



Cite this: *Nanoscale*, 2022, **14**, 4690

Versailles project on advanced materials and standards (VAMAS) interlaboratory study on measuring the number concentration of colloidal gold nanoparticles†

Caterina Minelli,^{ID} *^a Magdalena Wywijas,[‡]^a Dorota Bartczak,^{ID} ^b
 Susana Cuello-Núñez,^{ID} ^b Heidi Goenaga Infante,^{ID} ^b Jerome Deumer,^c
 Christian Gollwitzer,^c Michael Krumrey,^{ID} ^c Karen E. Murphy,^d
 Monique E. Johnson,^{ID} ^d Antonio R. Montoro Bustos,^{ID} ^d Ingo H. Streng,^{ID} [§]^d
 Bertrand Faure,^{ID} ^e Peter Høghøj,^e Vivian Tong,^{ID} ^a Loïc Burr,^f Karin Norling,^g
 Fredrik Höök,^{ID} ^g Matthias Roesslein,^h Jovana Kocic,^{ID} ⁱ Lyndsey Hendriks,^{||}
 Vikram Kestens,^{ID} ^j Yannic Ramaye,^j Maria C. Contreras Lopez,^j Guy Auclair,^j
 Dora Mehn,^k Douglas Gilliland,^k Annegret Potthoff,^l Kathrin Oelschlägel,^{||}
 Jutta Tentschert,^m Harald Jungnickel,^m Benjamin C. Krause,^{ID} ^m
 Yves U. Hachenberger,^m Philipp Reichardt,^m Andreas Luch,^{ID} ^m
 Thomas E. Whittaker,^{**}ⁿ Molly M. Stevens,ⁿ Shalini Gupta,^{ID} ^o Akash Singh,^{††}^o
 Fang-hsin Lin,^{ID} ^p Yi-Hung Liu,^p Anna Luisa Costa,^{ID} ^q Carlo Baldisserri,^q
 Rid Jawad,^r Samir E. L. Andaloussi,^r Margaret N. Holme,^{r,n} Tae Geol Lee,^s
 Minjeong Kwak,^s Jaeseok Kim,^s Johanna Ziebel,^t Cedric Guignard,^t
 Sebastien Cambier,^t Servane Contal,^t Arno C. Gutleb,^t Jan "Kuba" Tatarkiewicz,^u
 Bartłomiej J. Jankiewicz,^{ID} ^v Bartosz Bartosewicz,^{ID} ^v Xiaochun Wu,^{ID} ^w
 Jeffrey A. Fagan,^{ID} ^d Elisabeth Elje,^{x,y} Elise Rundén-Pran,^{ID} ^x Maria Dusinska,^x
 Inder Preet Kaur,^z David Price,^{aa} Ian Nesbitt,^{ab} Sarah O' Reilly,^{ab}
 Ruud J. B. Peters,^{ID} ^{ac} Guillaume Bucher,^{ad} Dennis Coleman,^{ae} Angela J. Harrison,^{ae}
 Antoine Ghanem,^{af} Anne Gering,^{af} Eileen McCarron,^{ag} Niamh Fitzgerald,^{ag}
 Geert Cornelis,^{ID} ^{ah} Jani Tuoriniemi,^{ah} Midori Sakai,^{ai} Hidehisa Tsuchida,^{ai}
 Ciarán Maguire,^{‡‡}^{aj} Adriele Prina-Mello,^{ID} ^{aj} Alan J. Lawlor,^{ak} Jessica Adams,^{ak}
 Carolin L. Schultz,^{ID} ^{al} Doru Constantin,^{am} Nguyen Thi Kim Thanh,^{ID} ^{an}
 Le Duc Tung,^{an} Luca Panariello,^{ID} ^{§§}^{ao} Spyridon Damilos,^{†‡}^{ao}
 Asterios Gavrilidis,^{ID} ^{ao} Iseult Lynch,^{ID} ^{ap} Benjamin Fryer,^{ap}
 Ana Carrasco Quevedo,^{ap} Emily Guggenheim,^{ap} Sophie Briffa,^{ap}
 Eugenia Valsami-Jones,^{ID} ^{ap} Yuxiong Huang,^{ID} ^{aq} Arturo A. Keller,^{ID} ^{aq}
 Virva-Tuuli Kinnunen,^{ar} Siiri Perämäki,^{ID} ^{ar} Zeljka Krpetic,^{ID} ^{as} Michael Greenwood^{as}
 and Alexander G. Shard^{ID} ^a

We describe the outcome of a large international interlaboratory study of the measurement of particle number concentration of colloidal nanoparticles, project 10 of the technical working area 34, "Nanoparticle Populations" of the Versailles Project on Advanced Materials and Standards (VAMAS). A total of 50 laboratories delivered results for the number concentration of 30 nm gold colloidal nanoparticles measured using particle tracking analysis (PTA), single particle inductively coupled plasma mass spectrometry (spICP-MS), ultraviolet-visible (UV-Vis) light spectroscopy, centrifugal liquid sedimentation (CLS) and small angle X-ray scattering (SAXS). The study provides quantitative data to evaluate the repeatability of these methods and their reproducibility in the measurement of number concentration of model nanoparticle systems following a common measurement protocol. We find that the population-averaging methods of SAXS, CLS and UV-Vis have high measurement repeatability and reproducibility, with between-labs variability of 2.6%, 11% and 1.4% respectively. However, results may be significantly biased

