.08

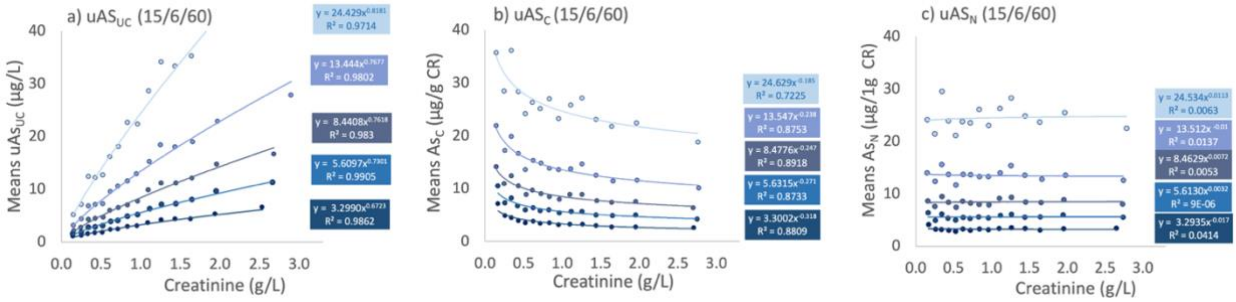
ANOVA / Model summary	df	F	p	R	R^2	SE of the estimate
	2	38,440.65	<.001	0.97	0.93	5.57

Appendix 6: Multilinear regression analysis of contributions of uCR and uAsUC to %CRCE. Significance was established by ANOVA. Standardized regression beta coefficients (β) were calculated for uAs_{UC} and uCR. Formally, linear multi regression is not quite appropriate here because %CRCE, as the dependent variable, is not unrelated to the independent variables uCR and uAs. Also, all three variables are not normally distributed. Despite these limitations, the beta coefficient, which is significantly higher for uCR in all 3 cases, signals a stronger influence of uCR than uAs_{UC} on %CRCE.

