Possibilities and limitations of current technologies for quantification of biological extracellular vesicles and synthetic mimics

Sybren L.N. Maas a,b, Jeroen de Vrij a,b, Els J. van der Vlist c, Biaina Geragousian a,b, Louis van Bloois d, Enrico Mastrobattista d, Raymond M. Schifflers e, Marca H.M. Wauben c, Marike L.D. Broekman a,b, Esther N.M. Nolte-'t Hoen c,*

a Department of Neurosurgery, University Medical Center Utrecht, The Netherlands
b Department of Pharmaceutics, Utrecht Institute for Pharmaceutical Sciences, Utrecht University, The Netherlands.
c Department of Biochemistry & Cell Biology, Faculty of Veterinary Medicine, Utrecht University, The Netherlands
d Department of Pharmaceutics, Utrecht Institute for Pharmaceutical Sciences, Utrecht University, Utrecht, The Netherlands.
e Department of Clinical Chemistry and Hematology, University Medical Center Utrecht, The Netherlands.

ABSTRACT

Nano-sized extracellular vesicles (EVs) released by various cell types play important roles in a plethora of (patho)physiological processes and are increasingly recognized as biomarkers for disease. In addition, engineered EV and EV-inspired liposomes hold great potential as drug delivery systems. Major technologies developed for high-throughput analysis of individual EV include nanoparticle tracking analysis (NTA), tunable resistive pulse sensing (tRPS) and high-resolution flow cytometry (hFC). Currently, there is a need for comparative studies on the available technologies to improve standardization of vesicle analysis in diagnostic or therapeutic settings.

We investigated the possibilities, limitations and comparability of NTA, tRPS and hFC for analysis of tumor cell-derived EVs and synthetic mimics (i.e. differently sized liposomes). NTA and tRPS instrument settings were identified that significantly affected the quantification of these particles. Furthermore, we detailed the differences in absolute quantification of EVs and liposomes using the three technologies. This study increases our understanding of possibilities and pitfalls of NTA, tRPS and hFC, which will benefit standardized and large-scale clinical application of (engineered) EVs and EV-mimics in the future.

#1. Introduction

Extracellular vesicles (EVs) are lipid membrane-enclosed vesicles released by cells and present in bodily fluids. EVs are heterogeneous in composition and size, ranging from approximately 50 to 1000 nm, with the vast majority <200 nm in size [1,2]. EVs originate from their donor cell as a result of outward budding of the plasma membrane. Alternatively, EVs form as a result of intracellular budding within late endosomes, from which vesicles are released upon fusion of these multivesicular bodies with the plasma membrane [3]. Regardless of their size and origin, ‘EVs’ is the collective term adopted to designate any type of cell-derived vesicle in the extracellular space. In recent years, multiple reports have demonstrated EVs to play an important role in (patho)physiological processes, such as immune responses [4], blood coagulation [5], tissue repair [6] and tumor growth [7,8]. Current research focuses on obtaining improved insight into the formation and function of EVs and on studying the potential of EVs for medical applications. One of these applications is to use EVs present in body fluids as biomarkers for diagnosis and monitoring of diseases [9,10]. In cancer, tumor-derived EVs can serve as biomarkers since they contain proteins and RNAs from their malignant donor cells [7,8]. Since tumor-derived EVs are released in easily accessible bodily fluids, such as blood or urine [7,11], analysis of these EVs for disease monitoring may circumvent biopsies [11], thereby reducing biopsy related morbidity and mortality. A second important application of EV in the medical field is their use as drug delivery systems. Although liposomes, which share the bilayered membrane structure with EVs, have been employed as drug delivery systems for many years, cross-pollination of knowledge in the lipidome and EV research fields now holds high promise for improvement of current delivery systems. Various studies have indicated that EVs can be exploited as carriers for delivery of exogenous therapeutic cargoes, e.g. siRNAs, in vivo [12]. EV characteristics that facilitate efficient delivery of biological drugs include their capacity to traverse intact biological barriers (e.g. blood–brain barrier) and to deliver functional RNA into cells, as well as their stability in blood (reviewed in [13]).
Current research focuses on exploiting these features to either engineer natural EV for drug delivery to specific tissues, or to design EV mimics formulated as liposomes containing relevant EV components [14].

Even though EVs are increasingly recognized as important biological and therapeutic entities, standardized methods for their analysis are still lacking [15]. Establishment of such methods is crucial for safe application of (engineered) EV in clinical practice, but EV quantification has proven technically difficult due to the small size of EVs and their heterogeneity in size and composition.