Received 24th November 2021,
 Accepted 11th February 2022

DOI: 10.1039/d1nr07775a

rsc.li/nanoscale



Nanoscale, 2022, 14, 4690–4704 | 4691

In this work, we undertook a large international interlaboratory comparison of five measurement methods for number-based particle concentration of colloidal gold suspensions under the umbrella of the Versailles Project on Advanced Materials and Standards (VAMAS), namely project 10 of the VAMAS technical working area (TWA) 34 (Nanoparticle Populations). VAMAS is an international organisation that supports world trade in products dependent on advanced materials technologies, through international collaborative projects aimed at providing the technical basis for harmonised measurements, testing, specifications, and standards. The lead organisation, the UK National Physical Laboratory, provided 54 laboratories across the world with colloidal gold nanoparticle samples, together with a common measurement protocol²⁴ and a reporting form in October 2017. Results were

The reference value of the number concentration of the colloidal gold test sample, used for comparability purposes of the results obtained under the VAMAS study reported here, was determined by LGC through characterisation of the same material independently of this study using spICP-MS with

Nanoscale, 2022, 14, 4690–4704 | 4693

CLS

CLS techniques quantify the separation of colloidal particles from a liquid under the presence of a centrifugal force. The particle migratory motion (settling) occurs in presence of a density contrast between the particles and the liquid. The driving force for the particle motion is the centrifugal force, which is opposed by the drag force. As a result, the settling velocity depends on particle characteristics such as size, density and shape. CLS instruments consist of either disc or cuvette centrifuges. However, all reports submitted to this study involved the use of a disc centrifuge, in which case the method is also known as differential centrifugal sedimentation (DCS). In these types of instruments,³¹ the sample is injected at the centre of the disc and sediments towards the edge. Its light extinction is detected close to the edge of the disc where a laser diode and a photodetector are aligned at opposite sides of the disc. Using Stokes' law and assumptions on the particle density and spherical geometry, light extinction-based particle size distributions are generated and converted to particle mass-based distributions using Mie light scattering theory. Participants were requested to report the experimental settings selected for the measurements and to record results assuming an effective particle density of 15.0 g cm^{-3} . We note that this value differs from that of gold's bulk density and is only valid for particles with the same size and shape as those used in this work. Reasons for this include that the value takes account of all the material enclosed within the particles Stokes volume, which also comprises citrate and solvent molecules bound and adsorbed at the surface of the gold particle. Participants were also instructed to weigh the syringe before and after each injection as to determine the mass of the injected dispersion. Presented results are normalised to the injection volume.

UV-Vis spectroscopy

UV-Vis spectroscopy measures the relative intensity of a monochromatic beam light passing through a sample for a range of radiation wavelengths. Particle number concentration is derived from the light absorption according to Beer–Lambert law and Mie theory.¹¹ Absorption spectra at wavelengths between 390 nm and 700 nm were acquired in triplicate with increments of 1 nm. Average absorption at 450 nm was ultimately chosen for the measurement of the particle number concentration, together with a recommended molar extinction coefficient of $(2.16 \pm 0.12) \times 10^9 \text{ M}^{-1} \text{ cm}^{-1}$.

PTA

In PTA, a diluted suspension of nanoparticles is illuminated with a laser light and the light scattered by individual particles is visualised with an optical microscope and recorded over time. The equivalent spherical hydrodynamic diameter of individual particles is determined by tracking them over time and measuring their translational diffusion coefficient, while the number concentration results from counting the particles in a known volume. Participants were requested to dilute the five

samples provided and perform 5 independent measurements, one for each vial. For each measurement, participants were instructed to prepare several aliquots according to their *in-house* procedure and acquire a minimum of 5 independent videos of 60 seconds for each aliquot under repeatability conditions. It was recommended to analyse a QC material at the beginning, in the middle and at the end of each measurement session. Because QC materials were not provided for the study, the participants who performed such measurements were asked to identify suitable materials and record relevant information.

spICP-MS

In spICP-MS, a diluted sample is introduced into the ICP-MS instrument at a set flow rate. A plasma atomises and ionises the particles producing an ion cloud for detection by mass spectrometry. The technique works by acquiring individual intensity readings with very short dwell times to ensure that signals from single particles can be resolved. With most current instrumentations, typical dwell times between 50 μs and 10 ms are used. To satisfy the single particle rule, the number of detected peaks should not exceed 10% of the maximum number of possible events per minute. As an example, using a dwell time of 3 ms, a maximum of 20 000 particles can be registered per minute, but it is recommended that the number of pulses in the time scan does not exceed ≈ 1200 per minute.^{32,33} The overall number of peaks detected per minute is directly proportional to the number of particles in the suspension. Participants were requested to dilute the five samples and perform 5 independent measurements, one for each vial. For each measurement, participants were instructed to prepare a number of aliquots according to their *in-house* procedure and acquire a minimum of 5 measurement replicates of 60 s scan duration per aliquot under repeatability conditions. In addition, participants were requested to determine the instrument transport efficiency by following the *in-house* practice, possibly at the start, middle and at the end of the measurement session to be able to assess the variability of this parameter. A known drawback in ICP-MS is the incomplete transport of solution from the nebulizer into the plasma. Consequently, the transport efficiency of the system needs to be determined in order to accurately equate the number of detected nanoparticles to the number concentration of the original sample. Several approaches exist,^{34,35} but none was specifically recommended against the local laboratory practice. The participants were requested to measure a QC material at the beginning, in the middle and at the end of each measurement session. Because no such materials were provided for the study, the participants who performed such measurements were asked to identify suitable materials and record relevant information.