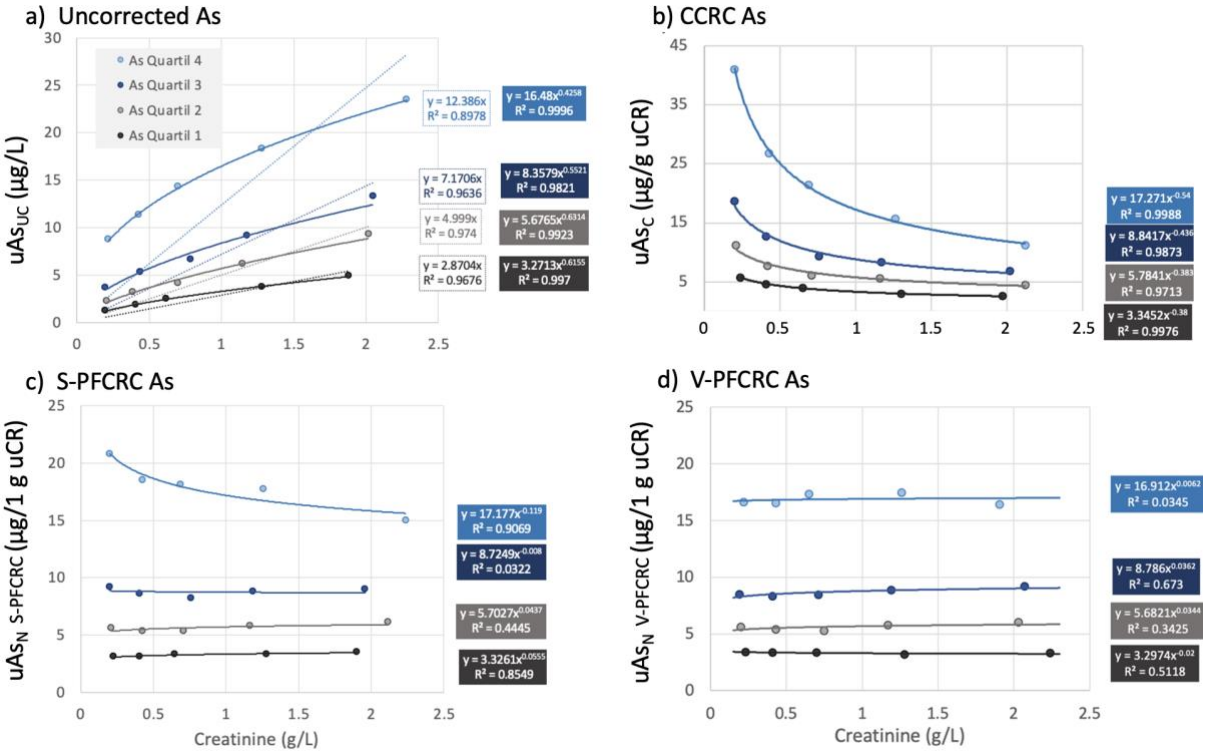


Appendix 7: Sex- and Age-differentiated efficacy of standardization. Simple power regression functions of the form $y = a \times x^b$ (x = uCR, y = uAs) and the corresponding R^2 values were calculated for various subsets of a total of n=5599 samples (all ages, $uAs_{UC} \leq 500$ µg/L) grouped by age (App.6a/c) and sex (App.6b/d). Exponent b and R^2 as measures for the dependency between uCR and uAs are plotted for uAs_{UC} (light blue bars), uAs_C (dark blue bars), and uAs_N (green bars). Numbers in brackets: sample numbers (n)/subset. Male

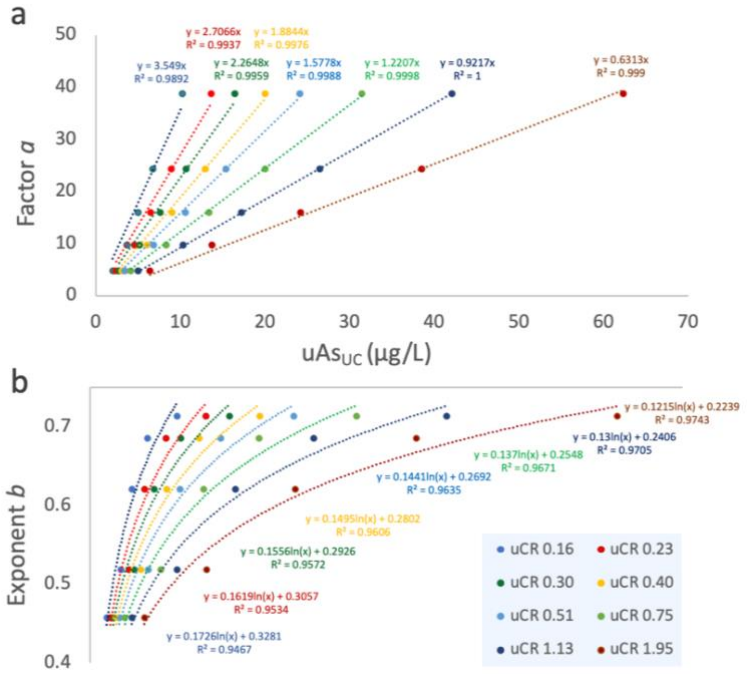
and female samples (App.6b/d) were normalized, with Eq.8 generated either by gender aggregated (Fig.2d) or by two separate, sex-differentiated analyses (indicated as Sex Dif).



Appendix 8a: Enlarged illustration of uCR/uAs-relations in sextiles 1-5 for uAsUC, uAsC, and uAsN.
Section of Fig. 2b-d with ordinate truncated to 40.

