Intrinsic variability of current MESF calibrators impacts the assignment of absolute fluorescence values to nanoparticles and extracellular vesicles by flow cytometry

Estefanía Lozano-Andrés[1, #], Tina Van Den Broeck[2, #], Lili Wang[3], Majid Mehrpouyan[4], Ye Tian[5], Xiaomei Yan[5], Marca H.M. Wauben*[1, #], Ger J.A. Arkesteijn*[1]

1 Department of Biomolecular Health Sciences, Faculty of Veterinary Medicine, Utrecht University, Utrecht, The Netherlands

2 BD Biosciences, Erembodegem, Belgium

3 Biosystems and Biomaterials Division, National Institutes of Standards and Technology (NIST), Gaithersburg, MD 20899

4 BD Biosciences, San Jose, CA 95131

5 Department of Chemical Biology, MOE Key Laboratory of Spectrochemical Analysis & Instrumentation, Key Laboratory for Chemical Biology of Fujian Province, College of Chemistry and Chemical Engineering, Xiamen University, Xiamen 361005, People's Republic of China

TRAIN-EV Marie Skłodowska-Curie Action-Innovative Training Network, train-ev.eu

*Both authors contributed equally

Corresponding author:

Estefanía Lozano-Andrés

Department of Biomolecular Health Sciences

Faculty of Veterinary Sciences

Utrecht University

Yalelaan 2, 3584 CM

Utrecht, The Netherlands

T: +31 30 2535451/2535833

E: e.lozanoandres@uu.nl

Abstract

Flow cytometry is a promising technique to characterize nanoparticles (NPs) and extracellular vesicles (EVs). However, the majority of reported experiments use arbitrary units to indicate fluorescence intensity. This hampers comparison of results from different laboratories and different platforms. We investigated the advised use of calibrated molecules of equivalent soluble fluorophores (MESF)-beads for assignment of absolute fluorescence to NPs and EVs. Firstly, we evaluated the use of two different FITC-MESF bead sets as calibrators on three different flow cytometry platforms (BD Influx™, CytoFLEX LX™ and SORP BD FACSCelesta™). Secondly, NPs and biological 4T1 mammary carcinoma-EVs were analyzed using the BD Influx to calibrate their fluorescence signals by using four different sets of FITC and PE-MESF beads. Fluorescence calibration using bright calibrators designed for cellular flow cytometry makes inter-platform comparison possible. However, the intrinsic uncertainty in the assignment of the currently available MESF beads, which have a far brighter fluorescence range than NPs and EVs, hampers a reliable assignment of absolute MESF numbers based on extrapolation into the dim fluorescence range.

The findings that we present in this study demonstrate that for robust comparison of calibrated values to submicron sized particles the same calibration materials should be used. Furthermore, our results illustrate the need for dim calibration materials specifically designed for calibration of NPs and EVs to enable a reliable assignment of absolute fluorescence values in the lower fluorescent ranges and to perform inter-platform comparison.

Key words: fluorescence, calibration, standardization, MESF, extracellular vesicles, nanoparticles, flow cytometry

Introduction

The assignment of a specific value is inextricably bound to a variation around this value. The calibration methods used in flow cytometry are no exception to this rule. A well-known calibration method in flow cytometry is the use of beads to which a specific fluorescence value is assigned and expressed in molecules of equivalent soluble fluorophores (MESF) or equivalent number of reference fluorophore (ERF). These are units established and accepted by the National Institute of Standards and Technology (NIST). Calibration beads are commercially available and suited for a large number of fluorescence channels. Flow cytometry is most commonly used for the purpose of cell analysis by determining the level of fluorescence intensity for the marker of interest on labeled cells. The fluorescence intensity on the calibrator beads matches the intensity that can be expected on labeled cells. Therefore the calibrated cellular fluorescence variation that has to be taken into account closely compares to the intrinsic variation on the calibrator, allowing for data interpolation. Besides cellular measurements, flow cytometry has attracted the attention of many researchers working in the field of nanoparticles (NPs) and cell-derived extracellular vesicles (EVs) as a valuable tool for high-throughput single particle-based analysis [1, 2]. It allows for quantitative

71 and qualitative multiparametric characterization of subsets of EVs based on the detection of light
72 scatter signals and fluorescence. Although flow cytometry is one of the methods of choice for EV
73 analysis, it is not free of challenges and pitfalls, mainly because EV signals are lying close to the
74 instrument's detection limit. Important requirements need to be met and data need to be reported
75 in a proper way to improve rigor and reproducibility of flow cytometric EV analysis [3]. Despite
76 the advancements made in the past years, the vast majority of EV flow cytometric data is still
77 published in arbitrary units and results are expressed as percentages of subpopulations or median
78 fluorescence intensity (MFI) relative to the experimental controls. This approach is cumbersome
79 for the analysis of dim and small particles, whereby positively stained events cannot be fully
80 discriminated from negative counterparts and background signals. Furthermore, although the
81 number of flow cytometers that are suitable for NP and EV analysis is increasing, these platforms
82 are very diverse in terms of fluidics, optics and electronics. Calibrated measurements will not only
83 improve inter-platform comparison but also contributes to determine the sensitivity of a flow
84 cytometer, which is especially relevant when a fluorescent threshold is applied. Altogether, the use
85 of standardized units of fluorescence (MESF or ERF) will strongly contribute to rigor and
86 reproducibility of EV flow cytometric analysis.. The three-society ISEV-ISAC-ISTH EV flow
87 cytometry working group has recommended in the published MIFlowCyt-EV framework for
88 standardized reporting of EV flow cytometry experiments, the use of MESF beads for fluorescence
89 calibration [3]. Fluorescence calibration is successfully applied for cell flow cytometry in which
90 arbitrary units are translated to standardized units of fluorescence [4-7]. Key in proper calibration
91 is the precision and accuracy of the available calibrators. Importantly, a NIST/ISAC
92 standardization study reported in 2012 observed differences in the assigned standardized number
93 to calibrators provided by different manufacturers [8]. Therefore we decided in our study to focus
94 our evaluation on different calibrator sets provided by the same manufacturer to exclude any