Uncertainties

All results are expressed as average values and relative or absolute standard deviations ($n = 1$). As such, reported errors do not include a full uncertainty budget for the various methods.



SAXS measurements were performed by two laboratories, L44 and L45, as shown in Table 1. The instrumentations utilised by the two laboratories were significantly different, one using X-rays from a synchrotron radiation facility (L45) and the other one using a commercial instrument (L44). L45 used two silicon diodes (one in front and one behind the sample) to normalize the measured SAXS signal by the incoming photon flux to get the measured signal in absolute units, while L44 used the automatic normalization of the instrument, based on the simultaneous measurement of the transmitted intensity.



Table 1 Participant instrument details for the SAXS method

Laboratory code	X-ray source	Photon energy (keV)	Beam size (mm ²)	Detector	Pixel size (μm)	Type of sample holder	Sample thickness (mm)	Sample-detector distance (mm)
L44	Sealed tube	8.04 ± 0.02	0.8 × 0.8	Hybrid-pixel PILATUS 200k	172	Glass flow cell	1.4	938
L45	Synchrotron radiation (bending magnet)	8.0000 ± 0.0008	0.5 × 0.5	Hybrid-pixel PILATUS 1M	172	Glass capillary	1	4616.3 ± 0.6

L45 used a model fitting approach to convert the measured scattering intensities into number concentration, whereas L44 used an approach based on the *Expectation Maximization algorithm*^{36,37} to determine the number-based size distribution histogram.

Taking the energy-dependence of the electron density contrast into account, the mean concentration values for SAXS resulted in $1.33 \times 10^{14} \text{ kg}^{-1}$ (L45) and $1.38 \times 10^{14} \text{ kg}^{-1}$ (L44), which deviated from the reference value by 9.5% (L45) and 6.1% (L44) respectively. Despite differences

in the experimental set-up and the evaluation procedure, the resulting average number concentration values were within 3%, with a measurement repeatability standard deviation of 5% of the undiluted samples for both laboratories (Fig. 2A). The 3% difference can be explained by experimental uncertainties, or differences of the number concentration of the samples respectively. Both laboratories used a different method to analyse the scattering data, which may also cause some variations in the concentration values.

