Seminars in Arthritis and Rheumatism 63 (2023) 152303

Contents lists available at ScienceDirect

Seminars in Arthritis and Rheumatism

journal homepage: www.elsevier.com/locate/semarthrit

The statistical challenge of analysing changes in dual energy computed tomography (DECT) urate volumes in people with gout

Sarah Stewart [a,b,*], Greg Gamble [b], Anthony J Doyle [b,c], Chang-Nam Son [d], Opetaia Aati [b], Kieran Latto [b], Anne Horne [b], Lisa K Stamp [e], Nicola Dalbeth [b]

[a] School of Clinical Sciences, Faculty of Health and Environmental Sciences, Auckland University of Technology, 90 Akoranga Drive, Northcote Auckland 0627, New Zealand
[b] Department of Medicine, Faculty of Medical and Health Sciences, University of Auckland, 85 Park Road, Grafton, Auckland 1023, New Zealand
[c] Te Whatu Ora Health New Zealand, Te Toka Tumai Auckland, Radiology, Private Bag 92 024, Auckland 1142, New Zealand
[d] Department of Rheumatology, Uijeongbu Eulji Medical Center, Eulji University School of Medicine, 712 Dongil-ro, Uijeongbu 11749, South Korea
[e] Department of Medicine, University of Otago, Christchurch, 2 Riccarton Avenue, Christchurch 8011, New Zealand

ARTICLE INFO

Keywords:
Gout
Urate
Statistical analysis
Dual-energy computed tomography

ABSTRACT

Background: Dual energy computed tomography (DECT) allows direct visualization of monosodium urate crystal deposition in gout. However, DECT urate volume data are often highly skewed (mostly small volumes with the remainder considerably larger), making statistical analyses challenging in longitudinal research. The aim of this study was to explore the ability of various analysis methods to normalise DECT urate volume data and determine change in DECT urate volumes over time.

Methods: Simulated datasets containing baseline and year 1 DECT urate volumes for 100 people with gout were created from two randomised controlled trials. Five methods were used to transform the DECT urate volume data prior to analysis: log-transformation, Box-Cox transformation, log(X-(min(X)-1)) transformation; inverse hyperbolic sine transformation, and rank order. Linear regression analyses were undertaken to determine the change in DECT urate volume between baseline and year 1. Cohen's d were calculated as a measure of effect size for each data treatment method. These analyses were then tested in a validation clinical trial dataset containing baseline and year 1 DECT urate volumes from 91 people with gout.

Results: No data treatment method successfully normalised the distribution of DECT urate volumes. For both simulated and validation data sets, significant reductions in DECT urate volumes were observed between baseline and Year 1 across all data treatment methods and there were no significant differences in Cohen's d effect sizes.

Conclusions: Normalising highly skewed DECT urate volume data is challenging. Adopting commonly used transformation techniques may not significantly improve the ability to determine differences in measures of central tendency when comparing the change in DECT urate volumes over time.

Abbreviations

CI	Confidence interval
DECT	Dual-energy computed tomography
MSU	Monosodium urate
SE	Standard Error

1. Background

Monosodium urate (MSU) crystal deposition in joints and soft tissue structures plays a central role in the pathology and clinical manifestations of gout, including painful gout flares, tophus formation, and bone damage [1–3]. Dual-energy computed tomography (DECT) is an imaging modality which has the ability to directly visualise MSU crystal deposition through colour-coding urate based on its chemical composition [4]. The development of specific software has also allowed for the automated measurement of regional urate volume [5].

Reporting DECT urate volumes over time is increasingly common in gout observational research and clinical trials [5–11]. However, DECT urate volume is highly variable across different patients; most have very small urate volumes, and the remainder considerably larger [5,9]. DECT urate volume data therefore display non-normal distributions, which

* Corresponding author.
E-mail address: sarah.stewart@aut.ac.nz (S. Stewart).

https://doi.org/10.1016/j.semarthrit.2023.152303

Available online 2 November 2023
0049-0172/© 2023 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

result in a loss of statistical power when detecting differences over time. This variation has resulted in inconsistent statistical methods used to analyse change in DECT volume over time in response to various urate lowering therapies, including the use of a range of non-parametric tests [5,12-14] and data transformations [6]. Non-normal distributions may also contribute to the variable sensitivity to change of DECT urate volume that has been demonstrated amongst longitudinal studies of people with gout assessing response to urate lowering therapy [5-7,9,15]. The aim of this study was to explore the ability of various analysis methods to normalise DECT urate volume data and determine change in DECT urate volumes over time.

2. Methods

2.1. Data sets

Data collected as part of two randomised controlled trials which measured the total DECT urate volume within the foot/ankle region of each participant (a continuous variable measured in cm^3) were used to demonstrate various analysis methods for the purpose of the current paper. Data from the first trial consisted of 104 participants with erosive gout and baseline serum urate >0.30 mmol/L (67.3% with at least one subcutaneous tophus at baseline) who were randomly assigned to urate lowering therapy titrated to a serum urate target <0.30 mmol/L (control), or to a urate target <0.20 mmol/L (intervention) [16]. The second dataset comprised of 125 participants with gout and baseline serum urate >0.36 mmol/L (45.6% with at least one subcutaneous tophus at baseline) who were randomly assigned to continue their current allopurinol dose (control) or to allopurinol dose titration to a serum urate target <0.36 mmol/L (intervention) [6]. The first trial was approved by the Southern Health and Disability Ethics Committee, New Zealand and the second trial by the Multi-Regional Ethics Committee, New Zealand. All participants in both trials provided written informed consent.

Only baseline and year 1 data were used for the current analysis. Ten participants from the first trial and 37 participants from the second trial were excluded due to incomplete paired data from DECT scans at baseline and after 1 year, resulting in the inclusion of 94 and 88 participants from each trial, respectively.

In both trials, DECT scans of the feet/ankle region were performed on a dual X-ray tube 128 detector row scanner (Somatom Definition Flash, Siemens Healthineers, Erlangen, Germany). All scans were performed with the same image protocol; acquisition at 128×0.6 mm and pitch of 0.7. X-ray tube 1 is operated at 80 kV/260 mA and tube 2 at 140 kV/130 mA, with tin filtration. The images were reconstructed on a medium-soft bone algorithm (Siemens kernel D30s), 512 matrix, to 0.75 mm slices with a 0.5 mm increment. The images were viewed both as 0.75 mm slices and reconstructed 3 mm slices on a Picture Archiving Communication System (PACS). Volume of urate deposition was measured using the Siemens proprietary workstation (MultiModality Workspace, Siemens Healthineers) and the proprietary syngo gout package. For both trials, two independent experienced readers calculated total urate volumes in the feet/ankle regions for each participant. Readers removed obvious artefact from regions of interest before urate volume assessment. Inter-reader intraclass correlation coefficients were ≥0.99 for both trials. The mean score from both readers was used in the analysis.