In recent years, several instruments have become available that allow detection and characterization of individual EVs. These techniques include nanoparticle tracking analysis (NTA) [16,17], tunable resistive pulse sensing (tRPS) [18] and high-resolution flow cytometry (hFC) [19]. EV detection and quantification with these single-particle analysis techniques rely on distinct principles. NTA is based on the illumination of particles in suspension with a laser beam, followed by the recording of the scattered light by a light-microscope. The Brownian motion of each particle is individually tracked to determine the mean square displacement of the individual particle. Since temperature and viscosity of the suspension are known and controlled, the Stokes–Einstein equation can be used to determine the hydrodynamic diameter of each individual particle. The total number of particles is used for particle concentration estimation [16,20]. In tRPS, a non-conductive polyurethane membrane, punctured to contain a single opening, separates two fluid cells [21]. By applying a voltage across the membrane a flow of ions is induced. Once a particle moves through the nanopore, the flow of ions is altered resulting in a brief “resistive pulse” which is recorded by the instrument [22]. The size-distribution [23] and concentration [24,25] of particles can be calculated by referring to the observed pulse height and rate to pulses induced by reference particles of known volume and concentration. Flow cytometric analysis of particles involves the sequential excitation of individual, fluorescently labeled particles in a liquid stream and detection of emitted light by diodes or photomultipliers [26]. In hFC, a high-end flow cytometer is optimized for the analysis of nano-particles. This optimization consists of light scattering detection at customized angles, the usage of high power lasers and high-performance photomultiplier tubes for more sensitive light detection, and application of fluorescence-based thresholding to distinguish particles of interest from noise signals [19]. In-depth description of the technical backgrounds of the techniques is beyond the scope of this manuscript and described elsewhere for NTA [16,20,27,28], tRPS [22–24] and hFC [19,29].

For accurate EV quantification and characterization, it is important to know to what extent instrument-specific variables influence particle characterization. For NTA, studies on how instrument settings affect the analysis of heterogeneous EV populations are limited [20,28,17], and the effects of specific variables on EV quantification and size-profiling by tRPS are largely unknown. For hFC, detailed reports on optimizing the instrument configuration and settings for accurate analysis of EVs and other nano-sized particles have recently been published [19,29]. In a few studies, two or three of the above described techniques have been compared. However, these studies either focused on size-profiling of synthetic beads [30,31], or did not address effects of instrument settings on EV characterization and quantification [32,33].

Here, we report a comprehensive comparative study on NTA, tRPS and hFC for analysis of populations of heterogeneous nano-sized EVs and synthetic mimics (i.e. polystyrene beads and calcine-loaded liposomes). We identified different NTA- and tRPS-variables that significantly influenced the quantification of these particles. Furthermore, we assessed the comparability of NTA, tRPS and hFC in absolute quantification of liposomes and EVs. Based on these data, we stress the importance of technical knowledge of the instruments, awareness of analytical variables, and recognition of how instrument settings affect measurements when analyzing EV populations with unknown concentration and size heterogeneity.

### 2. Materials and methods

#### 2.1. Polystyrene beads

115 and 203 nm polystyrene beads (Izon Science, Christchurch, New Zealand) were analyzed using tRPS and NTA. For hFC, fluorescent 100 and 200 nm polystyrene beads (yellow–green-fluorescent FluoSpheres, Invitrogen) were used.

#### 2.2. Liposome preparation and characterization

Egg phosphatidylcholine (EPC), egg phosphatidylglycerol (EPG) (Lipoid GmbH, Ludwigshafen, Germany) and cholesterol (Sigma-Aldrich Chemie B.V., Zwijndrecht, The Netherlands) were dissolved in chloroform/methanol (1:1, v/v) in a round-bottom flask in a molar ratio of 2:0:0:1, respectively. A lipid film was prepared by rotary evaporation (Rotavapor R3, Büchi Labortechnik AG, Flawil, Switzerland), followed by drying under a stream of nitrogen. The lipid film was hydrated with 10 mM calcine for 105 nm liposomes or 250 μM calcine for “L146” and “L212” liposomes in HEPES buffered saline (HBS, 10 mM HEPES, 137 mM NaCl, pH 7.4). Liposomes were sized by multiple extrusion under nitrogen pressure using polycarbonate membranes (Nuclepore, Pleasanton, CA, USA) with pore sizes of 200 nm and 100 nm in a Lipex high pressure extruder (Lipex, Northern Lipids, Vancouver, Canada) or a Liposofast Extruder (Avestin, Inc, Ottawa, Canada). Non-entrapped calcine was removed with dialysis against HBS for at least 3 days using Slide-A-Lyzer dialysis cassettes with a cut off of 10 kD (Thermo Scientific, Bremen, Germany). The mean particle size of the liposomes and the polydispersity index (PDI) was determined by means of dynamic light scattering (DLS) using a Malvern ALV CGS-3 with a He–Ne laser source (Malvern Instruments, Malvern, UK). Liposome sizes (L146 and L212) were 146 nm with a PDI of 0.03 and 212 nm with a PDI of 0.07. The zeta-potential of the liposomes (ζ potential) was determined using a Malvern Zetasizer Nano-Z (Malvern Instruments, Malvern, UK). The phosphate concentrations of the liposomes were determined with a phosphate assay described by Rouser et al. [34]. For final use, L146 and L212 liposomes were diluted with HBS till a final total lipid (including cholesterol) concentration of 65 mM.

#### 2.3. Cell culture and EV isolation

The human glioblastoma cell line U87-MG and the lymphoblastoma cell line RN were cultured in medium containing FCS depleted from bovine EVs as described previously [18,19]. After 24 h of incubation the supernatant was isolated and centrifuged at 200 × g for 10 min, two times at 500 × g for 10 min, followed by 10,000 × g for 30 min. 100,000 × g pelleted EVs were resuspended in phosphate buffered saline (PBS) containing 0.2% BSA from an ultracentrifuged stock solution [29]. EVs were fluorescently labeled with 7.5 μM PKH67 (Sigma-Aldrich), mixed with 2.5 M sucrose, overlaid with a linear sucrose gradient (2.0–0.4 M sucrose in PBS) in an SW60 tube (Beckman) and floated into the gradient by centrifugation for 16 h at 192,000 × g [29]. Gradient fractions were collected, diluted in PBS and analyzed. Fraction densities were determined by refractometry.