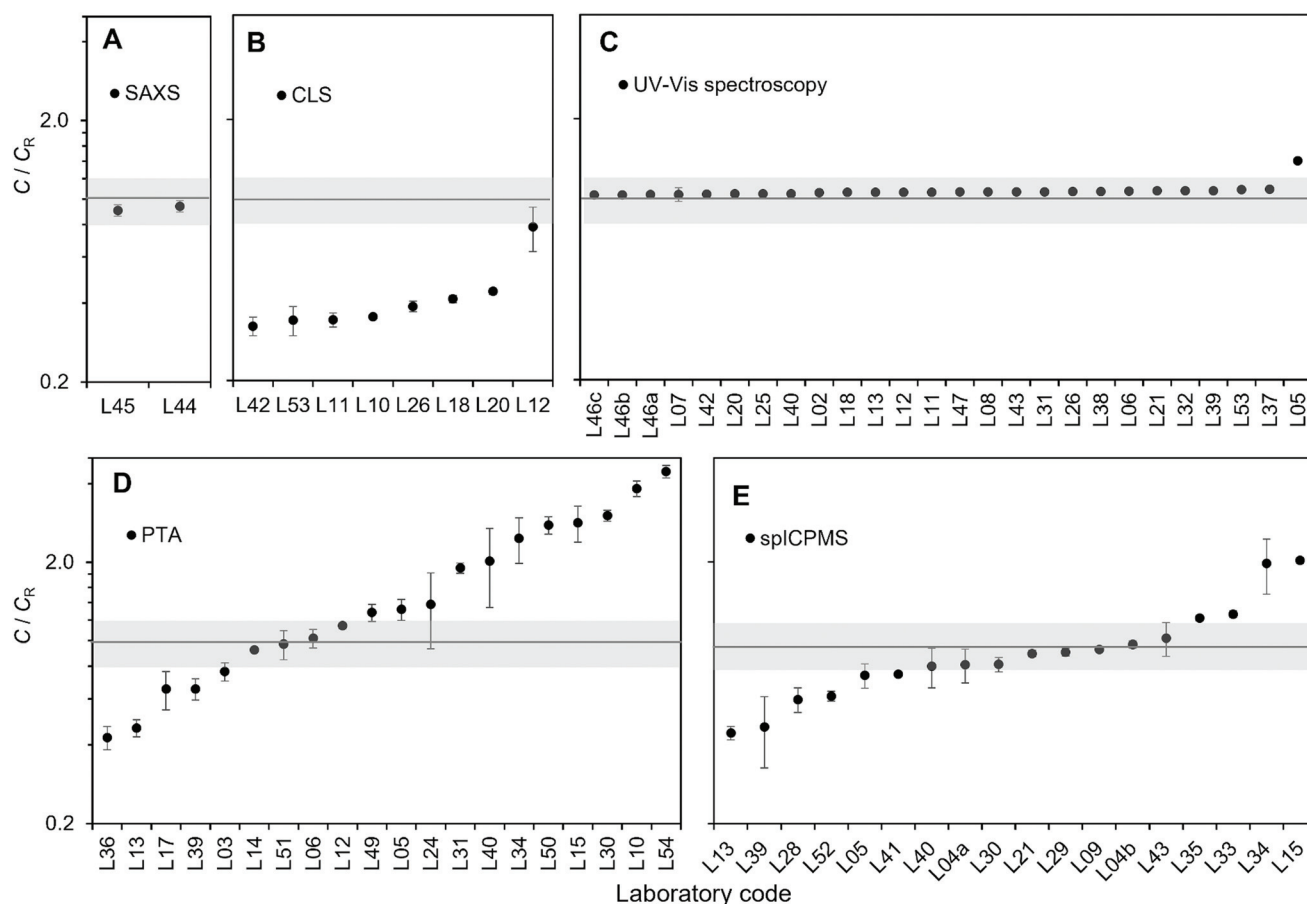


Fig. 2 Results of the number concentration C of the test material as measured by the laboratories participating in the study by the methods of (A) SAXS, (B) CLS, (C) UV-Vis spectroscopy, (D) PTA and (E) spICP-MS expressed relative to the accepted number concentration reference value C_R . The X-axis shows the laboratory codes. Error bars show the repeatability standard deviation (in-lab variability). The grey box shows the 95% confidence interval of the assigned reference value for the number concentration of the test material.



Population-averaging methods: UV-Vis spectroscopy

Twenty-three laboratories participating in the study performed measurements using the UV-Vis method. Samples were measured both undiluted and diluted by a factor 5 and 10 (Table 3). Most laboratories performed gravimetric dilutions, except for three participants who used volumetric dilutions. Results for the undiluted samples are shown in Fig. 2C. The average measurement repeatability within each laboratory was excellent with repeatability standard deviations below 0.5%. Excluding laboratory L05, between-labs variability was of 1.4%. Laboratory L05 detected problems in the instrument photometer shortly after the study and it is likely that these problems were also present at the time of the study.

These low variabilities are particularly significant if we take into account the large number of different instrument models and instrument set-ups that were used, for example including both single and dual beam configurations.

The results reported here for the measurement of particle number concentration are based on the absorption of the colloidal suspension measured at a wavelength of 450 nm. It is found in literature that the absorption at the maximum of the localised surface plasmon resonance peak (LSPR) is at times used for the measurement of the particle concentration.²⁵ The between-labs variability of this value resulted in 1.9% *i.e.* using the absorption at 450 nm delivered better method reproducibility. It should be noted that the reproducibility standard deviation of the wavelength of the LSPR maximum was within 0.11% across the laboratories, which is also a good indication of the quality and integrity of the sample product.

The mean UV-Vis method result was 4.4% above the reference value of the test material. As in the case of CLS, this bias depends on the accuracy of certain parameter used in the calculations, specifically the molar extinction coefficient of the particles at 450 nm. The most significant challenge of using UV-Vis for the measurement of particle number concentration resides in the knowledge of the particle molar extinction coefficient. This can be calculated using Mie theory, for which size, shape and optical constants are required.¹¹ Alternatively, it can be measured by absorption spectroscopy based on the knowl-

Laboratory code	Gravimetric dilutions	Instrument model and laser wavelength	Calibrant	Rotational speed (rpm)	Sucrose temperature before and after injection (°C)
L10	Yes	CPS Instruments DC24000 UHR, 405 nm	264 nm PVC, CPS Instruments Lot#123	20 000	n.r. ^a –31.5
L11	Yes	CPS Instruments DC24000 UHR, 405 nm	237 nm PVC, CPS Instruments Lot#150	18 000	24.9–34.9
L12	Yes	CPS Instruments DC24000 UHR, 405 nm	237 nm PVC, CPS Instruments Lot#150	24 000	23.6–29.1
L18	Yes	CPS Instruments DC24000 UHR, 405 nm	145 nm silica, CPS Instruments Lot#148	20 065	21.0–26.5
L20	Yes	CPS Instruments DC24000 UHR, 405 nm	145 nm silica, China University of Petroleum, Lot#GBW(E)120059	20 000	16.6–16.6
L26	Yes	CPS Instruments DC24000 UHR, 405 nm	237 nm PVC, CPS Instruments Lot#105	24 000	20.7–33.9
L42	n.a.	CPS Instruments DC24000 UHR, 405 nm	476 nm PVC, Analytik UK	20 000	n.r.
L53	Yes	CPS Instruments DC24000 UHR, 405 nm	179 nm silica, nanoComposix, Lot#JEA0224	24 000	28.0–37.6

^a n.r. = not recorded.