95 additional variations due to inter-manufacturer discrepancy. However, inherent to the method used,
96 an intrinsic uncertainty of the assignment of absolute numbers of MESF or ERF to calibrator beads
97 will always exist. This uncertainty will translate in a variation that remains within acceptable range
98 as long as the sample values are within the range of this calibrator. However, the currently available
99 calibrator materials are designed for the fluorescence calibration of labeled cells, being much
100 brighter in fluorescence when compared to the fluorescence of stained EVs and other NPs. We here
101 analyzed the usability of the available calibrator materials for assigning absolute fluorescence
102 values for EVs and NPs.
103 We investigated how the variation within the number of MESF or ERF of the calibrator impacts
104 slope and position of the regression line and consequently influenced the assignment of absolute
105 units of fluorescence to NPs and EVs that are far out of the range of these calibrator intensity levels.
106 To this end, we evaluated the performance of different custom-made calibrator beads sets provided
107 by the same manufacturer and their inter-platform variation on three different flow cytometers
108 capable of single NP and EV analysis. Finally, we studied how the use of different bead sets would
109 affect the calibration of fluorescence signals from NPs and fluorescently stained and antibody
110 labeled EVs on the same platform.
111

112 **Material and Methods**

113 **Calibration beads**

114 For calibration of the fluorescence axis in the fluorescein isothiocyanate (FITC) channel we used
115 two different sets of FITC MESF beads (custom-made, 6 µm lot MM2307 #131-10; #131-8; #130-
116 6; #130-5; #130-3 and 2 µm lot MM2307#156; #159.1; #159.2; #122.3, BD Biosciences, San Jose,
117 CA). For calibration of the fluorescence axis in the PE channel, we used two different sets of PE

118 MESF beads (6 µm, commercial QuantiBrite, Catalog No. 340495 lot 62981 and 2 µm, custom
119 made, lot MM2327#153.1; #153.2; #153.4; #153.5; #153.6, BD Biosciences, San Jose, CA).
120 The 6 µm FITC MESF beads were prepared by reacting various concentration of FITC with PMMA
121 Beads (Bangs Labs) in borate buffer at pH 9.2. The 2 µm FITC beads were prepared by reacting
122 various concentrations of FITC-BSA (with a FITC/BSA molar ratio of 2) with 2 µm carboxylic
123 beads (Bangs Labs) using EDC/NHS chemistry. The 2 µm PE beads were made as described above
124 except various concentrations of PE were used with 2 µm carboxylic beads in EDC/NHS chemistry.
125 These beads were analyzed on a BD LSRFortessa™ (BD Biosciences) and their MESF values were
126 assigned by cross-calibration using commercially available MESF beads (Flow Cytometry
127 Standards Corp.). The FITC ERF values were assigned to both 2 µm and 6 µm beads using a
128 specific lot of FITC-FC Bead (BD Biosciences) as a calibrator with known ERF value, which has
129 been assigned by NIST. This provided us with two distinct calibrator bead sets that were produced
130 through the same manufacturing process and assigned using the same instruments and the same
131 internal NIST traceable calibrator to exclude internal processing variations.
132 In addition, we measured commercially available Quantum™ FITC-5 MESF (7 µm, Catalog No.
133 555, lot 14609, Bangs Laboratories) and AccuCheck ERF Reference Particles Kit (3 µm, Catalog
134 No. A55950, lot #081220207, #081220203, #081220208, Thermo Fisher) which were prepared
135 according to the manufacturer's instructions.
136 All calibration bead sets were measured with gain or voltage settings as would be used for the
137 analysis of small particles (i.e. EVs). In addition, not all beads could be measured on every
138 instrument. For fair cross-platform comparison of the slopes of the regression lines, only the same
139 bead populations that were measured on all instruments were included for linear regression
140 analysis.
141

142 **Flow cytometer platforms**

143 In this study three flow cytometers were used. A jet in air-based BD Influx (BD Biosciences, San
144 Jose, CA), a BC CytoFLEX LX (Beckman Coulter, Brea, CA) with a cuvette-based system and a
145 cuvette-based SORP BD FACSCelesta™ (BD Biosciences, San Jose, CA) equipped with a
146 prototype small particle side scatter module.