#### 2.4. NTA

An LM14 Nanosight instrument (Nanosight Ltd, Salisbury, UK) equipped with a CMOS camera (Hamamatsu Photonics, Hamamatsu, Japan) and a 488 nm laser was used. Data acquisition and processing were performed using NTA software 2.3 build 0025. Background extraction was applied, and automatic settings were applied for the minimum expected particle size, minimum track length and blur settings. Since samples were diluted at least 20 times in PBS, viscosity settings for water were applied and automatically corrected for the temperature used. Detection threshold and camera level settings varied as described.
in the Results section. Five movies of 60 s at 25 frames per second were recorded and designated as a single measurement. Only measurements with at least 1000 completed tracks were further analyzed [17]. For polystyrene bead dilutions, single measurements were performed for each dilution, whereas triplicates were recorded for liposome and EV samples.

We excluded data obtained at camera-level 15 (shutter: 1200, gain: 500) as this camera-level resulted in the detection of substantial amounts of background detection, obscuring accurate data interpretation.

2.5. tRPS

For tRPS, the qNano instrument (Izon Science Ltd, Christchurch, New Zealand) was used as described [35]. Data was recorded and analyzed using the Izon Control Suite Software version 2.2.2.111. The default minimum blockade height (0.05 nA) for particle detection was used. For sample calibration and serial dilution experiments, polystyrene beads supplied by the qNano manufacturer were used. Both 115 and 203 nm polystyrene bead dilutions were recorded using NP100 nanopores. Liposome dilutions were recorded using two different NP100 nanopores at 0.8 kPa and 1.2 kPa pressure settings. EV samples were analyzed using both an NP200 (1.2 kPa pressure) and NP150 nanopore (1.4 kPa). The buffers of EVs and calibration beads were kept identical by diluting the calibration beads in the appropriate fraction of a (mock-loaded) sucrose-based density gradient.

2.6. hFC

High-resolution flow cytometric analysis of individual EVs was performed using the BD Influx flow cytometer (Becton Dickinson, San Jose, CA, USA) with an optimized configuration, as described in detail before [29]. Light scattering was measured with a collection angle of 15°–25° (reduced wide-angle FSC) and detection was performed in log mode. Samples were run at low pressure (5 PSI on the sheath fluid and 4.2 PSI on the sample) using a 140 μm nozzle. The calculated flow rate at these settings was 52.2 μl per minute, as determined by weighing the volume aspirated during 30 min. Fluorescent 100 nm and 200 nm polystyrene beads (yellow–green-fluorescent Fluospheres, Invitrogen) were used for calibration of the fluorescence, reduced wide-angle FSC, and SSC settings. EVs in sucrose fractions were diluted in PBS at least 20 times and time-based quantitative measurements were performed as described before [29]. Data was acquired using Spigot software version 6.1 (Becton Dickinson). Data was further analyzed using FCS Express software (De Novo Software, Los Angeles, USA).

2.7. Statistical analysis

Data analysis was performed using Prism version 5.0 (GraphPad Software, La Jolla, CA, USA) or Microsoft Excel 2010 (Microsoft, Seattle, WA, USA). Two-tailed independent t-tests were used to test for significant differences in means. One-way ANOVA followed by Tukey’s post-test was performed to test differences between multiple groups. Pearson’s correlation coefficient was used to determine the correlation between dilution and the measured concentration. Significance was determined and indicated as (*) p-value ≤ 0.05, (**) p-value ≤ 0.01 and (***) p-value ≤ 0.001. Error bars represent the standard deviation (s.d.) unless stated otherwise.

3. Results

3.1. NTA-based particle quantification

Detection of nano-sized particles with NTA is influenced by two parameters: the camera-level (shutter speed and camera gain), which is set prior to data acquisition, and the detection threshold, i.e. the scattering intensity threshold above which particles are traced (set at data processing). Here, we tested how these parameters affected the quantification of nano-sized particles that differ in refractive index, size, and heterogeneity.

First, we determined the accuracy for quantification of homogeneous populations of 115 or 203 nm sized polystyrene beads, which have a high refractive index (r.i.) and consequently cause extensive light scattering. Within a 32-fold dilution range (0.9–290 × 10^6/ml) the measured concentrations approximated the expected concentrations for 115 and 203 nm beads (Fig. 1A) (R²: 0.969 and 0.998 respectively).

For the 115 nm beads we were also able to obtain valid measurements (>1000 completed tracks) outside this range. However, the resulting s-shaped curve (Fig. 1A) indicates an overestimation of particles below 0.9 × 10^6/ml and an underestimation above 290 × 10^6/ml, resulting in decreased correlation accuracy (R²: 0.859 for all measurements). Although the range of refractive indices that EVs can exhibit is largely unknown, polystyrene beads most likely have an r.i. that is substantially higher than the r.i. of the majority of EVs [17]. Next, we tested how camera level and detection threshold settings affect the quantification of calcine-labeled liposomes, which, similar to EVs, are enclosed by a lipid bilayer. Movies were recorded, at camera level 6 (shutter: 150, gain: 250), camera level 9 (shutter: 450, gain: 250), and camera level 12 (shutter: 600, gain: 350), which represent preprogrammed NTA settings. After data acquisition each movie was processed at detection threshold 4, 6, 8 and 10 (standard software setting).