Laboratory code	Gravimetric dilutions	Instrument model and light source	Cuvette	Temperature (°C)
L02	Yes	Analytik Jena Specord Plus 200, deuterium/halogen, dual beam	Brandt 759081D, PMMA, 1 cm, 2.5 mL	Room temp.
L05	Yes	Varian Cary 50 Bio, xenon, dual beam	Ratiolab 2712120, PS, 1 cm, 1 mL	20
L06	Yes	Hach Lange DR5000, deuterium/halogen, single beam	Sarstedt 67.754, PS, 1 cm, 2 mL	22.0
L07	Yes	Shimadzu UV-2600, deuterium, dual beam	Greiner bio-one, polypropylene, 1 cm, 1 mL	27
L08	Yes	Hach Lange DR 3900, halogen, single beam	ONDA 000302 eng, quartz, 1 cm, 4 mL	26
L11	Yes	Jasco V 570, deuterium/halogen, dual beam	Brand macro, PMMA, 1 cm, 2 mL	24.2
L12	Yes	Shimadzu UV-1800, tungsten/halogen, dual beam	Hellma Analytics 111-QS, quartz, 1 cm, 2 mL	24
L13	Yes	Agilent Technologies 8453, deuterium/tungsten, single beam	Hellma Analytics Suprasil, quartz, 1 cm, 2 mL	25
L18	Yes	PerkinElmer Lambda 650, deuterium/tungsten, dual beam	Hellma Analytics QX High Precision Cell, quartz, 1 cm, 3 mL	23.3
L20	Yes	Agilent Technologies Cary 60 UV Vis, xenon, single beam	N.k., Glass, 1 cm, 2 mL	26
L21	Yes	Agilent Technologies Cary 5000, deuterium/halogen, dual beam	NSG precision cell, quartz, 1 cm, 2 mL	20.7
L25	No	Agilent Technologies Cary 8454, deuterium, single beam	N.k., quartz, 1 cm, 3 mL	20
L26	Yes	PerkinElmer Lambda 850, deuterium/tungsten/halogen, dual beam	Hellma Analytics 104F-10-K-40, quartz, 1 cm, 1.4 mL	Room temp.
L31	Yes	Jasco V-650, deuterium/tungsten/halogen, dual beam	Signa Aldrich z276669, quartz, 1 cm, 3 mL	21.0
L32	Yes	Shimadzu UV-3600, deuterium/halogen, dual beam	Hellma Analytics Suprasil, quartz, 1 cm, 3.5 cm	23.7
L37	Yes	Molecular Devices Spectramax M2, single beam	Fisher FB55143, plastic, 1 cm, 3 mL	24.5
L38	Yes	Agilent Technologies Cary 5000, deuterium/halogen, dual beam	VWR 634-0675, PS, 1 cm, 2.5 mL	18
L39	No	Jenway 6800, dual beam	Sarstedt, PS, 1 cm, 3 mL	25
L40	No	Shimadzu 1800, deuterium and tungsten, dual beam	Fisherbrand, PS, 1 cm, 3 mL	23
L42	n.a.	Agilent Technologies Cary 300, tungsten/deuterium, dual beam	Hellma, quartz, 1 cm, 1.5 mL	20
L43	Yes	PerkinElmer Lambda 35, deuterium/halogen, dual beam	Hellma, quartz, 1 cm, 4 mL	25
L46a	Yes	Biotech Synergy MX xenon, single beam, fixed bandwidth 5 nm	BrandTech Sci 759105, PMMA 1 cm, 2 mL	21.7
L46b	Yes	Thermo Spectronic Unicam UV540, tungsten/deuterium, dual beam, bandwidth 1 nm	BrandTech Sci 759105, PMMA 1 cm, 2 mL	21.7
L46c	Yes	Thermo Spectronic Unicam UV540, tungsten/deuterium, dual beam, bandwidth 4 nm	BrandTech Sci 759105, PMMA 1 cm, 2 mL	21.7
L47	Yes	Shimadzu UV2450, deuterium/tungsten/halogen, dual beam	Hellma Analytics, quartz, 1 cm, 0.8 mL	21.4
L53	Yes	Ocean Optics DT-MINI-2-GS 2000+ UV-VIS-ES, deuterium/tungsten/halpgen, dual beam	Brand UV-cuvette macro, proprietary plastic, 1 cm, 2 mL	28

In-lab variability varied between 1.2% and 33.6%, with an average repeatability standard deviation of 11.3%. However, measurement reproducibility across laboratories was poor, with 68% between-labs variability in the final number concentration results. The reasons for such a high variability in results may be multiple. Some laboratories invested great effort in minimising and quantifying the number concentration of background particles present in the dispersant used for the dilutions, but such efforts were not consistent across participants. The type of light source used to illuminate the sample in conjunction with the camera sensitivity may affect the minimum particle size detectable by the instrument, which would affect particle count, as it is the case also for image noise. The reliance of the PTA methods on the manual setting of camera levels and instrument software thresholds for the analysis of the images introduces a level of subjectivity in the measurements. No specific correlation or trend of the measurement results with experimental parameters such as camera level or instrument wavelength were identified.

^a grav. = gravitational; vol. = volumetric. ^b (U) = Malvern Panalytical software with concentration upgrade included.