147 The BD Influx flow cytometer was modified and optimized for detection of submicron-sized
148 particles [9]. In brief, FITC was excited with a 488 nm laser (Sapphire, Coherent 200 mW) and
149 fluorescence was collected through a 530/40 bandpass filter. PE was excited with a 562 nm laser
150 (Jive, Cobolt 150 mW) and fluorescence was collected through a 585/42 bandpass filter. Optical
151 configuration of the forward scatter detector was adapted by mounting a smaller pinhole and an
152 enlarged obscuration bar in order to reduce optical background. This reduced wide-angle FSC
153 (rwFSC) allowed detection of sub-micron particles above the background based on forward scatter
154 [9, 10]. Upon acquisition, all scatter and fluorescence parameters were set to a logarithmic scale.
155 To minimize day to day variations, the BD Influx was standardized at the beginning of each
156 experiment by running 100 and 200 nm yellow-green (505/515) FluoSphere beads (Invitrogen,
157 F8803 and F8848). The instrument was aligned until predefined MFI and scatter intensities where
158 reached with the smallest possible coefficient of variation (CV) for rwFSC, SSC and fluorescence.
159 After optimal alignment, PMT settings required no or minimal day to day adjustment and ensured
160 that each measurement was comparable. MESF beads and NPs were measured with a FSC
161 threshold set at 1.0 while for biological EVs a fluorescence threshold was set at 0.67 by allowing
162 an event rate of 10-20 events/second while acquiring a clean PBS control sample.

163 When performing quantitative and qualitative analysis of synthetic NPs and biological EVs,
164 preparations were diluted in PBS as indicated. Upon loading on the Influx, the sample was boosted
165 into the flow cytometer until events appeared, after which the system was allowed to stabilize for

166 30 seconds. Measurements were performed either by a fixed 30 second time or by setting a gate
167 around the spike-in beads and allowing to record a defined number of events in the gate (80 000
168 events) using BD FACS Sortware 1.01.654 (BD Biosciences).
169 The CytoFLEX LX was used without any tailor-made modifications in the configuration. Before
170 measurements, the manufacturer recommended startup and QC procedure were run first. All scatter
171 and fluorescence parameters were set to a logarithmic scale. FITC was measured with a 50 mW
172 488 nm laser and fluorescence was measured through a 525/40 band pass filter at gain 1.0. FITC
173 MESF beads were recorded with an FSC threshold at 1000. Measurements were performed using
174 CytExpert 2.1 (Beckman Coulter).
175 The SORP BD FACSCelesta™ was equipped with a prototype small particle SSC module for
176 improved scatter detection. Before measurement, the recommended CS&T performance check was
177 run to monitor performance on a daily basis and to optimize laser delay. All scatter and fluorescence
178 parameters were set to a logarithmic scale. 100 nm yellow-green (505/515) FluoSphere beads
179 (Invitrogen, F8803) were acquired and used to set optimal fluorescence (FITC detector) PMT-V
180 values. FITC was measured with a 100 mW 488 nm laser through a 530/30 band pass filter. FITC-
181 MESF beads were recorded with an SSC threshold at 200. Measurements were performed using
182 BD FACSDiva™ Software v8.0.3 (BD Biosciences).
183 Further descriptions of each instrument and methods are provided in Data S1 (MIFlowCyt
184 checklist) and Data S2 (MiFlowCyt-EV framework).
185

186 **Preparation of FITC-doped silica nanoparticles**

187 Synthetic silica nanoparticles (SiNPs) of 550 nm diameter with six different FITC fluorescence
188 intensities were produced by using a modified method of literature reports [11-13]. Briefly, the
189 amine reactive FITC molecules were covalently linked to the silane coupling agent, (3-

190 aminopropyl)-triethoxylsilane (APTES) in anhydrous ethanol. Monodisperse silica seeds of ~90
191 nm prepared by using amino acid as the base catalyst [11, 12] were suspended in a solvent mixture
192 containing ethanol, water and ammonia. Then tetraethyl orthosilicate (TEOS) and different
193 volumes of APTES−FITC solutions were added for growing FITC-doped SiNPs by a modified
194 Stöber method [13, 14]. Upon washing three times with anhydrous ethanol, the FITC-doped SiNPs
195 were reacted with TEOS in the solvent mixture to allow growth of a silica layer. The synthesized
196 SiNPs were washed three times with anhydrous ethanol and stocked in anhydrous ethanol. The
197 diameters of SiNPs were measured by transmission electron microscopy.
198
199 **Isolation and fluorescent staining of extracellular vesicles for flow cytometric analysis**
200 EV-containing samples were obtained from 4T1 mouse mammary carcinoma cell culture
201 supernatants (ATCC, Manassas, VA) as previously described [10, 15, 16]. EVs were stained with
202 5-(and-6)-Carboxyfluorescein diacetate succinimidyl ester (CFDA-SE, hereinafter referred as
203 CFSE) (Thermo Fisher, Catalog No. C1157) and separated as described previously [9]. Briefly, 2
204 µl of the isolated 4T1 EVs (corresponding to a concentration of 1.44 E12 particles/mL as
205 determined by nanoparticle tracking analysis) were mixed with 18 µl PBS/0.1% aggregate-depleted
206 (ad)BSA. For antibody labeling, samples were first resuspended in 15.5 µL PBS/0.1% adBSA and
207 incubated with 0.5 µg of rat anti-mouse CD9-PE (Clone: KMC8, IgG2a, κ, lot 7268877, BD
208 Biosciences) or matched isotype antibodies (Rat IgG2a, κ, PE-conjugated, lot 8096525, BD
209 Biosciences) for 1h at RT while protected from light. EVs were then stained with 40 µM CFSE in
210 a final volume of 40 µl. The sealed tube was incubated for 2h at 37°C while protected from light.
211 Next, staining was stopped by adding 260 µl PBS/0.1% adBSA. After fluorescent staining, EVs
212 were separated from protein aggregates and free reagents by bottom-up density gradient
213 centrifugation in sucrose for 17.30 h at 192,000 g and 4°C using a SW40 rotor (k-factor 144.5;