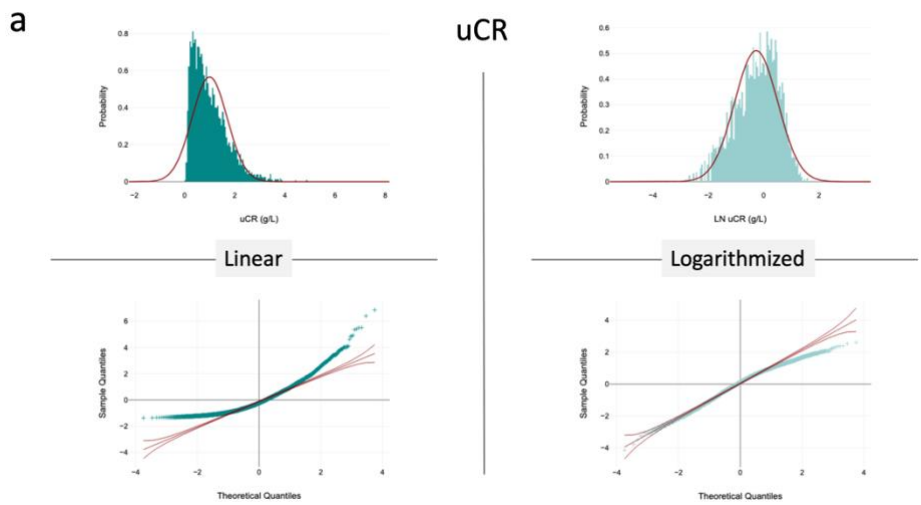


Appendix 8b: Compare quartiles of uAs$_{UC}$, uAs$_C$, and uAs$_N$ for S-PFCRC (c) and V-PFCRC (d).
Preliminary analysis of all n=298 samples from in-house patients with uAs$_{UC}$ < 100 µg/L.
S-PFCRC (c) still exhibits stronger residual dependencies (high R^2 and exponent b) in the marginal quartiles 1 and 4, which is resolved by V-PFCRC (d).