Nanoscale, 2022, 14, 4690–4704 | 4699

Table 5 Participant instrument details for the splCP-MS method

Laboratory code	Average dilution factor ^a	Instrument model	Average flow rate (mL min ⁻¹)	Transport efficiency (%)	Dwell time (ms)	QC material
L04a	62 310 000 (grav.)	TOFWERK icpTOF 2R	0.409	7.55	1.027	NIST RM 8013 60 nm Au NPs
L04b	40 030 000 (grav.)	Thermo ElementXR	0.089	10.8	1	NIST RM 8013 60 nm Au NPs
L05	5 373 000 (grav.)	Thermo XSeries II	0.513	5.40	3	NIST RM 8013 60 nm Au NPs
L09	3 979 000 (grav.)	PerkinElmer NexION 350 D	0.324	10.3	0.05	BBI solution 30 nm Au NPs
L13	1 088 000 (grav.)	PerkinElmer NexION 300 D	0.238	11.2	0.05	NIST RM 8013 60 nm Au NPs
L15	4 910 000 (grav.)	Agilent ICP-MS 7900	0.346	3.9	0.1	Nanocomposix Gold BioPure 30 nm Au NPs
L21	6 029 717 (grav.)	PerkinElmer NexION 350 D	0.0963	13.5	0.05	NIST RM 8012 30 nm Au NPs
L28	1 053 000 (grav.)	PerkinElmer NexION 2000	0.209	15.5	0.05	BBI solution 30 nm Au NPs
L29	2 796 000 (grav.)	PerkinElmer NexION 300	0.343	5.58	0.1	Nanocomposix Gold BioPure 60 nm Au NPs
L30	4 913 000 (grav.)	Thermo Icap Q	0.346	6.45	3	Nanocomposix Gold BioPure 60 nm Au NPs
L33	2 500 000 (vol.)	PerkinElmer NexION 350	9	6.97	0.08	PerkinElmer N18142303
L34	3 237 000 (grav.)	PerkinElmer NexION 350 D	0.260	2.7	0.05	BBI solution 60 nm Au NPs
L35	989 000 (grav.)	PerkinElmer NexION 350 S	0.215	7.1	0.05	NIST RM 8013 60 nm Au NPs
L39	1 092 000 (grav.)	PerkinElmer NexION 350 X	0.450	2.90	0.05	None
L40	5 000 000 (vol.)	Agilent 7900	0.346	6.57	0.1	Not reported
L41	1 807 000 (grav.)	PerkinElmer NexION 350 D	0.306	11.6	0.1	Nanocomposix Au NPs
L43	977 197 (grav.)	Agilent 7900	0.340	4.2	0.1	BBI solution 60 nm Au NPs
L52	1 020 250 (grav.)	PerkinElmer NexION 300 X	0.165	11	0.1	BBI solution 60 nm Au NPs

^a grav. = gravitational; vol. = volumetric.

calibrated the instrument transport efficiency using a nanoparticle material (Table 5, last column). Examination of the results with respect to the use of a reference material (*i.e.* RM 8012 or RM 8013) or a QC material from other commercial vendors, did not reveal any trends. Participating laboratories were free to choose the most appropriate transport efficiency calibration method and were not required to provide the details of which method was used.

Effect of dilution

Part of this study was designed to evaluate best practice in sample dilution and the impact this can have in measurement outcome and reproducibility. Fig. 3A compares the results of the measurements performed on the diluted samples with those of the undiluted samples for the population-averaging methods. For most laboratories, a deviation was observed, which was found to be within 20%. No significant trend was observed for laboratories that chose a gravimetric approach to sample dilution rather than volumetric dilution, although it is not possible to conduct a systematic analysis given the disparity in the number of laboratories that chose the two approaches.

For the UV-Vis method, the average repeatability standard deviation changed from below 0.5% to 0.6% and 1.2% for samples prepared with a dilution factor of 5 and 10 respectively. This result may reflect the analytical precision which depends upon signal intensity. In terms of measurement reproducibility, the between-labs variability changed from 1.4% to about 6% irrespective of the dilution factor. Here, we see the impact of the dilution step in the overall reproducibility of the measurement. A similar trend was observed for the CLS method. The average measurement repeatability in fact

slightly improved, the standard deviations varying from 5% to 4%, possibly indicating that the diluted samples had a more optimal concentration for the instrument. However, measurement reproducibility across the laboratories worsened, with between-labs variability changing from 11% to 16% and 19% for the measurements performed on the samples diluted a factor 5 and 10 respectively.

The laboratories that participated with particle-counting methods were asked to optimise the concentration of the sample according to the laboratory best practice. The chosen concentration varied across two orders of magnitude for both methods, with no significant resulting trend on measurement bias. For those laboratories that reported a measurement result deviating more than 20% from the accepted reference value of the number concentration of the test material, the dilution factor did not appear to correlate with the bias. We note, however, that for the lower dilution factors, PTA measurement results had negative bias. This effect is attributed to overlapping of particle signals during imaging.

Discussion

Fig. 4 summarises the results of this interlaboratory study in terms of method comparability and reproducibility. The results highlighted clear distinctions between population-averaging and particle-counting methods.

A first distinction is that the population-averaging methods show better agreement between laboratories than the particle-counting methods. For example, the reproducibility of the UV-Vis method is noteworthy, with results across 24 instruments in agreement within 1.4%. This is notwithstanding