214 Beckman Coulter, Fullerton, California, USA). Twelve fractions of 1 mL were then collected from
215 the top of the gradient and respective densities were determined by refractometry using an Atago
216 Illuminator (Japan). For analysis by flow cytometry, EV samples corresponding to a 1.14 g/mL
217 density were diluted 1:20 in PBS prior measurement. MIFlowCyt-EV guidelines [3] were followed
218 whenever applicable (Data S2).
219
220 **Concentration determination by using spike-in beads**
221 EV concentration was normalized using a spiked-in external standard containing 200 nm orange
222 (540/560) fluorescent beads (Invitrogen, F8809). The concentration of the beads was determined
223 by Flow NanoAnalyzer N30 (NanoFCM, Xiamen, China) and stocked at 5.7E10 particles/mL.
224 Beads were diluted 1:10^4 in PBS and added to the EV samples, mixed and measured on the flow
225 cytometer. Bead count was used to calculate the EV concentration for BD Influx measurements.
226
227 **Data analysis**
228 For fluorescence calibration each bead peak population was gated using FlowJo Version 10.5.0
229 and MFI were obtained for further least square linear regression analysis. Data was handled in
230 Microsoft Excel and figures were prepared using GraphPad Prism version 8.0 (GraphPad Software
231 Inc). The software FCMPASS Version v2.17 was used to generate files with calibrated axis units
232 in the histograms and dot plots shown [17] (Software is available on
233 http://go.cancer.gov/a/y4ZeFtA).
234
235 **Data availability**

236 All EV data of our experiments have been submitted to the EV-TRACK knowledgebase (EV-
237 TRACK ID: EV210047) [18]. All flow cytometric data files have been deposited at the Flow
238 Repository (FR-FCM-Z3FJ).
239
240 **Results**
241 **The inclusion of different numbers of bead populations impacts linear regression analysis of**
242 **custom-made and commercial MESF-bead sets**
243 We noticed that available calibrator sets differ in the number of fluorescent bead populations,
244 typically ranging from 3 to 5 populations. Therefore, we first assessed whether the inclusion of
245 more or less populations impacts the consistency of the slope generated by linear regression
246 analysis. To this end, we used custom-made standard FITC MESF beads (6 µm) and commercially
247 available beads from Bangs Labs, both containing five fluorescent bead populations with relatively
248 similar FITC MESF values (ranging respectively from 7094 to 715225 and 8452 to 1203595).
249 On the BD Influx, we could detect all 5 populations from brightest to dimmest for both FITC MESF
250 bead sets when using the EV analysis settings. Next, provided MESF values by the manufacturers
251 and measured MFI were transformed into log-log and three graphs with three equations were
252 obtained for each FITC MESF bead set (Figure 1). The graphs represent linear regressions by
253 including either: (i) all five bead populations, (ii) 4 bead populations (from brightest to second
254 dimmest) or (iii) 3 bead populations (from brightest to third dimmest). We found that the number
255 of populations that are used for the subsequent linear regression analysis has an impact on the slope
256 that belongs to that particular equation (Figure 1). This was observed for both custom-made and
257 commercially available beads and showed that when fluorescence calibration is performed ideally
258 the same bead populations should be used for linear regression analysis.
259

260 **Assessment of different MESF bead sets for the calibration of fluorescent signals across**
261 **platforms**

262 Two FITC MESF bead sets of 6 µm and 2 µm, containing either five or four bead fluorescent
263 populations were measured using three different flow cytometer platforms, namely a BD Influx, a
264 BC CytoFLEX and a SORP BD FACSCelesta™. For fair comparison between the three
265 instruments, we included 4 bead fluorescent populations from both FITC MESF bead sets, as these
266 were consistently measured on all three platforms. Respective singlet gated populations are
267 displayed as overlays in histograms showing the FITC fluorescence (Figure 2a). The 4 populations
268 in the histograms of both bead sets are indicated from dim to bright fluorescence as p1, p2, p3 and
269 p4 (Figure 2b). The bigger sized 6 µm FITC MESF beads contained overall brighter fluorescent
270 intensities, whose assigned values for the four populations used in the cross-platform comparison
271 range from 25,910 to 715,225 FITC MESF. The smaller sized 2 µm FITC MESF beads covered a
272 dimmer part of the fluorescence intensity range with assigned values ranging from 3,634 to 103,706
273 FITC MESF. As expected, when MFI values are expressed in arbitrary units, data cannot be directly
274 compared between instruments (Figure 2b). However, when the FITC MFI of every bead
275 population was next correlated to the FITC MESF values provided for each population using least
276 square linear regression, we obtained comparable FITC MESF units (Figure 2c). Importantly, the
277 two bead sets revealed a different slope, consistent within each platform, indicating that this
278 variation was introduced by an inherent attribute of the beads. To evaluate the precision and
279 accuracy of FITC MESF assignments we used specific bead populations from one set, referred to
280 as 'unknown' in Figure 2d-h, to calculate the FITC MESF value using the regression line from the
281 other bead set. The 'unknown' samples used were: (i) p1 and p4 of the 2 µm bead set for which the
282 FITC MESF values were calculated using the regression line of the 6 µm bead set and (ii) p1 and
283 p2 of the 6 µm bead set for which the FITC MESF values were calculated using the regression line