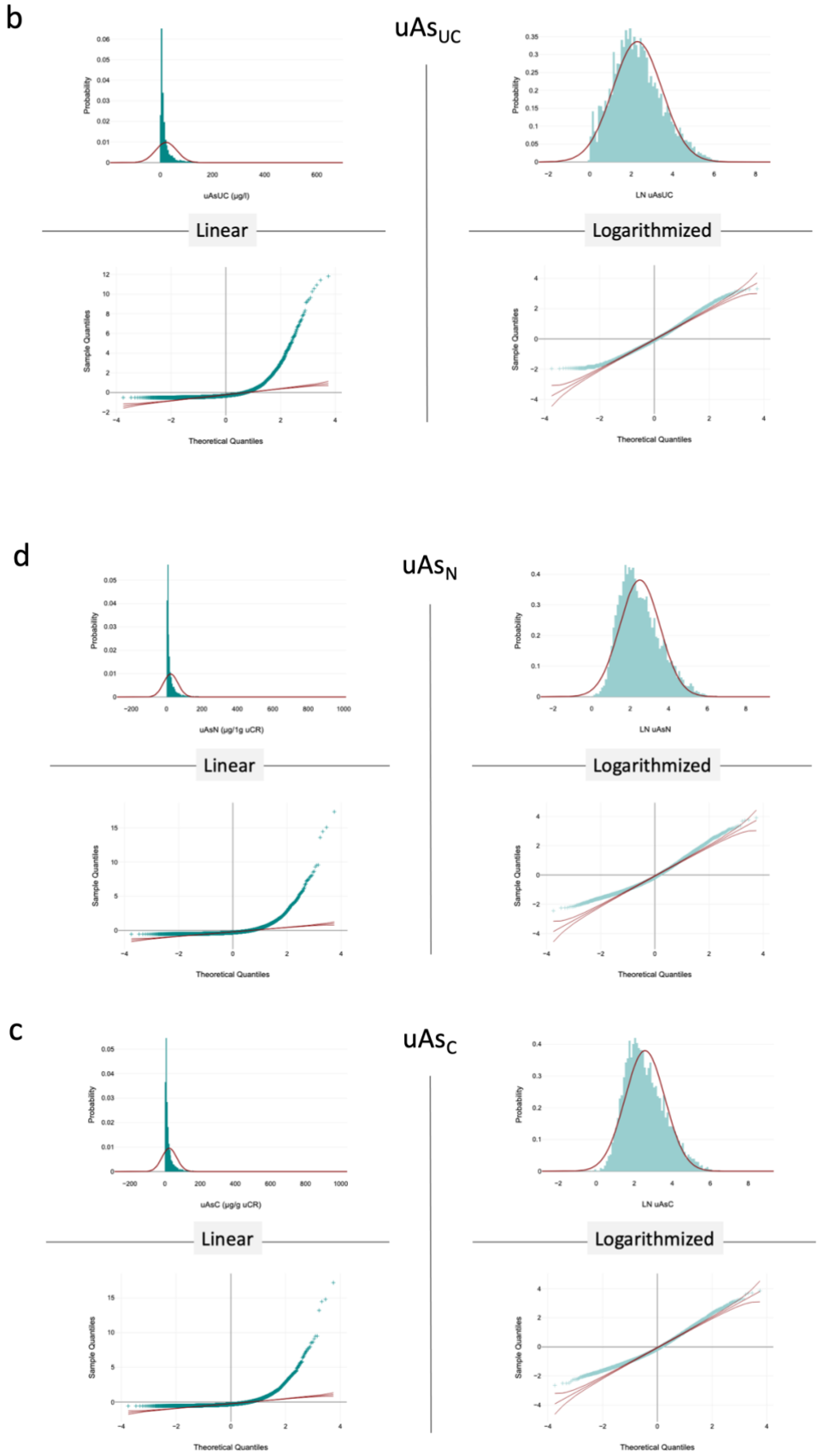


Appendix 9a: The uCR-differential linear relationship between coefficient *a* and uAs_{UC}.
Data were apportioned in 8 uCR-ranges and 5 uAsUC-quintiles and analyzed by power regression analog to Figs.2a/b, yielding 5 curves of the form $uAs_{UC} = a \times uCR^{b}$. Linear relation between uAs_{UC} (mean of respective quintile) and *a* of the corresponding curve were established by linear regression for all 8 uCR ranges (uCR-means in the legend of App.9b). The gradual decrease of line slopes with urinary concentrations (uCR) signifies uAs_N (= *a*) to take on smaller values than uAs_{UC} in samples with uCR > 1 g/L and higher than uAs_{UC} with uCR < 1 g/L**.**

Appendix 9b: The uCR-differential logarithmical relationship between exponent *b* and uAs_{UC}.
The positive logarithmic dependence of exponent *b* on uAs_{UC} is illustrated by the 5 power equations of App.9b for all 8 uCR. Higher renal As-loads may raise the fractional excretion rate of uAs, thus converging it with the fractional excretion of uCR, which under ordinary conditions is not reabsorbed at all. The consequences are more negligible CRCEs at higher uAs_{UC} due to the approximation of *b* to the value of 1. In addition to the uAs_{UC}-dependent increment of exponent *b*, a negative dependence of the curve slopes on uCR was observed. In all eight uCR ranges, the curves become gradually steeper with urinary dilution, indicating that closer to linear relations between uCR and uAS_{UC} are attained already at lower uAs_{UC}.



Appendix 10a: Effect of logarithmization on data distribution of uCR. Histograms and Quantile-Quantile plots of all n=5599 samples with $uAs_{UC} \leq 500$ µg/L are given in linear (left) and logarithmized (right) form for uCR (a), uAs_{UC} (b), uAs (c).



Appendix 10b-d: Effect of logarithmization on data distribution of uAs. Histograms and Quantile-Quantile plots of all n=5599 samples with $uAs_{UC} \leq 500$ µg/L are given in linear (left) and logarithmized (right) form for uCR(a), uAs_{UC}(b), uAs(c).