To reduce variability due to differences in trial-specific factors, data from half of the participants from each trial were randomly chosen and combined to create a discovery dataset and a validation dataset. Equal numbers of cases and controls from each trial were included in each dataset. Baseline demographic and clinical characteristics for participants in the discovery and validation cohorts are presented inTable 1.

Table 1
Baseline demographic and clinical characteristics of participants in discovery and validation cohorts.

		Discovery cohort	Validation cohort
N		91	91
Baseline serum urate, mean (SD), mmol/L		0.38 (0.07)	0.41 (0.10)
Year 1 serum urate, mean (SD), mmol/L		0.29 (0.09)	0.32 (0.12)
Age, mean (SD), years		61.0 (11.9)	61.1 (12.7)
Sex, n (%)	Female	6 (6.6%)	5 (5.5%)
	Male	85 (93.4%)	86 (94.5%)
Gout disease duration, mean (SD), years		19.5 (12.5)	22.0 (12.6)
Gout flares in past month, mean (SD)		1.7 (5.5)	1.2 (4.6)
At least one subcutaneous tophus, n (%)		48 (52.7%)	50 (54.9%)
Body mass index, mean (SD), kg/m^3		33.6 (7.3)	35.5 (8.6)
Diabetes, n (%)		17 (18.7%)	22 (24.2%)
Hypertension, n (%)		62 (68.1%)	57 (62.6%)
Cardiovascular disease, n (%)		22 (24.2%)	26 (28.6%)
Dyslipidaemia, n (%)		48 (52.7%)	49 (53.8%)
Renal impairment, n (%)		28 (30.8%)	38 (41.8%)
Chronic kidney disease, n (%)		15 (16.5%)	20 (22.0%)
Pain visual analogue scale, mean (SD)		1.3 (2.1)	1.6 (2.3)

2.2. Statistical analysis

2.2.1. Discovery cohort simulated datasets

Data from the non-zero containing DECT urate volumes from the discovery cohort were used to create two simulated datasets each containing baseline and year 1 DECT urate volumes, and baseline serum urate levels for 100 participants. The datasets were created in RStudio (v.2022.07.1) using the rlnorm and simstudy packages and a correlation matrix derived from the original discovery cohort. Gamma distributions were used for both simulated datasets to reflect the distribution of DECT urate volumes in the original discovery cohort. The first simulated dataset used dispersion parameters of 5 to approximate the proportion of DECT urate volumes in the original discovery cohort in which 28/91 (30.8%) participants had DECT urate volumes <0.05cm^3. Dispersion parameters were increased to 7.5 in the second simulated dataset to approximate a larger proportion of very low DECT urate volumes. The two simulated datasets contained approximately 30% DECT urate volumes <0.05cm^3 (Fig. 1a) and 50% DECT urate volumes <0.05cm^3 (Fig. 1b). Frequency distribution tables for the two simulated datasets are also shown in **Supplementary Table 1**. Volumes <0.05cm^3 may be associated with artefacts seen in nail beds, thickened plantar skin (calluses) and noise-related submillimetre specks [17].

In addition to using simulated non-transformed continuous measures of DECT urate volume (data treatment method 1) from each of the two simulated datasets, five approaches were used to transform the simulated DECT urate volume data from each time point prior to analysis: a log-transformed continuous measure of DECT urate volume (data treatment method 2); a Box-Cox transformed continuous measure of DECT urate volume (with lambda determined empirically by SPSS) [18] (data treatment method 3); a log(X-(min(X)-1)) transformed continuous measure of DECT urate volume (data treatment method 4); an inverse hyperbolic sine transformed continuous measure of DECT urate volume (data treatment method 5); and the rank of continuous DECT urate volumes (data treatment method 6) (Table 2). For data treatment method 2 (log transformation), a value of 0.005 was added to any DECT urate volume = zero, prior to transformation. A total of 12 simulated data sets (2 simulated datasets x 6 data treatment methods) were created from the discovery cohort.

Mixed linear regression models were used to determine the mean (95% confidence interval (CI)) change in DECT urate volume between baseline and year 1. Prior to analyses, the distribution of residuals from the models were examined via visual inspection of histograms and normal Q-Q plots and calculation of Kolmogorov Smirnov and Shapiro

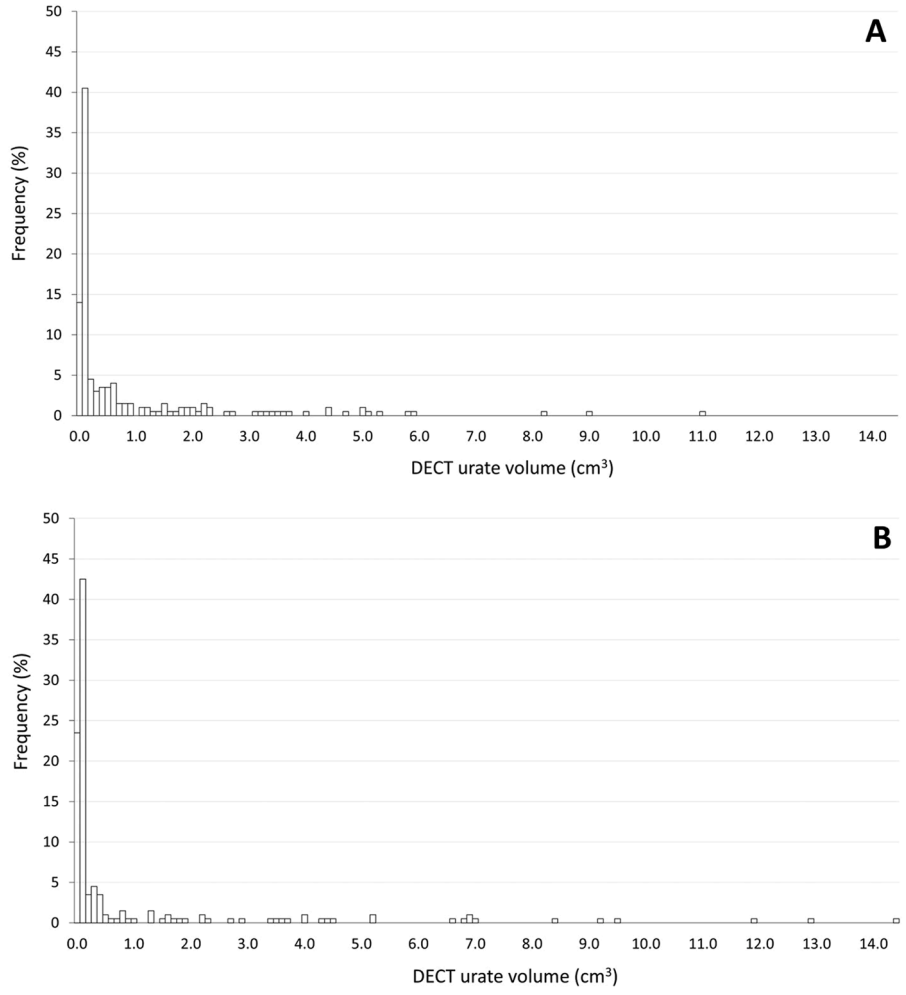

Fig. 1. Histograms showing distribution of simulated DECT urate volumes with 30% of volumes <0.05cm^3 (**A**); and 50% of volumes <0.05cm^3 (**B**).