284 of the 2 µm bead set (Figure 2c). Using this approach, the extrapolated FITC MESF values of the
285 'unknown' p4 from the 2 µm beads and the 'unknown' p2 from the 6 µm beads could successfully
286 be compared between platforms, proving the precision of the fluorescence calibration and
287 facilitating the interpretation of data acquired on the three different platforms (Figure 2d-e and
288 Figure 2h).
289 To quantify accuracy, we decided to set a range of 20% difference as acceptable variation of the
290 assigned values (10% above or 10% below actual values, as indicated by the green band in Figure
291 2d-e). We found less than 20% difference from the expected value for all calculated FITC MESF
292 values of the 'unknown' signal from the p4 of the 2 µm beads on the 6 µm bead-set regression line
293 and the p2 of the 6 µm beads on the 2 µm bead-set regression line. This observation was
294 independent of the platform and bead set used. Thus, further demonstrating not only precision but
295 also the accuracy of fluorescence calibration (Figure 2d-e and Figure 2h).
296 However, when the 'unknown' p1 bead signals of the 2 and 6 µm beads were calculated using
297 respectively the 6 or 2 µm bead-set regression lines (Figure 2f-g) , we found that although
298 calculated FITC MESF values were again comparable between platforms and hence calibration
299 proved to be precise, the accuracy was uncertain. Comparison of the calculated MESF values of
300 the 'unknown' beads to the expected MESF values in these cases revealed more than a 20%
301 difference, leading to either an underestimation or an overestimation of the absolute number of
302 FITC MESF (Figure 2f-h).
303 This can be explained by the fact that slight differences in the slope of the calibration lines of the
304 different bead sets become more prominent when extrapolation needs to be extended into the dim
305 area beyond the fluorescence intensities of the calibration beads themselves (Figure 2f). It is
306 unlikely that differences in slopes are caused by instrument non-linearity since it is a recurring
307 pattern on all three platforms (Figure 2c) and linear plotted values do not show non-linearity

308 (Figure S1a-c). Instrument linearity was further confirmed on the BD Influx following the approach
309 described by Bagwell et al [19] (Figure S2). Additionally, variability between separate
310 measurements could be excluded (Figure S3). Furthermore, by using additional commercially
311 available calibration beads, slope variability was confirmed to be inherent to the use of calibrator
312 beads in general (Figure 1) further indicating that this is not a phenomenon restricted to the custom-
313 made beads only. Analysis of a second commercial set of ERF Reference Particles showed that
314 only two bead populations could be detected using our EV settings. Although for this specific bead
315 set additional information about the uncertainty of the fluorescent value of the beads was provided,
316 the robustness of the linear regression analysis based on only two data points is questionable
317 (Figure S4). Finally, a similar phenomenon as described for the FITC MESF calibrator sets was
318 observed when 6 µm and 2 µm calibrator bead sets for PE MESF were used, albeit in this case the
319 dimmer 2 µm beads displayed a steeper slope than the brighter 6 µm beads (Figure S1d-f).
320 Altogether, our observations are pointing towards an uncertainty in the MESF assignment of both
321 the FITC and the PE MESF beads rather than towards instrument or other external factors.
322

323 **MESF assignments of FITC fluorescence intensities from synthetic silica nanoparticles**
324 **depend on the MESF-bead calibrator set used**

325 So far, we have demonstrated that MESF assignment of unknow particles potentially has a higher
326 chance of being inaccurate when the signal of the unknown particle is significantly higher/lower
327 than the fluorescence intensities of the used MESF calibrators themselves. We next investigated
328 how calibration with the FITC MESF bead sets impacts absolute fluorescent assignment when
329 nanoparticles with even dimmer fluorescence were measured. For this purpose, we focused on one
330 platform (BD Influx) to measure and analyze 550 nm silica NPs, containing 6 populations with
331 different FITC fluorescence intensities. Singlets were gated (Figure S5) and an overlay histogram

332 showing 6 different FITC fluorescence intensities was generated (Figure 3a, left). We next
333 performed fluorescence calibration and generated two additional histograms showing calibrated
334 FITC MESF units based on either the 6 µm (Figure 3a, middle) or the 2 µm FITC MESF calibrator
335 bead set (Figure 3a, right). The obtained MFI and CV values for each population, as well as the
336 calculated FITC MESF values based on the two different calibrator sets are shown in Figure 2b.
337 Our data show that when using the 6 µm calibration beads, the calculated FITC MESF values were
338 always lower as compared to the use of the 2 µm bead set regression line (Figure 3b). In addition,
339 the multi-intensity peak analysis of the silica NPs in one measurement revealed that the difference
340 in the absolute FITC MESF values of these NPs obtained by the two calibrator bead sets increased
341 in the dimmer range of the fluorescence, with a 27.3% variation for the brightest fluorescent peak
342 (p6) and 76.5% variation for the dimmest population of these NPs (Figure 3b). To further visualize
343 this, we used the calculated FITC MESF values (Figure 3b) to generate an overlay bar graph
344 showing the differences in calculated FITC MESF from the six FITC fluorescent populations
345 within the 550 nm NPs (Figure 3c). Our results demonstrate that the variation between using the 6
346 µm (dark green) or 2 µm (light green) FITC MESF bead set progressively increased from brightest
347 (p6) to dimmest (p1) NP populations (Figure 3c), as expected. This phenomenon can be ascribed
348 to the differences in the slopes of the regression lines of the two bead sets and the extrapolation
349 into the dim area. To confirm that this phenomenon was not limited to the use of FITC MESF beads
350 we additionally analysed 6 µm and 2 µm PE MESF calibrator bead sets on the BD Influx (Figure
351 S1d-f). By measuring a sample containing 200 nm broad spectrum fluorescent polystyrene NPs we
352 could determine its fluorescent intensity in terms of ERF (Figure S6a). The variation in the PE ERF
353 value obtained by the two calibrator sets was in this case 41.3% (Figure S6b), thereby
354 demonstrating that the phenomenon observed with FITC MESF beads also occurs with PE MESF
355 calibrator beads.