As expected, at increased camera levels the particles appeared brighter and increased detection was observed of weak-scattering particles (Fig. 1B, top-panel). The number of detectable particles was also increased by reducing the detection threshold (Fig. 1B, bottom-panel). A complete overview of screenshots at different camera levels and detection thresholds is provided in Suppl. Fig. S1A). Numerical analysis of these data revealed that the quantification of liposomes is significantly influenced by the NTA settings, with measurement of higher concentrations after increasing the camera level or decreasing the detection threshold (Fig. 1C and Suppl. Fig. S1B). At both camera levels 9 and 12, accurate linearity in measured concentration was observed for multiple dilutions of liposomes (applying camera level 6 resulted in an inadequate number of completed tracks) (Fig. 1D). Thus, relative comparison of liposome concentrations is feasible with different NTA settings, but measurement of the exact concentration strongly depends on the camera-level and detection threshold settings. Of interest, the increased number of liposome detection after increasing the camera level is not accompanied by increased detection of smaller sized liposomes (Suppl. Fig. S1C). It has previously been suggested that smaller particles may be over-scattered by larger particles, which would especially occur after concentrating samples [17,36]. However, this appears not to occur for liposomes (Suppl. Fig. S1D).

Next, we tested to what extent the camera-level and detection threshold influence the quantification of EVs, which are more variable in size and r.i. than liposomes. For these experiments, we used EVs derived from the RN lymphoblastoma and U87-MG glioblastoma cell lines that were purified from contaminating protein aggregates by sucrose density gradient ultracentrifugation. As expected, increasing the camera level resulted in an increased brightness of detected particles (Suppl. Fig. S2A). Similar to the liposome analysis, different EV quantification data were obtained at different camera level settings (Fig. 1E). A maximum fold change of 3 was observed (camera-level 6 versus 12 at detection threshold 10). The influence of detection threshold on particle quantification was less prominent for EVs compared to liposomes (Fig. 1E and Suppl. Fig. S2B). Similar data were obtained for the U87-MG derived EVs (data not shown). Sample dilution did not significantly influence measurement of the raw EV concentration (i.e. the measured sample concentration multiplied by the dilution factor) (Fig. 1F). This was corroborated by the observation that the size-distributions and mode sizes were similar at the different dilutions (Suppl. Fig. S2C and S2D). Thus, although EVs are more heterogeneous
in size than liposomes, the presence of infrequent EV that displayed a higher level of scattering did not substantially affect detection of neighboring EVs.

In conclusion, camera-level and detection threshold variables were found to affect the NTA-based quantification of liposomes and EVs. The influences were more profound for the relatively homogeneous liposomes than for the heterogeneous EVs.

3.2. trPS-based particle quantification

As an alternative to NTA, we tested trPS for liposome and EV quantification, with a specific focus on establishing the most suitable measurement conditions. As trPS-based quantification requires a linear correlation between the particle count rate (particles per minute) and the concentration of particles, we first measured a dilution range of polystyrene calibration beads. We observed linearity over a 64-fold dilution range for the 115 nm beads ($R^2: 0.979$) and over a 32-fold range for the 203 nm beads ($R^2: 0.994$) (Fig. 2A).

Particle detection above threshold levels is dependent on the blockade height (resistive pulse) generated by a particle moving through a nanopore. This blockade height is determined by the particle's volume relative to the volume of the nanopore opening, the applied voltage, and buffer used. These parameters together determine a “trPS setup” and thus determine if particles surpass threshold levels (0.05 nA at default software settings). High-sensitivity trPS setups can be used to detect the smallest particles. To obtain a high-sensitivity setup one should apply a high voltage, low stretch (to establish a minimal opening size of the nanopore) and a small nanopore (NP100/NP150) [32,35]. Nanopore characteristics are known to differ between individual nanopores, as well as over time [37]. To assess the effect of this on liposome quantification, we compared three cases. First we compared two new NP100 nanopores (setups #1 and #2). Subsequently, the nanopore used for setup #1 was tested again after approximately 7 h
of usage (termed setup #3). The most optimal (i.e. high-sensitivity) settings were applied in all three cases. Measurement of 115 nm polystyrene beads for tRPS calibration showed different mode blockade heights detected for setups #1, #2, and #3, with #2 > #1 > #3 (Fig. 2B, left panel). This indicated that the lower detection limit, as determined by the height of the calibration bead blockade relative to the threshold level of 0.05 nA, was different at the different setups. This is illustrated by reconstructing the mean blockade heights of 115 nm calibration particles and liposomes for the three setups (Fig. 2B, right panel). Using setup #3, the peaks of the 115 nm calibration particles are closer to the detection threshold level. The lower detection limit in setup #3 is therefore higher than in setups #1 and #2, implicating that the blockade height induced by smaller liposomes may not surpass the detection threshold. This could result in the detection of only larger-sized liposomes. Secondly, these observed differences indicate that characteristics of nanopores, such as resolution, may change over time.