Table 2
Description of six methods of data treatment prior to analysis of change in DECT urate volume between baseline and year 1.

Data treatment	Description
1	A non-transformed continuous measure of DECT urate volume
2	A log-transformed continuous measure of DECT urate volume (with an offset of 0.005 for volumes = zero)
3	A Box-Cox transformation of continuous measure of DECT urate volume (with lambda empirically defined by SPSS)
4	A log(X-(min(X)-1)) transformation
5	Inverse hyperbolic sine transformation
6	The rank of continuous DECT urate volumes

Wilks statistics. The DECT urate volume was entered as the dependant variable, and the time point (baseline/year 1) as the independent variable. Year 1 (i.e., baseline) serum urate status was also included in the models as a covariate (i.e., participants <0.36 mmol/L vs participants ≥0.36 mmol/L to reflect the urate target recommended in the 2020 American College of Rheumatology gout management guidelines [19]). This allowed determination of differences in urate volume between the two time points while adjusting for baseline urate target status. This analysis was performed for each of the 12 simulated data sets. The models were not adjusted for any further covariates due to the different interactions between the transformed DECT volume values and covariates. An unstructured covariance structure was used in all repeated measures mixed models after establishing that it provided the lowest Akaike Information Criteria (AIC) value [20]. All regression models were run in IBM SPSS Statistics v. 25.

2.2.1.1. Method comparison. Cohen's d and it's 95% CI were calculated from the regression estimates as a measure of effect size of the differences in DECT urate volume means between baseline and Year 1. Cohen's d and their 95% CIs were used to compare the different data treatment methods. Cohen's d ranges from 0 to 1, with values of 0.2 considered a small effect size, values of 0.5 a medium effect size, and values of 0.8 a large effect size [21]. Based on a fixed sample size, the smaller the Cohen's d value, the greater the ability to detect a smaller treatment effect between the two time points for the same sample size.

2.3. Validation study

The above analyses were then undertaken in the validation cohort (comprised of the other random half of the combined datasets from the two trials), with the addition of the intervention group (i.e., intervention vs control) included as a covariate and the interaction effect of intervention group*timepoint also included. The validation cohort contained 23 controls and 21 cases from trial 1 and 23 controls and 24 cases from trial 2, totalling 91 participants. This cohort contained 37% of DECT urate volumes <0.05cm^3 and 59% of DECT urate volumes <0.25cm^3 (Fig. 2).

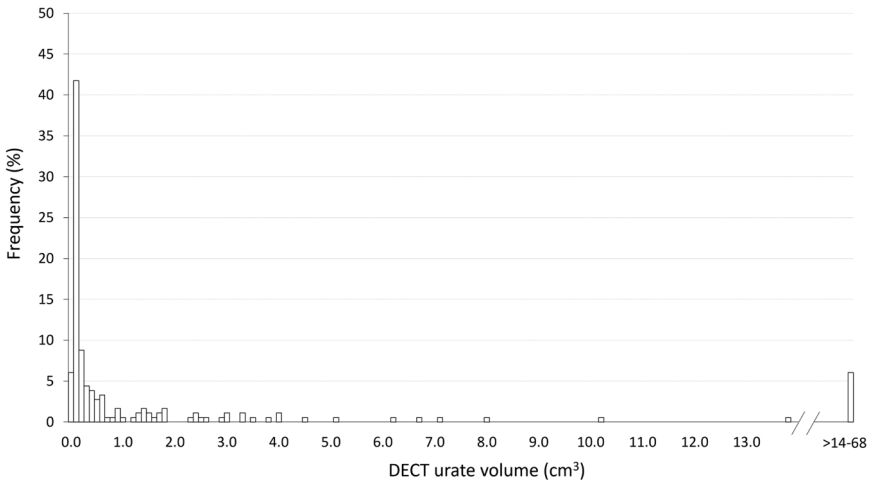

Fig. 2. Histograms showing distribution of DECT urate volumes in the validation cohort.

3. Results

3.1. Simulation study

Fig. 3 and Fig. 4 present the histograms and normal Q-Q plots, respectively, showing the distribution of standardised residuals from the linear regression models for each simulated dataset for each data treatment method. The Kolmogorov Smirnov and Shapiro Wilks statistics are shown in Table 3. Although data treatment method 2 (log transformation) and 6 (rank) produced the least skewed distributions, they were not completely normalized, with log transformation producing left skewed data and the rank transformation producing a more uniform distribution.

Significant reductions in DECT urate volumes were observed between baseline and Year 1 across all data treatment methods for the simulated data sets containing 30% and 50% of DECT urate volumes $<0.05cm^3$, respectively (Table 4). Back-transformation of the mean difference and 95% CIs for the log transformed variables gave $0.67cm^3$ (0.60, 0.75) and $0.63cm^3$ (0.48, 0.83) for the simulated dataset with 30% and 50% DECT urate volumes $<0.05cm^3$, respectively (note that the difference between the logarithms of the two geometric means, is the logarithm of their ratio, not of their difference and so these anti logged values are bound a dimensionless ratio).

For the simulated data set containing 30% DECT urate volumes $<0.05cm^3$, Cohen's d effect size was small for data treatment method 1, and medium for data treatment methods 2 to 6. For the simulated data set containing 50% DECT urate volumes $<0.05cm^3$, Cohen's d effect size was small for data treatment method 1 and 2, and medium for data treatment methods 3 to 6. There was no clinically important differences in Cohen's d values between the data treatment methods for any simulated data set evident by the overlapping confidence intervals (Fig. 5).

3.2. Validation study

Significant reductions in DECT urate volumes were observed between baseline and Year 1 for all data treatment methods (Table 5). Cohen's d was small for data treatment method 1 and medium for data treatment methods 2 to 6 (Table 5). There were no statistically significant differences in Cohen's d values between the six data treatment methods (Fig. 6). The back-transformed mean difference and 95% CIs for the log-transformed DECT urate volumes in the validation data set was 0.61 cm^3 (0.50, 0.75).

4. Discussion

This analysis compared six different statistical approaches to treating highly skewed DECT urate volume data in studies determining change in DECT urate volume over time (i.e., when assessing change during urate lowering therapy in people with gout). This analysis shows that the distribution of continuous DECT urate volume data is extremely challenging to normalise, especially when there are substantial numbers of participants with DECT urate volumes below the limits of detection/measurement. Despite the application of commonly used transformation techniques, the ability to determine differences in measures of central tendency (i.e., means) remains similar to using non-transformed data. This raises the question of what a measure of central tendency is when most DECT urate volumes are very low.