356

357 **Different MESF-bead sets assign different absolute numbers of ERF and MESF to**
358 **fluorescently CFSE stained and CD9 labeled extracellular vesicles**

359 Next, we investigated the impact of this phenomenon on the assignment of absolute FITC ERF
360 units and PE MESF units to an EV sample measured on the same platform by using the four
361 different calibration bead sets, i.e. 6 and 2 µm FITC MESF and the 6 and 2 µm PE MESF beads.
362 Since EVs scatter less light than beads and NPs, we used a fluorescence threshold to measure the
363 EVs [9]. To this end, we used CFSE as a luminal dye to fluorescently stain 4T1-derived EVs. We
364 analyzed the expression of CD9, a tetraspanin enriched on the surface of 4T1-derived EVs, by
365 using a tetraspanin CD9-PE antibody. As controls, we measured unstained EVs and CFSE stained
366 EVs with a matching isotype-PE control (Figure 4a). Since 200 nm spike-in beads were added to
367 the samples for EV concentration measurement, we defined an EV-gate and a spike-in bead gate
368 in these samples and relevant additional controls (Figure S7a). Our data revealed a 76.6% variation
369 in the MFI expressed in absolute CFSE ERF units on CFSE stained EV and a 156.9% variation in
370 the MFI expressed in absolute PE MESF units on CD9-PE labeled EVs when the two different
371 calibrator bead sets were used on the same platform (Figure 4a-b). The fluorescent threshold value
372 of 0.67 used on the BD Influx corresponds to an equivalent of 150 FITC MESF based on the 6 µm
373 beads or 300 FITC MESF based on the 2 µm beads (Figure 4a) and indicates the level of detection
374 using this fluorescence threshold. Therefore, especially when a fluorescence threshold is used to
375 measure EVs, accurate calibrators are required to be able to report the fluorescence threshold in
376 calibrated units. Lastly, in the overlay histograms it is shown how the fluorescence intensities of
377 the calibrators relate to the dimmer signals generated by CFSE stained and CD9-PE labeled EVs
378 (Figure 4c-d).

379

380 **Discussion**

381 The calibration of fluorescence measurements obtained by flow cytometry in terms of MESF or
382 ERF units is important for robust comparison of data between different instruments and
383 laboratories. The use of flow cytometry in the expanding EV and NP research field is evolving
384 very rapidly, where the definition of what and how much can be detected is crucial. Besides the
385 inter-comparability of data, the use of calibrated values generates awareness about the range of
386 fluorescence that can be expected for single EVs [3, 20].

387 In line with previous findings we here show that linear regression curves derived from bright
388 calibration beads can be used to obtain MESF/ERF values for dim NPs and EVs, allowing inter-
389 comparability of data when the same calibrator is being used [20]. Furthermore, calibration also
390 provides information on the detection sensitivity of the instrument and makes it possible to report
391 the fluorescence threshold in calibrated units. However, we show that the robustness and
392 reproducibility of a calculated MESF or ERF value is varying upon a series of parameters,
393 including (1) the number of bead subpopulations that are used from the same calibrator set and (2)
394 the use of different calibrator beads sets, even when they are from the same manufacturer.

395 Since earlier studies already pointed out towards variabilities in MESF or ERF assignments
396 between different manufacturer's [8], we designed our study to explore the usability of bright
397 calibration bead sets from the same manufacturer for assignment of fluorescence values to dim NPs
398 and EVs. We found that the calculation of MESF or ERF values for dim submicron sized particles
399 is substantially affected by the intrinsic variation within the assignment of MESF/ERF numbers to
400 the bright calibrator beads and that this intrinsic variation also differs within calibrator sets from
401 the same manufacturer. Consequently, the slope of the regression line generated with a calibrator
402 set varies with the number of data points used for linear regression within the calibrator bead set