The variation in detectable size range for setups 1–3 resulted in substantial differences in absolute quantification of the liposomes (difference setup #2 versus #3: 2.43 fold; Fig. 2C). Setup #3 allowed detection of >80 nm liposomes only (Fig. 2C, left panel), and consequently yielded the lowest liposome quantification. Differences in particle concentration (1.43 fold) were also observed for setups #1 and #2, despite the comparable efficiency in detection of small liposomes at these setups (Fig. 2C, left panel).

Besides absolute quantification of liposomes, we also determined how accurate a range of liposome dilutions could be quantified by tRPS (Fig. 2D). An NP100 nanopore was used for this test, and we concomitantly investigated whether the pressure level in flux influenced liposome quantification. For both applied pressure levels we observed accurate detection and linearity over a 4-fold dilution range. Surprisingly, changing the applied pressure led to significantly different liposome sizing estimations (Suppl. Fig. 3A and B).

Fig. 2. tRPS-based particle quantification. (A) Quantification of 115 and 203 nm polystyrene calibration beads. As tRPS quantification is based on the conversion of observed particle per minute counts to that of polystyrene calibration beads of known concentration, the read-out is displayed as "particles per minute". (B) Three tRPS setups displaying the observed blockade heights for the same 115 nm calibration beads (left-panel). The dashed line illustrates the detection threshold (both panels). Reconstruction of the recorded data for beads and liposomes at the three different setups (right-panel), illustrating that the lower detection limit is the highest for setup #3, followed by setups #1 and #2. Bin size 15 pA. (C) Representative liposome size-distributions obtained at the three different tRPS setups (left-panel). For each of the three setups the measured concentration was corrected for the dilution factor to obtain raw concentration estimations (n = 6) (right-panel). Bin size 5 nm. (D) Quantification of serially diluted liposomes at two different pressure levels (n = 3). (E) Representative size-distribution obtained for RN-derived EVs on an NP200 nanopore setup and an N150A nanopore. Bin size 5 nm. (F) Raw particle concentrations were determined for RN and U87-MG derived EVs at both the NP200 and NP150 nanopore setups (n = 3). Data are mean ± s.d.
EV measurements by tRPS indicated that the size distributions of RN (Fig. 2E) and U87-MG derived EVs resembled those obtained using the NTA, with the majority of EVs being 100–200 nm in size. Similar to what we observed for NTA, tRPS showed the presence of a small number of larger (200–600 nm) EVs. Due to the presence of large EVs, we tested two larger nanopores (NP200 and NP150) for EV quantification, to reduce clogging events. Even though frequent nanopore clogging was observed, overall particle detection was stable and reproducible for each triplicate of sample measurements (Suppl. Fig. S4A). Applying the NP150 nanopore, which theoretically allows for detection of 100–120 nm particles, yielded significantly higher EV particle concentrations as compared to the measurement with the NP200 pore (Fig. 2F) (difference RN-derived EVs 1.45 fold, U87-derived EVs 1.50 fold). The ability to measure smaller sized EVs with the NP150 nanopore (Suppl. Fig. S4B), led to significant differences in the calculated mean and mode sizes of the EVs (e.g. mode sizes of 136.3 nm (NP200) and 117.8 nm (NP150) for RN-derived EVs) (Suppl. Fig. S4C).

In conclusion, quantifications of liposomes and EVs can differ between (high sensitivity) nanopore setups and this is most likely related to the lower detection limit. Since the required lower detection limit may be unknown for liposomes and EVs, tRPS measurement may result in underestimation of the concentration.

3.3. hFC-based particle quantification

Reliable quantification of nano-sized particles using fluorescence-based hFC requires that sufficient numbers of fluorophores are associated to the particle to be detected above the fluorescence threshold, and that maximal sensitivity in fluorescence detection is obtained. The optimal configuration and settings for quantitative and qualitative analyses of nano-sized particles using the BD Influx have been determined previously [19,29]. Instrument settings that were found to affect EV measurements included the nozzle size and the applied sample/sheath fluid pressure. Using the optimal settings, 100 nm fluorescent polystyrene beads were efficiently detected above background noise. Furthermore, for a 16-fold dilution range, hFC accurately detected sample dilutions for both 100 (R²: 1.00) and 200 nm (R²: 0.999) fluorescent polystyrene beads (Fig. 3A).

Calcein labeled, 105 nm sized liposomes could also be detected above the fluorescence threshold (Fig. 3B). As expected, light scattering

Fig. 3. hFC-based particle quantification. (A) Quantification of serially diluted 100 and 200 nm fluorescent polystyrene beads. Indicated are the mean number of beads detected in a fixed time window of 30 s. (B) Dotplots indicating that calcein-loaded liposomes can be detected above the fluorescence threshold (solid horizontal line) that excludes detection of non-fluorescent noise events (left-panel) and that the FSC and SSC signals induced by these heterogeneous are highly variable. (C) hFC measurements indicated that PKH-67 labeled RN-derived EVs can be detected above the fluorescence detection threshold (left panel) and that the FSC and SSC signals induced by these heterogeneous are highly variable. (D) EVs were measured over an 8-fold range, and corrected for the dilution used to determine the raw concentration estimation. No statistical different raw concentration estimations between the dilutions were observed. Data are mean ± s.d. (n = 3).
(FSC and SSC) levels generated by low r.i. liposomes were low and could not be discriminated from those generated by noise, indicating the need for fluorescence-based analysis (data not shown). Within the 16-fold dilution range tested here, liposomes could be quantified with accurate linearity ($R^2 > 1.00$) (Fig. 3C).