The analyses in the current study were undertaken regardless of the distribution of residuals. Linear regression models are valid regardless of distribution when sample sizes are adequate [22]. Applying transformation with the aim of normalising data does not make a major difference to the ability to detect treatment effects. Furthermore, alternatives to untransformed data are likely not warranted in larger sample sizes [22]. However, it should be noted that this does not apply for prediction. Data transformation also creates an additional layer of complexity in interpretation of the results. Clinically meaningful interpretation of the data would require back-transformation of the results to the original volumetric units (i.e., cm^3). However, back-transformed confidence intervals often cannot be interpreted on this scale [23]. Even antilog transformation, which initially may be viewed as straightforward, will only produce the dimensionless ratio of the two geometric means, not their difference [24,25].

The distribution of DECT urate volumes used in the simulation and validation datasets in the current analysis were similar to those reported in other gout studies, in which most participants have very low volumes, and the remaining few proportionally higher [9,26-29]. For example, a cross-sectional multi-centre study of people with gout on allopurinol reported a total mean DECT urate volume of $0.16cm^3$ across the hands, feet and knees, but a wide range between $0.01cm^3$ and $19.53cm^3$ [26]. DECT urate volume is generally higher in patients with longer disease duration, higher gout flare frequency, and subcutaneous tophi [26,27]. While DECT readers removed obvious artefact from regions of interest before urate volume measurement, it should also be acknowledged that artefacts that contribute to small, automated urate volumes, would not be expected to change during urate-lowering therapy [17]. Challenges persist with DECT of small urate volumes; minor changes in scanning technique (including small fluctuations in x-ray beam voltage, detector sensitivity and even patient positioning), image reconstruction (such as

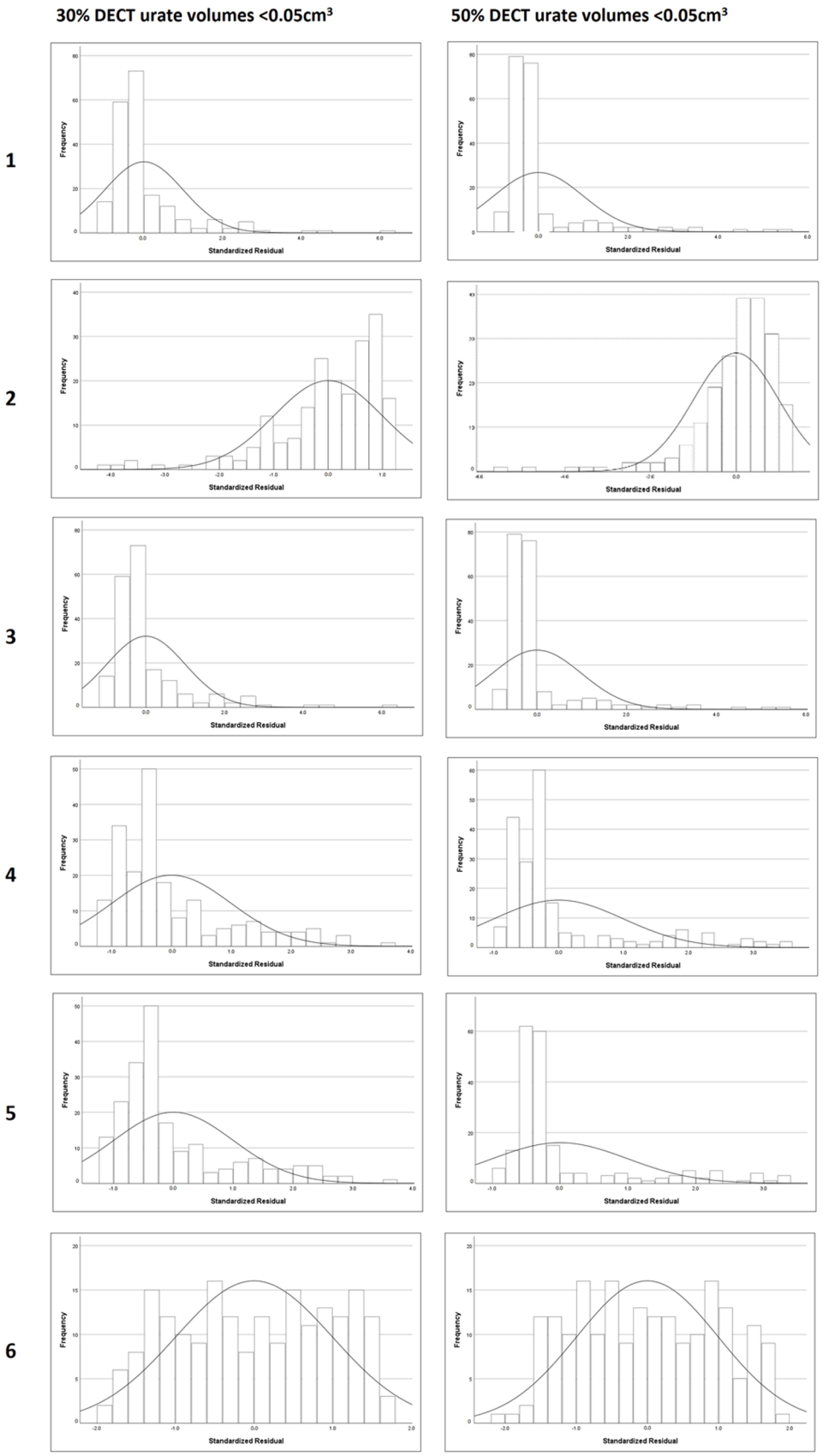

Fig. 3. Histogram montage showing distribution of standardised residuals from linear regression models for simulated dataset. Data treatment methods: **1** non-transformed; **2** log-transformed; **3** Box-Cox transformation; **4** log(X-(min(X)-1)) transformation; **5** inverse hyperbolic sine transformation; **6** rank.