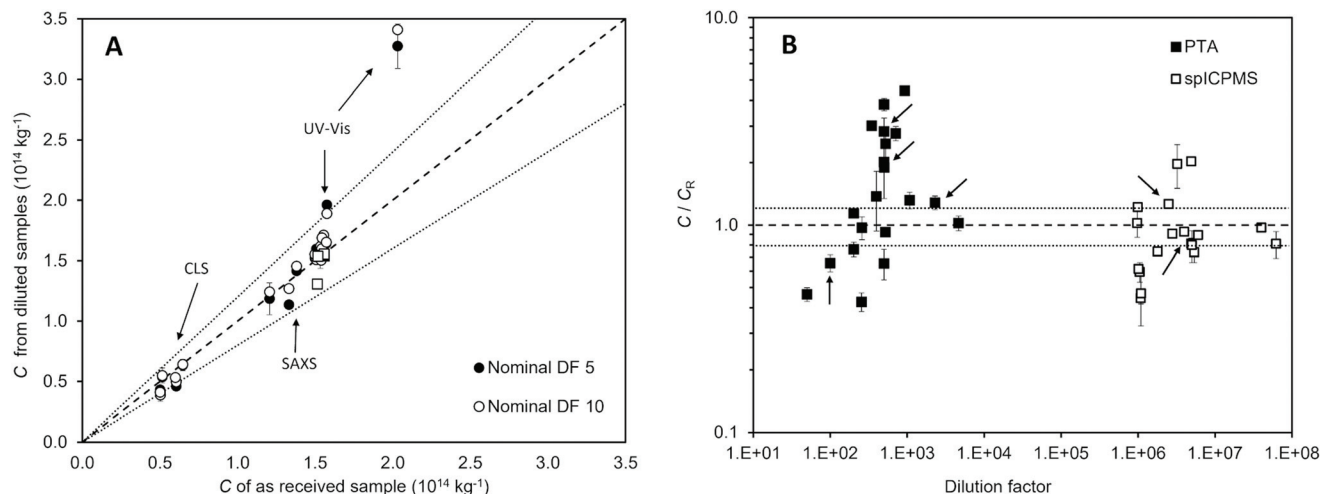


Fig. 3 (A) Population-averaging methods: raw number concentration C of test materials from measurements of the samples diluted 5 (black marker) and 10 (white marker) times as a function of the number concentration of the undiluted sample as measured by the same laboratory using the same method. Circles relate to dilutions performed gravimetrically and squares to those performed volumetrically. (B) Particle-counting methods: raw number concentration C from measurements of the diluted aliquots relative to the accepted reference value for the test materials concentration C_R as a function of the dilution factor. The arrows indicate the results from samples diluted volumetrically. In both figures, the dashed line is the identity relationship and the dotted lines show a deviation of 20%.

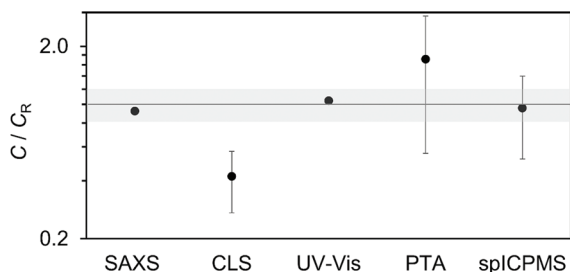


Fig. 4 Results of the number concentration C of the test material as measured by the 5 methods investigated in the VAMAS TWA 34 project 10 interlaboratory study relative to the accepted reference value C_R . Error bars show method between-labs variability (reproducibility standard deviation). The grey box shows the 95% confidence interval of the assigned reference value for the number concentration of the test material.

instruments being from 12 different manufacturers and utilising both single and dual beam configurations, and experimental set-ups differing from each other, for example in the type of sample cuvette that was used to hold the sample. This outcome is likely to arise from the robustness of the method and the minimal sample preparation required for the measurements. The robustness reflects the maturity of the technology underpinning the method, combined with the relatively straightforward protocol, which could be reduced to two light transmission measurements (the blank and the sample) at one single wavelength (*i.e.* 450 nm). CLS, instead, relies on a more complex technology and requires calibration. Nonetheless, a RDS value of 11% across seven laboratories was achieved. This consistency may, in part, be due to all the instruments in the

study being made by the same manufacturer. SAXS, as a population-averaging method, was performed with a lab-based and a synchrotron-based experimental set-up. The evaluated values of the mean number concentration of both set-ups only differ by 3%. This can be considered a very promising result, considering that both laboratories applied 2 different algorithms to analyse the scattering data. Similarly to UV-vis, this can be attributed to the simple sample preparation and maturity of the technique.

In contrast, the measurement reproducibility of particle-counting methods was significantly poorer across laboratories showing RSD values of 46% and 68% for the spICP-MS and PTA method respectively, despite average repeatability within each laboratory being below 12% RSD for both methods. We could not identify a trend between instrument models and measurement results and concluded that the variability is more likely due to differences between individual instruments, procedures and data processing across the laboratories. Some laboratories described at length the experimental protocol utilised to minimise potential sources of measurement uncertainty. For example, in the case of spICP-MS, some laboratories detailed how the transport efficiency was determined, the cut off value for distinguishing the nanoparticle from the continuous analyte signals identified and the background correction applied, but these practices did not appear to be in place in all laboratories. This is expected given the diverse background of the participating laboratories: while national measurement institutes may be experts in the application and standardisation of the methods, other laboratories may be using them as an internal QC and therefore be mainly concerned with method repeatability. We note that both PTA and spICP-MS may require the setting of thresholds for the discrimination of

uncertainty in materials' refractive index. However, particles are rarely spherical and something as common as the presence of a coating may alter their optical properties. This argument is even stronger for heterogeneous samples, for example containing polydisperse or diverse particle populations. The particle-counting methods would still hold the potential for accurate measurements, while the population-averaging methods would require substantial modelling to interpret the results. Where measurement accuracy in number concentration is required and particle-counting methods cannot be utilised, there is a need to advance methods for the measurement of other particle attributes, primarily the optical properties such as the complex refractive index.

As far as CLS is concerned, the lack of full knowledge of the experimental parameters that enter the underpinning model, such as the optical properties of the particles and the fluid, does not fully explain the bias in the measurement results. It is likely that particle losses within the system, for example at the injection port or by absorption onto the disc and syringe walls, also contribute to this bias. In a similar fashion to spICP-MS, the CLS method could be improved by performing a calibration of the transport efficiency. A number of factors would require evaluation for the design and selection of particle calibrants suitable for specific applications, including their bulk and surface chemistry. Particle reference materials with information on their mass concentration, and more recently on their number concentration, appear useful to this purpose. It is envisaged that implementing a calibration of the “transport efficiency” within the CLS measurement protocol could result in this method being both precise and more accurate.