hFC-mediated detection of RN and U87-MG derived EVs relied on fluorescent labeling of EV and efficient removal of unbound dye by sucrose-gradient ultra-centrifugation [29] (Fig. 3D and Suppl. Fig. 5). Although hFC does not allow for absolute size measurement of EV, variation in size and composition of EV are reflected in the light scattering (FSC and SSC) and fluorescence signals observed. Similar to what was observed in the NTA and tRPS measurements of the RN and U87-MG derived EVs, hFC-based analysis also indicated substantial heterogeneity within these EV populations based on light scattering and PKH67 fluorescence levels (Fig. 3D and Suppl. Fig. 5). Quantification by hFC indicated no significant differences in the estimation of EV concentrations over an 8-fold dilution range for both the RN and U87-derived EVs (Fig. 3E).

In conclusion, once sufficient numbers of fluorophores are associated to liposomes or EVs to allow their detection above the fluorescent threshold, hFC can be used for accurate quantitative analysis of fluorescently labeled liposomes and EVs in a range of sample dilutions.

4. Comparison of liposome and EV quantification using NTA, tRPS, and hFC

For clinical application and research purposes, it is of utmost importance to reliably determine the concentrations of (engineered) EVs or synthetic mimics. Ideally, measurements of identical samples with different technologies should yield comparable quantitative data. We therefore compared quantification data obtained by NTA, tRPS, and fluorescence-based hFC. Based on the previous experiments, a single setup was selected for each instrument. We performed measurements on relatively homogeneous populations of calcine-loaded liposomes with (DLS-based) sizes of 146 and 212 nm (referred to as L146 and L212 respectively), and a more heterogeneous population of purified and PKH67 labeled EVs. NTA-camera levels were selected based on the visually brightest detection of particles, without the occurrence of abundant over-scattering events. The tRPS settings were selected to allow for the highest-sensitivity measurement. More specifically, L146 measurements were performed with NTA-camera-level 12/detection-threshold 4 and nanopore NP100. For L212, camera-level 9/detection-threshold 4 and nanopore NP150 were selected. The RN-EVs were analyzed using NTA-camera-level 12/detection-threshold 10 and an NP150 nanopore. Optimized settings [19] were used for hFC, and hFC settings were identical for measurements of both lipidosome populations and EVs. On the three instruments L146 liposomes were quantified within a 12.5 fold difference (Fig. 4A, left-panel). The highest concentrations were measured with NTA ($1.86 \times 10^{14}$/ml), followed by tRPS ($5.33 \times 10^{13}$/ml), and hFC ($1.5 \times 10^{13}$/ml). Also for the L212 liposomes, NTA measurements yielded the highest concentrations ($7.73 \times 10^{13}$/ml), followed by tRPS ($3.27 \times 10^{13}$/ml) and hFC ($1.12 \times 10^{13}$/ml) (Fig. 4A, right-panel). Overall, the measured L212 concentrations on the three instruments were within a narrower absolute fold-range (6.92). We compared these quantifications with liposome concentration measurements based on dynamic light scattering (DLS)-sizing, lipid composition, and synthetic mimics. Ideally, measurements of identical samples with different technologies should yield comparable quantitative data. We therefore compared quantification data obtained by NTA, tRPS, and fluorescence-based hFC. Based on the previous experiments, a single setup was selected for each instrument. We performed measurements on relatively homogeneous populations of calcine-loaded liposomes with (DLS-based) sizes of 146 and 212 nm (referred to as L146 and L212 respectively), and a more heterogeneous population of purified and PKH67 labeled EVs. NTA-camera levels were selected based on the visually brightest detection of particles, without the occurrence of abundant over-scattering events. The tRPS settings were selected to allow for the highest-sensitivity measurement. More specifically, L146 measurements were performed with NTA-camera-level 12/detection-threshold 4 and nanopore NP100. For L212, camera-level 9/detection-threshold 4 and nanopore NP150 were selected. The RN-EVs were analyzed using NTA-camera-level 12/detection-threshold 10 and an NP150 nanopore. Optimized settings [19] were used for hFC, and hFC settings were identical for measurements of both lipidosome populations and EVs. On the three instruments L146 liposomes were quantified within a 12.5 fold difference (Fig. 4A, left-panel). The highest concentrations were measured with NTA ($1.86 \times 10^{14}$/ml), followed by tRPS ($5.33 \times 10^{13}$/ml), and hFC ($1.5 \times 10^{13}$/ml). Also for the L212 liposomes, NTA measurements yielded the highest concentrations ($7.73 \times 10^{13}$/ml), followed by tRPS ($3.27 \times 10^{13}$/ml) and hFC ($1.12 \times 10^{13}$/ml) (Fig. 4A, right-panel). Overall, the measured L212 concentrations on the three instruments were within a narrower absolute fold-range (6.92). We compared these quantifications with liposome concentration measurements based on dynamic light scattering (DLS)-sizing, lipid composition,
and phosphate quantification, as a standard in the liposome field [38]. Using this method, the calculated liposome concentrations were 2.27 × 10^{13}/ml for L146 and 7.46 × 10^{12}/ml for L212 liposomes (dotted horizontal lines in Fig. 4A). However, DLS is known to be heavily influenced by outliers [30], which may result in overestimated size measurements. When replacing DLS size measurement with averaged liposome sizing data obtained by NTA and tRPS (124 and 156 nm for the L146 and L212 liposomes, respectively; Fig. 4B and C), the calculated liposome concentrations were substantially higher (solid horizontal lines in Fig. 4A) and were most similar to the concentrations obtained by tRPS.