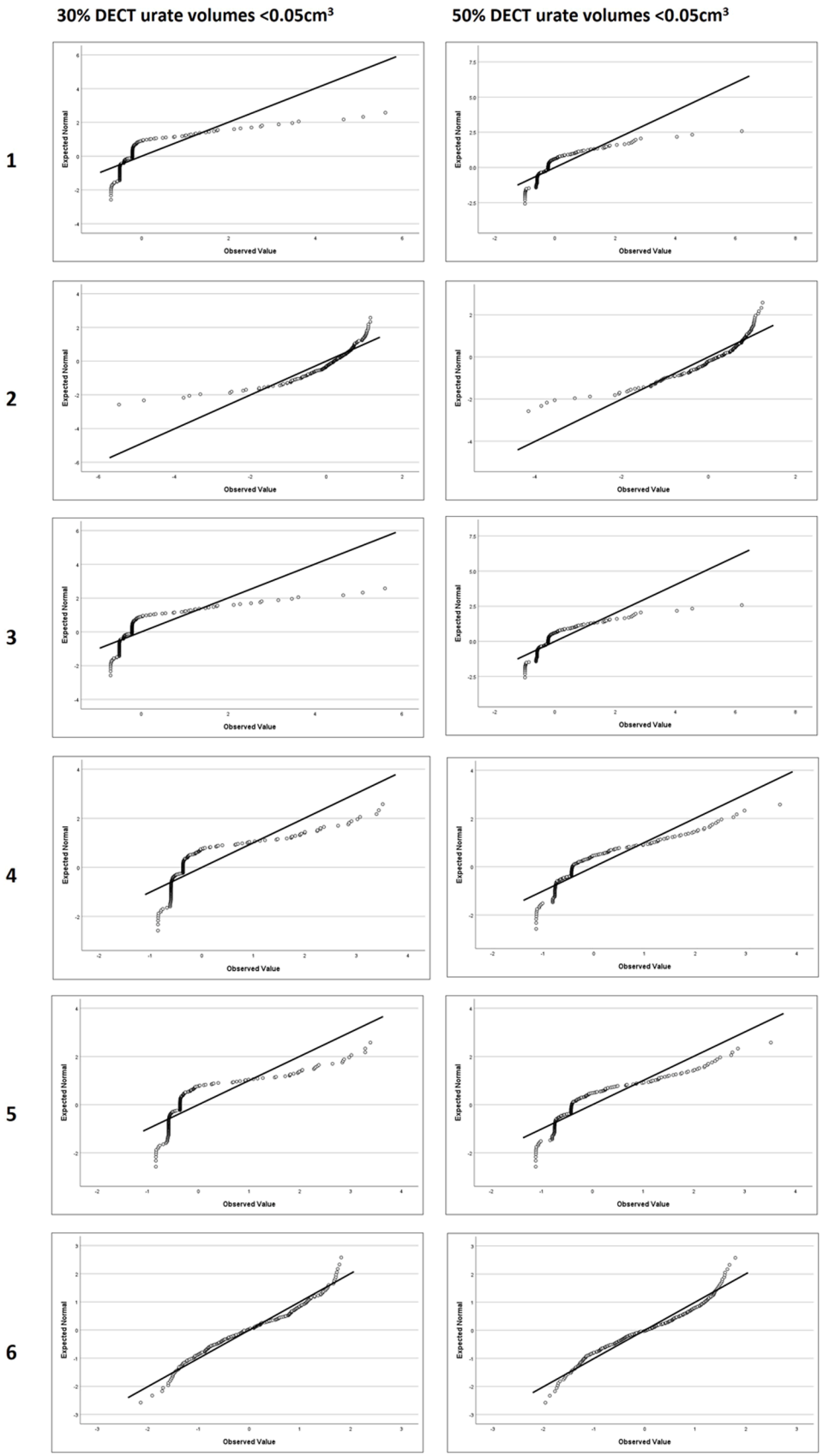

Fig. 4. Normal Q-Q plot montage showing distribution of standardised residuals from linear regression models for simulated dataset. Data treatment methods: **1** non-transformed; **2** log-transformed; **3** Box-Cox transformation; **4** log(X-(min(X)-1)) transformation; **5** inverse hyperbolic sine transformation; **6** rank.

Table 3
Kolmogorov-Smirnov and Shapiro-Wilk statistics for simulated data sets.

Data treatment method[a]	Simulated data set with DECT urate volumes with 30% <0.05cm3		Simulated data set with DECT urate volumes with 50% <0.05cm^3	
	Statistic	P-value	Statistic	P-value
Kolmogorov-Smirnov test				
1	0.247	<0.001	0.339	<0.001
2	0.118	<0.001	0.151	<0.001
3	0.247	<0.001	0.339	<0.001
4	0.201	<0.001	0.290	<0.001
5	0.208	<0.001	0.303	<0.001
6	0.070	0.019	0.082	0.002
Shapiro-Wilk test				
1	0.701	<0.001	0.564	<0.001
2	0.861	<0.001	0.798	<0.001
3	0.701	<0.001	0.564	<0.001
4	0.831	<0.001	0.676	<0.001
5	0.822	<0.001	0.668	<0.001
6	0.957	<0.001	0.966	<0.001

[a]Data treatment methods: 1 non-transformed; 2 log-transformed; 3 Box-Cox transformation; 4 log(X-(min(X)-1)) transformation; 5 inverse hyperbolic sine transformation; 6 rank.

kernel selection) and tissue decomposition algorithm settings can produce variations of the order of magnitude of the presumed urate deposits being investigated [30–33].

The challenges observed in analysing DECT urate volume data where the study sample frame typically incorporates a non-trivial proportion of participations without MSU deposition (and therefore negatively skewed volume data), shares similar analysis challenges observed in analysing tumour and infarct volumes, and in functional MRI activated voxel volumes [34–36]. However, DECT urate volume analysis has some important differences from approaches used in tumour and infarct volume analysis. DECT analysis currently focuses on average or total burden of urate volume for individuals rather than analysis of tracked changes in volume at specific locations on a pixel/voxel basis [37] or by assessment of distributions [38].

The results from this analysis have highlighted the impracticality of analysing change in urate volume over time with data containing very small DECT urate volumes. Even with transformations for skewed data, floor effects limit sensitivity to change in this setting. This has important implications for research design and sample size calculations, particularly when studying people with asymptomatic hyperuricemia or early in the gout disease course who have particularly small urate volumes (ranging from a mean of 0.02cm^3 to 0.06cm^3) [27-29,39]. Although these very small deposits may play a role in the diagnosis of gout [28], they may render parametric analyses of change over time challenging.

Other alternative approaches exist for studies with highly skewed and very small DECT urate volumes, including ordinal or multinomial analysis modelling. For example, replacing the continuous measure of DECT urate volume with a semi-quantitative DECT urate scoring system which has demonstrated adequate feasibility, reliability, and discriminative power, alongside addressing the issue of non-parametricity [9]. In distributions containing a large proportion of very small DECT urate volumes, a binomial model based on volumes above and below a minimally detectable difference (i.e., none vs. some, or, some vs. more) may also be more appropriate, as would a model comparing no change, an increase, or a decrease in DECT urate volume between time points. Alternatively, researchers may consider implementing a participant eligibility criterion that requires inclusion of those with larger DECT urate volumes at baseline (e.g., recruitment of study participants with baseline DECT urate volumes of >1.0cm^3). Using this screening approach, a hierarchical analysis may be useful to examine the proportion of participants with zero volumes at each follow up time point, as well as an analysis of change in DECT volume in those with measurable DECT urate volume at baseline. However, screening can be costly, and this approach would exclude those with no DECT evidence of MSU crystal deposition at baseline who may develop measurable DECT urate volumes during the follow-up period. Further research is also required to determine a baseline DECT urate volume threshold for

Table 4
Difference in DECT urate volumes between baseline and Year 1 for each data treatment method using the simulated datasets created from the discovery cohort. All estimates are presented in their transformed units.