403 and between different calibrator bead sets. Hence, there is an uncertainty for a particular point to
404 be assigned a number of MESF/ERF based on the spread of the points [21]. To measure the impact
405 of this intrinsic uncertainty on the assignment of absolute fluorescence values to dim NPs and EVs
406 we used two distinct calibrator bead sets assigned by using the same instruments and the same
407 internal NIST traceable calibrator. Hence following the same manufacturing process. The results
408 of the two bead sets demonstrate that the two separately obtained regression lines result in different
409 MESF/ERF values to particles with fluorescent intensities well below the intensities of the
410 calibrator beads themselves. Due to the further separation of the two regression lines at the lower
411 end, increasing uncertainties exist in the MESF/ERF assignment for these dim particles. The
412 currently available commercial MESF beads are designed to calibrate brighter fluorescence
413 intensities, in the range of fluorescently labeled cells. Furthermore, available calibrators do not
414 provide uncertainty values around the assignment of the calibrator beads themselves. It is important
415 to notice that EVs are orders of magnitude smaller than cells and their fluorescent intensities based
416 on generic staining and/or on antibody labeling are far dimmer than the calibrators. To assign
417 numbers of fluorophores to these weak signals, we need to rely on extrapolation of the regression
418 line. In this proof-of-principle study, we show that the intrinsic uncertainty of the number of
419 assigned fluorescent molecules per calibrator bead introduces a higher variation in the assigned
420 MESF/ERF numbers when values need to be extrapolated for dim particles.

421 It is clear from our results with the unknown synthetic NPs and EVs that the outcome of the
422 MESF/ERF values depends on the number of data points included for linear regression analysis
423 and the calibrator set used, even if beads are provided by the same manufacturer. Since calibration
424 is an essential part in the interpretation and inter-comparability of EV data, we therefore advise
425 that (1) for benchmarking studies, the same number of data points and the same calibrator sets

426 should be used and (2) that detailed description of calibration materials is included in the
427 manuscript. To that extend, we propose that not only calibrated fluorescence values need to be
428 reported but also the calibration bead set used, including all MFI (a.u.) and MESF/ERF values of
429 the beads themselves. Clearly, a calibrator set with MESF/ERF values closer to the range of
430 fluorescence intensities of the sample of interest and with a low uncertainty of assignment is highly
431 preferable. Importantly, novel state-of-the-art flow cytometers that are designed to measure small
432 particles rely obligately on sub-micron sized beads to perform MESF/ERF calibration. These state-
433 of-the-art platforms cannot measure the 'standard' 6 µm MESF beads [22]. Therefore, there is an
434 urgent need for small and dim calibration beads that are especially designed for NPs and EVs to be
435 widely used in cross-platform and inter-laboratory comparisons.
436 In summary, our results confirm that calibration of fluorescent signals allows for the comparison
437 between platforms and provides information on the detection sensitivity of the instrument, as
438 reported in previous studies [23, 24] but also urge for rigor and reproducibility when fluorescence
439 calibration is being employed for EV and NP measurements.
440 In general, for more robust assignments of fluorescence values to NPs and EVs, there is a need for
441 multi-institutional collaborations (between research labs, companies and metrology institutions,
442 such as NIST) to produce and validate calibration materials that have low and well-characterized
443 uncertainty of assigned fluorescent values, ideally allowing for data interpolation and with a size
444 range that is compatible for all flow cytometer platforms used for analysis of NPs and EVs.
445

446 **Funding Statement**

447 This research is supported by the European Union's Horizon 2020 research and innovation
448 programme under the Marie Skłodowska-Curie grant agreement No 722148 and by the National
449 Natural Science Foundation of China (21934004 and 21627811). E. L. A. is supported by the

European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 722148.

Author Contributions

E.L.A. designed and performed experiments, analyzed data and wrote the manuscript. T.B. performed experiments and gave conceptual advice. L.W. gave technical and conceptual advice. M.M., Y.T. and X.Y. prepared materials and gave technical advice. G.J.A.A. and M.H.M.W. supervised the research, designed (performed) experiments and wrote the manuscript. G.J.A.A. and M.H.M.W. contributed equally as senior author. All authors critically reviewed and edited the manuscript.

Acknowledgments

The authors would like to thank Prof. An Hendrix (Laboratory of Experimental Cancer Research, Ghent University, Belgium) for the possibility to prepare and analyze M4T1 derived EV in her lab, Dr. Joshua A. Welsh (National Cancer Institute, Bethesda, MD) for helpful discussion and Ludo Monheim (BD Biosciences, Erembodegem, Belgium) for helpful technical support. FITC-MESF and 2 µm PE-MESF beads were kindly provided by BD Biosciences (prepared by Dr. Majid Mehrpouyan) as part of the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 722148.

Declaration of interest disclosure

Tina Van Den Broeck and Majid Mehrpouyan are both employees of BD Biosciences, a business unit of Becton, Dickinson and Company. During the course of this study, the Wauben research group, Utrecht University, Faculty of Veterinary Medicine, Department of Biomolecular Health

Sciences and BD Biosciences collaborated as a co-joined partner in the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 722148. Xiaomei Yan declares competing financial interests as a cofounder of NanoFCM Inc., a company committed to commercializing the nano-flow cytometry (nFCM) technology.

Supporting Information

Additional supporting information can be found online in the corresponding section at the end of the article.

Data S1. Author Checklist: MIFlowCyt-Compliant Items.

Data S2. MIFlowCyt-EV framework.

Figure S1. Analysis of different MESF bead sets measured on the BD Influx.

Figure S2. Analysis of instrument linearity on the BD Influx.

Figure S3. Evaluation of measurement variability of 2 µm FITC MESF beads measured on the BD Influx.

Figure S4. Evaluation of commercially available ERF Reference Particles measured on the BD Influx.

Figure S5. Gating strategy for synthetic silica NPs.