The absolute concentration measurements of EV on the three instruments were within a smaller fold-range difference compared to the measurements of liposomes (4.44 versus 6.92 (L146) or 12.5 (L212); Fig. 4D). Interestingly, quantification of EVs by tRPS and hFC yielded absolute particle concentrations in the same range (1.01 × 10^9/ml and 1.40 × 10^9/ml). However, similar to the liposome measurements, NTA yielded substantially higher values for the EV concentrations (4.50 × 10^9/ml) (Fig. 4D).

In conclusion, the absolute quantifications as observed for both homogeneous calcine-loaded liposomes and a purified population of more heterogeneous, PKH67 labeled RN derived EVs differed significantly between the instruments. For liposomes, the difference in quantifications between the instruments decreased when measuring liposomes that were larger in size. The smallest difference in absolute concentration measurements between the instruments was found when measuring the more heterogeneous population of EVs, for which tRPS and hFC yielded highly similar results.

5. Discussion

Over the last decade, the interest in EVs has greatly intensified due to their proposed role in various biological processes and their potential as biomarkers for disease and as drug delivery systems. Approaches for accurate and standardized quantification of such nano-particles have not yet been established, but are crucial for safe application of EV(-mimics) in clinical settings. Here, we compared quantification of different nano-sized particles, i.e. polystyrene beads, calcine-labeled fluorescent liposomes and purified, PKH67-labeled EVs using three prominent single-EV analysis platforms; NTA, tRPS, and hFC. Moreover, we identified variables that significantly influenced particle quantification using NTA and tRPS.

The particle concentration range at which accurate quantification data could be obtained differed between the instruments. For NTA, the optimal concentration range was 9.0 × 10^7/ml–2.9 × 10^9/ml, which is a slightly larger dilution series than previously reported [16]. For tRPS, the required concentration for particle analysis increased as the particle volume decreased. Consequently, 203 nm beads were analyzed at 9.1 × 10^7/ml–2.9 × 10^9/ml, whereas 115 nm particles were analyzed at 3.6 × 10^6/ml–2.3 × 10^8/ml. hFC allows accurate quantification at lower particle concentrations (a range of 4.6 × 10^6/ml–7.3 × 10^7/ml was analyzed in the current study). Our recent data indicate that concentrations up to 1.0 × 10^10/ml can be reliably measured with hFC (manuscript in submission).

We identified the NTA camera level and detection threshold to be significant factors in the quantification of liposomes (Fig. 1C). In contrast, the absolute differences induced by changing these variable settings were less prominent for quantification of EVs (Fig. 1D). This may be a result of the relatively higher light-scattering properties of EVs (due to the presence of surface/luminal proteins and/or m(i) RNAs), combined with increased heterogeneity in this population, which may make NTA-based detection of EVs less sensitive to differences in settings as compared to the detection of homogeneous liposomes. Besides the empty liposomes used in this study, liposomes engineered to contain proteins and/or nucleic acids show more structural resemblance with EV and quantification of such particles may accordingly be less sensitive to NTA detection thresholding.

Our tRPS analyses showed inter-experimental variation in the sensitivity of liposome and EV detection (Fig. 2C and E), which translated into differences in concentration measurement. This sensitivity of tRPS-based measurements is determined by the size of the smallest detectable particle. For quantification of homogeneous particle populations of a known size, such as calibration beads, the most suitable nanopore setup for detection of all particles can easily be selected. However, for samples with an unknown size-distribution (e.g. EVs) this is more difficult and the obtained size detection range may be insufficient for detection of all particles. Besides this, we also noted slight differences in tRPS-based concentration measurements (up to ~1.4-fold) between set-ups in which the liposome size-distribution profiles and detection limits were similar (Fig. 2C). We hypothesize that subtle differences in nanopore size due to batch variations and nanopore longevity could have caused these variations in particle quantification. The observed differences stress the importance of comparing samples using exactly the same tRPS setup. Electro-kinetic forces were recently suggested [39] to influence the movement of particles through smaller tRPS nanopore. In case particles possess a different surface-charge compared to the polystyrene calibration particles, one of the two particle types may be more likely to pass through the nanopore. Since this may cause inaccuracy in the calculated particle/minute to concentration calculation, the manufacturer suggested to perform quantifications at two or more pressure levels, after which the tRPS software can determine a surface-charge corrected concentration. Since we observed no difference in the measured liposome concentration at two different pressure levels (Fig. 2D), we conclude that electrokinetic forces at these settings do not significantly influence the quantification of particles. The surface charge of the studied liposomes was −43.0 ± 0.87 mV, which is similar to the reported surface charge characteristics of EVs [40–42]. Single-pressure tRPS quantifications can therefore suffice for accurate EV quantifications. The difference in blockade height when measuring the 115 nm calibration particles at the two pressure settings (Suppl. Fig. S3A, left-panel) was unexpected, because the applied pressure does not change the particle volume and nanopore diameter. Implications of this phenomenon for particle characterization need to be further studied. When comparing the liposome size-distributions obtained by tRPS and NTA (Fig. 2C and Suppl. Fig. S2C), we conclude that both NTA and tRPS allowed detection of liposomes as small as 55–60 nm in size, which for NTA is the theoretical lower limit of liposome detection, limited by the r.i. of the particle [16,20].