Data treatment method[a]	Time point	Least-squares mean	Diff	SE	P	Cohen's d d	95% CI Lower	Upper
Simulated data set with DECT urate volumes with 30% <0.05cm^3								
1	Baseline	1.13	-0.59	0.12	<0.001	0.48	0.26	0.69
	Year 1	0.54						
2	Baseline	-1.41	-0.40	0.06	<0.001	0.71	0.49	0.93
	Year 1	-1.82						
3	Baseline	0.12	-0.24	0.04	<0.001	0.57	0.34	0.79
	Year 1	-0.12						
4	Baseline	0.48	-0.18	0.03	<0.001	0.67	0.45	0.88
	Year 1	0.30						
5	Baseline	0.61	-0.24	0.04	<0.001	0.67	0.45	0.88
	Year 1	0.38						
6	Baseline	107.3	-13.6	1.75	<0.001	0.78	0.55	1.00
	Year 1	93.7						
Simulated data set with DECT urate volumes with 50% <0.05cm^3								
1	Baseline	1.26	-0.68	0.18	0.003	0.38	0.16	0.59
	Year 1	0.59						
2	Baseline	-2.42	-0.46	0.14	0.002	0.32	0.12	0.52
	Year 1	-2.88						
3	Baseline	0.09	-0.18	0.05	<0.001	0.40	0.19	0.61
	Year 1	-0.09						
4	Baseline	0.42	-0.15	0.03	<0.001	0.54	0.32	0.74
	Year 1	0.27						
5	Baseline	0.53	-0.19	0.04	<0.001	0.54	0.33	0.75
	Year 1	0.34						
6	Baseline	105.4	-9.9	2.07	<0.001	0.48	0.27	0.68
	Year 1	95.6						

[a]Data treatment methods: 1 non-transformed; 2 log-transformed; 3 Box-Cox transformation; 4 log(X-(min(X)-1)) transformation; 5 inverse hyperbolic sine transformation; 6 rank. For Method 2, a lambda of -2 was used for 30% dataset, and -2.5 for 50% dataset.

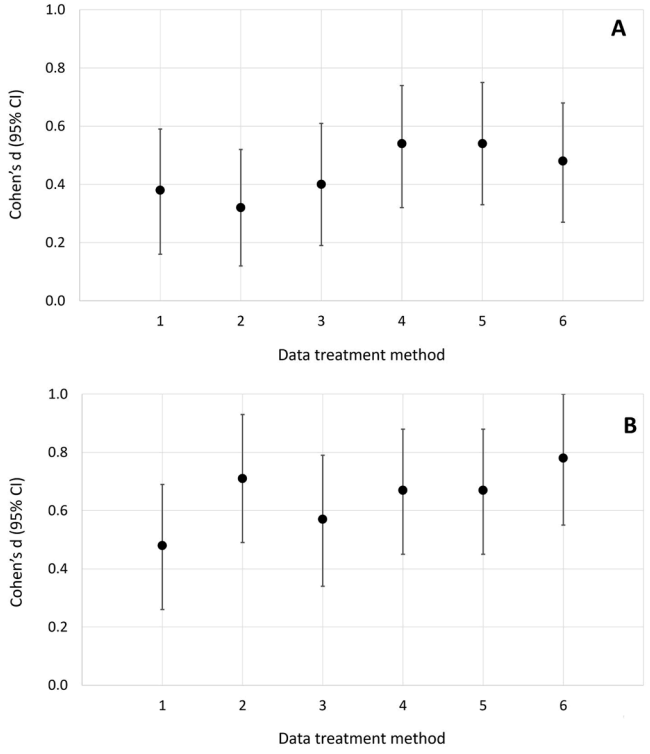

Fig. 5. Cohen's d and 95% confidence intervals (CI). **A** Simulated data set with DECT urate volumes with 30% <0.05cm^3; **B** Simulated data set with DECT urate volumes with 50% <0.05cm^3. Data treatment methods: **1** non-transformed; **2** log-transformed; **3** Box-Cox transformation; **4** log(X-(min(X)-1)) transformation; **5** inverse hyperbolic sine transformation; **6** rank.

participant screening that would result in a more parametric distribution of volume measures. It should be noted that the dispersion parameters chosen to create the simulated data sets in the current study were based on what we believe are clinically reflective of DECT urate volumes in people with gout, and statistical approaches may differ in other populations with very high and very low proportions of near-zero DECT urate volumes. Other alternative analysis models, which do not require assuming distributions, may also be potentially appropriate, including bootstrapping or generalised additive models (GAMMLS) where the mean and variance can be separately modelled.

Although some may consider analysis of percent change in DECT urate volumes from baseline as a further alternative, this method is discouraged as statistical inefficient [40]. Furthermore, analysis of percent change from baseline are unlikely to yield meaningful results in observational studies when there is a correlation between the exposure and baseline measures of the outcome [41]. A mixed-models approach to repeated measures was used in the current study with absolute DECT volume as the dependant variable as this method is robust to moderate levels of missingness and permits all participants with some data to be included in the analysis.

In conclusion, the analyses presented in this paper have highlighted the challenge in normalising highly skewed DECT urate volume data. Commonly adopted transformation techniques to address this, may not improve the ability to determine differences in measures of central tendency.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRediT authorship contribution statement

Sarah Stewart: Conceptualization, Methodology, Formal analysis, Data curation, Writing – original draft, Writing – review & editing. **Greg Gamble:** Conceptualization, Methodology, Formal analysis, Data curation, Supervision, Writing – original draft, Writing – review & editing. **Anthony J Doyle:** Investigation, Data curation, Writing – review & editing. **Chang-Nam Son:** Investigation, Data curation, Writing – review & editing. **Opetaia Aati:** Investigation, Data curation, Writing – review

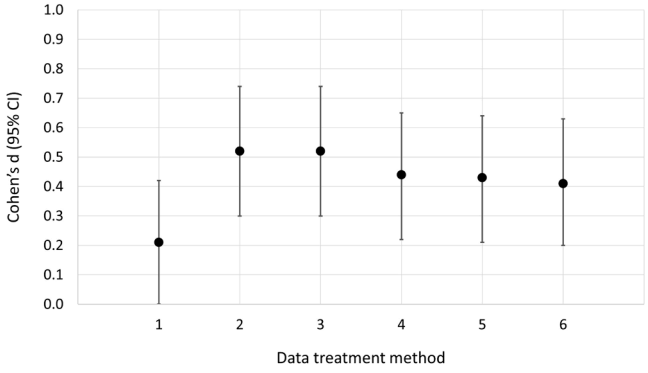

Fig. 6. Cohen's d and 95% confidence intervals (CI) for validation data set. Data treatment methods: **1** non-transformed; **2** log-transformed; **3** Box-Cox transformation; **4** log(X-(min(X)-1)) transformation; **5** inverse hyperbolic sine transformation; **6** rank.