More information on the performance of methods for the measurement of particle number concentration will be available through the pilot study P194 of the Consultative Committee for Amount of Substance: Metrology in Chemistry and Biology (CCQM) of the International Bureau of Weights and measurements (BIPM) that was led by LGC and ran in parallel with this VAMAS study. P194 consisted of an interlaboratory comparison among the international community of national measurement institutes of the measurement of colloidal number concentration of the same batch of particles utilised for the VAMAS study. While the final outcome of the P194 pilot study has not been formally announced, preliminary results indicated that the accepted reference value for the number concentration of the test sample used in this VAMAS study is within 3.5% of the number concentration measurement resulting from the pilot study.

Overall, this study indicates that the choice of the method to adopt for the measurement of the number concentration of colloidal particles primarily depends on the field of application. If the measurements are part of a QC/QA framework, precision is probably the driving criterion for the choice. However, where accuracy is required, the choice may be different. There are certainly other types of considerations to take into account: not all methods are suitable for all materials or particle size ranges; materials may become unstable upon dilution, in which case methods requiring minimal sample preparation are preferred; samples may be expensive and limited in volume, in which case

methods requiring high dilutions may be the only viable approach; samples may be susceptible to agglomeration, in which case methods with high size resolution are preferred.⁴ Furthermore, environmentally relevant samples contain nanoparticles in a concentration range from 10^{-3} g kg⁻¹ to 10^{-6} g kg⁻¹, meaning PTA and spICP-MS may be suitable for measuring them, but not UV-Vis nor CLS. In this respect, it is important for the community to have a range of methods of choice available. Studies like the current one are then important to evaluate method comparability and provide a means to validate different measurement approaches. This is particularly important for regulatory purposes. For the laboratories that took part in the study, this project also provided a means to benchmark their measurement capability. Furthermore, through this study some laboratories were able to identify problems with their equipment or update their operating procedure. The study was also useful to inform the community of the availability of the QC material LGCQC5050, the first of its kind (*i.e.* with an assigned value for the number concentration of the particles) and whose use will improve measurement confidence and outcome. For instrument manufacturers, the study provided an independent benchmarking and validation of their instrument models for the measurement of nanoparticle number concentration. Furthermore, the study highlighted areas for potential improvement of measurement outcome such as, for example, automation in sample handling and signal processing and improved calibration routines. The trust that the community have in such capabilities has grown as a result of this study and the knowledge developed through the preparation, delivery and outcome of the study are currently informing the development of new documentary standards and reference materials.

There is a need to extend these types of studies to a broader selection of particles that mimic more closely the attributes of real nanoparticle-based products. For example, the nanoparticle user community would benefit from the evaluation of current measurement technology for the measurement of the number concentration of (i) nanoparticles below 10 nm, which is a size range close or below the limit of detection of many liquid-based methods, but relevant for products ranging from quantum dots to viral vectors; (ii) polydisperse particle samples, with size distributions that are closer to real manufactured particle products; (iii) particles with refractive index close to that of water, for example of biological nature, with poor light scattering properties and low mass; (iv) multimodal samples, among which samples that contain populations with average sizes both in the nanometre and the micrometre size range present significant technical challenges; (v) non-spherical particles.

Conclusions

This interlaboratory study of the measurement of number concentration of colloidal nanoparticles carried out as project 10 of VAMAS TWA 34 provided the stakeholder community with important quantitative information for the comparative evaluation and validation of available measurement methods.

Broadly, particle-counting methods were found to be potentially less biased than the population-averaging methods, but the latter had superior reproducibility. We discussed some of the underpinning motivations for this distinction. Importantly, we highlighted the impact on measurement outcome of experimental practice in both sample preparation and measurement execution. Despite the nanoparticles used in this study being to some extent an ideal sample, it is significant that they are currently commercially available to the user community for comparative testing. This study thus provides a substantial dataset to build confidence in the methods available for the measurement of the number concentration of colloidal samples and contributes to optimise laboratory practice and inform new documentary standards.

Author contributions

CM contributed to conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, resources, supervision, validation, visualization, writing the original draft, review and editing the manuscript. MW contributed to data curation, investigation, methodology, resources and validation. DB, SCN and HGI contributed to data curation, formal analysis, investigation, methodology, resources, validation, writing the original draft, review and editing the manuscript. JD, CG and MK, KEM, MEJ and ARMB contributed to data curation, formal analysis, investigation, methodology, validation, writing the original draft, review and editing the manuscript. AGS contributed to conceptualization, formal analysis, funding acquisition, methodology, supervision, writing the original draft, review and editing the manuscript. All other authors contributed to investigation, validation and review and editing the manuscript.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This work was partially funded by the 14IND12 Innanopart and 18SIP01 ISOCONCur projects of the European Union through the European Metrology Programme for Innovation and Research (EMPIR). EMPIR is jointly funded by the EMPIR participating countries within EURAMET and the European Union. The work was also funded by the UK Department of Business, Energy and Industrial Strategy through the National Measurement System programme (projects NMS/ST18-21). NILU acknowledge funding from the EU H2020 project Nanoreg2 (Grant agreement No. 646221). The University of Birmingham acknowledge funding from the EU H2020 project ACEnano (Grant agreement No. 720952). The authors thank all the people that have been involved in the delivery of this large interlaboratory study and may not be included in the list of authors.