In contrast to NTA and tRPS, for hFC the threshold for particle detection is based on fluorescence intensity. Although the sensitivity for detection is largely improved by the use of high power lasers and by increasing the dwell time of the vesicles in the laser beam, particles with low fluorescence intensity (e.g. due to low PKH67 labeling efficiency or because of small size) may not be detected using this technique. Furthermore, the removal of unbound fluorescent dye by density gradient ultracentrifugation can be seen as a time-consuming procedure. However, the same procedure also allows separation of EVs from protein aggregates that are abundantly present in culture media and body fluids. This is essential, since such aggregates can mistakenly be recorded as vesicles by the technologies discussed here.

Comparability analysis of the three techniques indicated that substantially larger differences in quantification were obtained for liposomes, compared to EVs (Fig. 4A and D). In fact, no significant difference in raw concentration estimation was observed for quantification of EVs by tRPS versus hFC. One potential explanation is that the EVs exhibit higher fluorescence levels compared to the liposomes, either because EVs are larger in size and incorporate more dye or because of differences in labeling efficiency. Differences between the other instruments are difficult to account for. For both liposome batches and EVs, higher raw concentration estimations were obtained by NTA compared to tRPS. We tested whether background particle detection (from the buffer in
which particles were diluted) could explain the observed differences. However, measurement of PBS background particles at camera-level 9 and 12 revealed maximum concentrations of only 2.60 × 10⁷/ml and 3.3 × 10⁷/ml, respectively (data not shown), accounting for 3.29% (L146) and 2.33% (L212) of the measured concentrations of the liposomes.

So far, only one other study has directly compared NTA, tRPS and flow cytometry (using a different high-end flow cytometer) by analyzing the size distributions of polystyrene beads and urine-derived EVs [32]. Interestingly, comparable EV quantifications by NTA and tRPS were reported, whereas the flow cytometry-based EV quantification was 15 times lower. However, a direct comparison of these data to our current study is difficult, because a crude preparation (1550 × g centrifugation followed by 0.2 μm filtration) of EVs from a different biological source (urine) was analyzed and because the flow cytometric measurements in that study were light scatter-based. However, it is interesting to note that, in contrast to our findings, EV quantifications by NTA and tRPS were found to be comparable. This could imply that the type of EV and the degree of EV purification may also influence quantification by the different instruments.

Several strategies have previously been suggested to calibrate particle quantification in EV samples. For tRPS we spiked biological fluids with polystyrene beads of known size and concentration to improve EV quantification accuracy [18]. For NTA, on the contrary, this approach seems less suitable since the methodology does not allow for accurate discrimination of particles of interest from beads with a similar size [17,20,30]. Secondly, spiking a sample with large (>500 nm) silica beads could lead to over-scattering of the EVs and skew characterization [20]. An alternative that has been proposed for NTA calibration is a correction factor, based on the measured concentration of silica beads compared to the expected concentration of these beads [17]. Although promising and potentially valuable for measuring relatively homogenous populations of EVs, such a calibration method is unsuitable for analysis of the heterogeneous EV preparations studied here. As NTA is less accurate in the detection of size-based subpopulations [17,20,30], one would have to apply a multitude of silica calibration beads, each covering a subpopulation of EVs and subsequently aggregate analysis of these subpopulations. More research into the accuracy of such a calibration system will be essential before it can be broadly applied.

In conclusion, we identified NTA and tRPS instrument settings that affect particle quantification and showed that the impact of these parameters on quantification varies with the types of nano-sized particles analyzed (i.e. polystyrene beads, liposomes and EVs). Our data clearly indicate that absolute quantification of EVs and liposomes substantially differs using the three different technologies and that a golden standard for quantification of such particles is not available yet. Moreover, our data strongly underlie the importance of technical knowledge of the instruments for correct data interpretation, and plead for awareness of the effects of instrument settings in case vesicle populations with unknown concentration and size heterogeneity are measured. Increased understanding of the possibilities and pitfalls of these technologies will benefit standardized and large-scale clinical application of (engineered) EVs and EV mimics in the future.

Acknowledgments

This work has been financially supported, in part, by the Hersenstichting Nederland (2013-01) (foundation concerned with diseases of the brain), the Schumacher Kramer Stichting (11-04-2014), and the T & P Bohnenn Foundation (SLNM, JDV, BG, MLB). ENMN-TH receives funding from the European Research Council under the European Union’s Seventh Framework Programme (FP/2007–2013)/ERC grant agreement number 337581.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jconrel.2014.12.041.

References