Table 5
Difference in DECT urate volumes between baseline and Year 1 in the validation cohort. All estimates are presented in their transformed units.

Data treatment method[a]	Time point	Least-squares mean	Diff	SE	P	Cohen's d		
						d	95% CI Lower	Upper
1	Baseline	3.39	-1.06	0.522	0.045	0.21	0.00	0.42
	Year 1	2.34						
2	Baseline	-1.48	-0.49	0.101	<0.001	0.52	0.30	0.74
	Year 1	-1.98						
3	Baseline	0.13	-0.22	0.047	<0.001	0.52	0.30	0.74
	Year 1	-0.09						
4	Baseline	0.67	-0.20	0.048	<0.001	0.44	0.22	0.65
	Year 1	0.47						
5	Baseline	0.82	-0.25	0.059	<0.001	0.43	0.21	0.64
	Year 1	0.57						
6	Baseline	95.9	-9.19	2.348	<0.001	0.41	0.20	0.63
	Year 1	86.7						

[a]Data treatment methods: 1 non-transformed; 2 log-transformed; 3 Box-Cox transformation; 4 log(X-(min(X)-1)) transformation; 5 inverse hyperbolic sine transformation; 6 rank. For Method 2, a lambda of -1.5 was used.

& editing. **Kieran Latto:** Investigation, Data curation, Writing – review & editing. **Anne Horne:** Investigation, Data curation, Writing – review & editing. **Lisa K Stamp:** Resources, Data curation, Writing – review & editing. **Nicola Dalbeth:** Conceptualization, Methodology, Investigation, Resources, Data curation, Supervision, Project administration, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

Lisa Stamp reports grants or contracts from Health Research Council of New Zealand, personal royalties or licenses from UptoDate, and consulting fees from Pharmac. **Nicola Dalbeth** reports grants or contracts from Health Research Council of New Zealand and Novotech, personal consulting fees from AstraZeneca, Dyve Biosciences, Horizon, Selecta, Arthrosi, JW Pharmaceutical Corporation, PK Med, PTC Therapeutics, Protalix, Cello Health, JPI, Unlocked Labs, and LG, payment or honoraria from Novartis and Hikma. The other authors declare no conflicts of interest.

Acknowledgements

None.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.semarthrit.2023.152303.

References

[1] Dalbeth N, Clark B, Gregory K, Gamble G, Sheehan T, Doyle A, McQueen FM. Mechanisms of bone erosion in gout: a quantitative analysis using plain radiography and computed tomography. Ann Rheum Dis 2009;68:1290–5. https://doi.org/10.1136/ard.2008.094201.
[2] Dalbeth N, Pool B, Gamble GD, Smith T, Callon KE, McQueen FM, Cornish J. Cellular characterization of the gouty tophus: a quantitative analysis. Arthritis Rheum 2010;62:1549–56. https://doi.org/10.1002/art.27356.
[3] Faires J, Mccarty Jr D. Acute arthritis in man and dog after intrasynovial injection of sodium urate crystals. Lancet 1962;280:682–5. https://doi.org/10.1016/s0140-6736(62)91050-4.
[4] Choi HK, Al-Arfaj AM, Eftekhari A, Munk PL, Shojania K, Reid G, Nicolaou S. Dual energy computed tomography in tophaceous gout. Ann Rheum Dis 2009;68:1609–12. https://doi.org/10.1136/ard.2008.099713.
[5] Rajan A, Aati O, Kalluru R, Gamble GD, Horne A, Doyle AJ, McQueen FM, Dalbeth N. Lack of change in urate deposition by dual-energy computed tomography among clinically stable patients with long-standing tophaceous gout: a prospective longitudinal study. Arthritis Res Ther 2013;15:R160. https://doi.org/10.1186/ar4343.
[6] Dalbeth N, Billington K, Doyle A, Frampton C, Tan P, Aati O, Allan J, Drake J, Horne A, Stamp LK. Effects of allopurinol dose escalation on bone erosion and urate volume in gout: a dual-energy computed tomography imaging study within a randomized, controlled trial. Arthritis Rheumatol 2019;71:1739–46. https://doi.org/10.1002/art.40929.
[7] Araujo EG, Bayat S, Petsch C, Englbrecht M, Faustini F, Kleyer A, Hueber AJ, Cavallaro A, Lell M, Dalbeth N, et al. Tophus resolution with pegloticase: a prospective dual-energy CT study. RMD Open 2015;1:e000075. https://doi.org/10.1136/rmdopen-2015-000075.
[8] Pascart T, Grandjean A, Capon B, Legrand J, Namane N, Ducoulombier V, Motte M, Vandecandelaere M, Luraschi H, Godart C, et al. Monosodium urate burden assessed with dual-energy computed tomography predicts the risk of flares in gout: a 12-month observational study : MSU burden and risk of gout flare. Arthritis Res Ther 2018;20:210. https://doi.org/10.1186/s13075-018-1714-9.
[9] Bayat S, Aati O, Rech J, Sapsford M, Cavallaro A, Lell M, Araujo E, Petsch C, Stamp LK, Schett G, et al. Development of a dual-energy computed tomography scoring system for measurement of urate deposition in gout. Arthritis Care Res 2016;68:769–75. https://doi.org/10.1002/acr.22754.
[10] Dalbeth N, Becce F, Botson JK, Zhao L, Kumar A. Dual-energy CT assessment of rapid monosodium urate depletion and bone erosion remodelling during pegloticase plus methotrexate co-therapy. Rheumatology 2022;61:4898–904. https://doi.org/10.1093/rheumatology/keac173.
[11] Uhlig T, Eskild T, Karoliussen LF, Sexton J, Kvien TK, Haavardsholm EA, Dalbeth N, Hammer HB. Two-year reduction of dual-energy CT urate depositions during a treat-to-target strategy in gout in the NOR-Gout longitudinal study. Rheumatology 2022;61. https://doi.org/10.1093/rheumatology/keab533. Si81-5.
[12] Chui CSK, Choi AKY, Lam MMY, Kwan TH, Li OC, Leng Y, Chow DLY. Volumetric reduction and dissolution prediction of monosodium urate crystal during urate-lowering therapy - a study using dual-energy computed tomography. Mod Rheumatol 2021;31:875–84. https://doi.org/10.1080/14397595.2020.1825053.
[13] Ellmann H, Bayat S, Araujo E, Manger B, Kleyer A, Cavallaro A, Lell M, Schenker H, Simon D, Tascilar K, et al. Effects of conventional uric acid-lowering therapy on monosodium urate crystal deposits. Arthritis Rheumatol 2020;72:150–6. https://doi.org/10.1002/art.41063.
[14] Zhang Z, Zhang X, Sun Y, Chen H, Kong X, Zhou J, Zhou Y, Ma L, Jiang L. New urate depositions on dual-energy computed tomography in gouty arthritis during urate-lowering therapy. Rheumatol Int 2017;37:1365–72. https://doi.org/10.1007/s00296-017-3724-3.
[15] Sun Y, Chen H, Zhang Z, Ma L, Zhou J, Zhou Y, Ding Y, Jin X, Jiang L. Dual-energy computed tomography for monitoring the effect of urate-lowering therapy in gouty arthritis. Int J Rheum Dis 2015;18:880–5. https://doi.org/10.1111/1756-185X.12634.
[16] Dalbeth N, Doyle AJ, Billington K, Gamble GD, Tan P, Latto K, Parshu Ram T, Narang R, Murdoch R, Bursill D, et al. Intensive serum urate lowering with oral urate-lowering therapy for erosive gout: a randomized double-blind controlled trial. Arthritis Rheumatol 2022;74:1059–69. https://doi.org/10.1002/art.42055.
[17] Mallinson PI, Coupal T, Reisinger C, Chou H, Munk PL, Nicolaou S, Ouellette H. Artifacts in dual-energy CT gout protocol: a review of 50 suspected cases with an artifact identification guide. AJR Am J Roentgenol 2014;203:W103–9. https://doi.org/10.2214/AJR.13.11396.
[18] Box GEP, Cox DR. An analysis of transformations. R Stat Soc Series B Stat Methodol 1964;26:211–52.
[19] FitzGerald JD, Dalbeth N, Mikuls T, Brignardello-Petersen R, Guyatt G, Abeles AM, Gelber AC, Harrold LR, Khanna D, King C, et al. 2020 American College of Rheumatology guideline for the management of gout. Arthritis Care Res 2020;72:744–60. https://doi.org/10.1002/acr.24180.
[20] Konishi S, Kitagawa G. Information criteria and statistical modeling. Springer; 2008.
[21] Cohen J. Statistical power analysis for the behavioral sciences and eii. Hillsdale NJ Eribaum; 1988.
[22] Lumley T, Diehr P, Emerson S, Chen L. The importance of the normality assumption in large public health data sets. Annu Rev Public Health 2002;23:151–69. https://doi.org/10.1146/annurev.publhealth.23.100901.140546.
[23] Bland JM, Altman DG. Statistics Notes: The use of transformation when comparing two means. Br Med J 1996;312:1153.
[24] Bland JM, Altman DG. Statistics notes: Transformations, means, and confidence intervals. Br Med J 1996;312:1079.
[25] Bland JM, Altman DG. Statistics Notes: Logarithms. Br Med J 1996;312:700.
[26] Dalbeth N, Nicolaou S, Baumgartner S, Hu J, Fung M, Choi HK. Presence of monosodium urate crystal deposition by dual-energy CT in patients with gout treated with allopurinol. Ann Rheum Dis 2018;77:364–70. https://doi.org/10.1136/annrheumdis-2017-212046.
[27] Jia E, Zhu J, Huang W, Chen X, Li J. Dual-energy computed tomography has limited diagnostic sensitivity for short-term gout. Clin Rheumatol 2018;37:773–7. https://doi.org/10.1007/s10067-017-3753-z.
[28] Zhang Y, Di A, Yuan H. Characteristics of monosodium urate crystal deposition in the foot in the different stages of gout by dual-energy computed tomography. Int J Rheum Dis 2022;25:147–53. https://doi.org/10.1111/1756-185X.14249.
[29] Zhang X, Tang CX, Zhou F, Lin PH, Yin CQ, Gao QY, Zhou LL, Zhou CS, Lu GM, Zhang LJ. Burden and distribution of MSU deposition in patients with hyperuricemia and gout: a large cohort Chinese population based DECT study. Res Square; 2022. https://doi.org/10.21203/rs.3.rs-1622091/v1.
[30] McQueen FMF, Doyle A, Reeves Q, Gamble GD, Dalbeth N. DECT urate deposits: now you see them, now you don't. Ann Rheum Dis 2013;72:458–9. https://doi.org/10.1136/annrheumdis-2012-202452.
[31] Miller J, DiMaso L, Huang-Vredevoogd J, Shah J, Lawless M. Characterization of size-specific effects during dual-energy CT material decomposition of non-iodine materials. J Appl Clin Med Phys 2021;22:168–76. https://doi.org/10.1002/acm2.13471.
[32] Chen H, Zhao L, Liu F, Chen S, Hu Z, Chen L, Ma Y, Guo K, Ji A, Merriman T, Min J. Urate in fingernail represents the deposition of urate burden in gout patients. Sci Rep 2020;10:15575. https://doi.org/10.1038/s41598-020-72505-6.
[33] Leng Y, et al. Distribution of urate crystal deposition in the hands and wrists of patients with chronic gout. Hong Kong J Radiol 2021;24:87–91. https://doi.org/10.12809/hkjr2117173.
[34] Gill RR, Richard WG, Yeap BY, Matsuoka S, Wolf AS, Gerbaudo VH, et al. Epithelial malignant pleural mesothelioma after extrapleural pneumonectomy: Stratification of survival with CT-derived tumor volume. Am J Roentgenol 2012;198:359–63. https://doi.org/10.2214/AJR.11.7015.
[35] Wise AM, Stamey TA, McNeal JE, Clayton JL. Morphologic and clinical significance of multifocal prostate cancers in radical prostatectomy specimens. Urology 2002;60:264–9. https://doi.org/10.1016/s0090-4295(02)01728-4.
[36] Ospel JM, Menon BK, Qiu W, Kashani N, Mayank A, Singh N, et al. A detailed analysis of infarct patterns and volumes at 24-hour noncontrast CT and diffusion-weighted MRI in acute ischemic stroke due to large vessel occlusion: results from the ESCAPE-NA1 trial. Radiology 2021;300:152–9. https://doi.org/10.1148/radiol.2021203964.
[37] Bowman FD, Guo Y, Derado G. Statistical approaches to functional neuroimaging data. Neuroimaging Clin N Am 2007;17:441–58. https://doi.org/10.1016/j.nic.2007.09.002.
[38] Just N: improving tumour heterogeneity MRI assessment with histograms. Br J Cancer 2014;111:2205–13. https://doi.org/10.1038/bjc.2014.512.
[39] Dalbeth N, House ME, Aati O, Tan P, Franklin C, Horne A, Gamble GD, Stamp LK, Doyle AJ, McQueen FM. Urate crystal deposition in asymptomatic hyperuricaemia

and symptomatic gout: a dual energy CT study. Ann Rheum Dis 2015;74:908–11. https://doi.org/10.1136/annrheumdis-2014-206397.

[40] Vickers AJ. The use of percentage change from baseline as an outcome in a controlled trial is statistically inefficient: a simulation study. BMC Med Res Methodol 2001;1:6. https://doi.org/10.1186/1471-2288-1-6.

[41] Senn S. Change from baseline and analysis of covariance revisited. Stat Med 2006; 25:4334–44. https://doi.org/10.1002/sim.2682.