Figure S6. 200 nm fluorescent polystyrene NPs.

Figure S7. Gating strategy for external spiked-in beads and biological EV samples.

ORCID

Estefanía Lozano-Andrés https://orcid.org/0000-0002-7305-0776

Lili Wang https://orcid.org/0000-0003-2456-898X

Xiaomei Yan http://orcid.org/0000-0002-7482-6863

Marca H.M. Wauben https://orcid.org/0000-0003-0360-0311

Ger. J.A. Arkesteijn https://orcid.org/0000-0001-7739-582X

References

1. Nolan, J.P., *Flow Cytometry of Extracellular Vesicles: Potential, Pitfalls, and Prospects.* Curr Protoc Cytom, 2015. **73**: p. 13 14 1-16.
2. Lian, H., et al., *Flow Cytometric Analysis of Nanoscale Biological Particles and Organelles.* Annu Rev Anal Chem (Palo Alto Calif), 2019. **12**(1): p. 389-409.
3. Welsh, J.A., et al., *MIFlowCyt-EV: a framework for standardized reporting of extracellular vesicle flow cytometry experiments.* J Extracell Vesicles, 2020. **9**(1): p. 1713526.
4. Hoffman, R.A., *Standardization, calibration, and control in flow cytometry.* Curr Protoc Cytom, 2005. **Chapter 1**: p. Unit 1 3.
5. Wang, L. and R.A. Hoffman, *Standardization, Calibration, and Control in Flow Cytometry.* Curr Protoc Cytom, 2017. **79**: p. 1 3 1-1 3 27.
6. Gaigalas, A.K., et al., *Quantitating Fluorescence Intensity From Fluorophore: Assignment of MESF Values.* J Res Natl Inst Stand Technol, 2005. **110**(2): p. 101-14.
7. Wang, L., et al., *Quantitating Fluorescence Intensity From Fluorophores: Practical Use of MESF Values.* J Res Natl Inst Stand Technol, 2002. **107**(4): p. 339-53.
8. Hoffman, R.A., et al., *NIST/ISAC standardization study: variability in assignment of intensity values to fluorescence standard beads and in cross calibration of standard beads to hard dyed beads.* Cytometry A, 2012. **81**(9): p. 785-96.
9. van der Vlist, E.J., et al., *Fluorescent labeling of nano-sized vesicles released by cells and subsequent quantitative and qualitative analysis by high-resolution flow cytometry.* Nat Protoc, 2012. **7**(7): p. 1311-26.
10. Arkesteijn, G.J.A., et al., *Improved Flow Cytometric Light Scatter Detection of Submicron-Sized Particles by Reduction of Optical Background Signals.* Cytometry A, 2020. **97**(6): p. 610-619.
11. Yokoi, T., et al., *Periodic arrangement of silica nanospheres assisted by amino acids.* J Am Chem Soc, 2006. **128**(42): p. 13664-5.
12. Yokoi, T., et al., *Mechanism of Formation of Uniform-Sized Silica Nanospheres Catalyzed by Basic Amino Acids.* Chemistry of Materials, 2009. **21**(15): p. 3719-3729.
13. Stöber, W., A. Fink, and E. Bohn, *Controlled growth of monodisperse silica spheres in the micron size range.* Journal of Colloid and Interface Science, 1968. **26**(1): p. 62-69.
14. Giesche, H., *Synthesis of monodispersed silica powders II. Controlled growth reaction and continuous production process.* Journal of the European Ceramic Society, 1994. **14**(3): p. 205-214.
15. Vergauwen, G., et al., *Confounding factors of ultrafiltration and protein analysis in extracellular vesicle research.* Sci Rep, 2017. **7**(1): p. 2704.
16. Geeurickx, E., et al., *The generation and use of recombinant extracellular vesicles as biological reference material.* Nat Commun, 2019. **10**(1): p. 3288.
17. Welsh, J.A., et al., *FCMPASS Software Aids Extracellular Vesicle Light Scatter Standardization.* Cytometry A, 2019.
18. Consortium, E.-T., et al., *EV-TRACK: transparent reporting and centralizing knowledge in extracellular vesicle research.* Nat Methods, 2017. **14**(3): p. 228-232.
19. Bagwell, C.B., et al., *A simple and rapid method for determining the linearity of a flow cytometer amplification system.* Cytometry, 1989. **10**(6): p. 689-94.
20. Welsh, J.A. and J.C. Jones, *Small Particle Fluorescence and Light Scatter Calibration Using FCMPASS Software.* Curr Protoc Cytom, 2020. **94**(1): p. e79.

21. Draper, N.R. and H. Smith, *Applied regression analysis*. 1981, New York: Wiley.
22. Tian, Y., et al., *Protein Profiling and Sizing of Extracellular Vesicles from Colorectal Cancer Patients via Flow Cytometry.* ACS Nano, 2018. **12**(1): p. 671-680.
23. Welsh, J.A., J.C. Jones, and V.A. Tang, *Fluorescence and Light Scatter Calibration Allow Comparisons of Small Particle Data in Standard Units across Different Flow Cytometry Platforms and Detector Settings.* Cytometry A, 2020. **97**(6): p. 592-601.
24. Nolan, J.P. and S.A. Stoner, *A trigger channel threshold artifact in nanoparticle analysis.* Cytometry A, 2013. **83**(3): p. 301-5.

Main figure legends

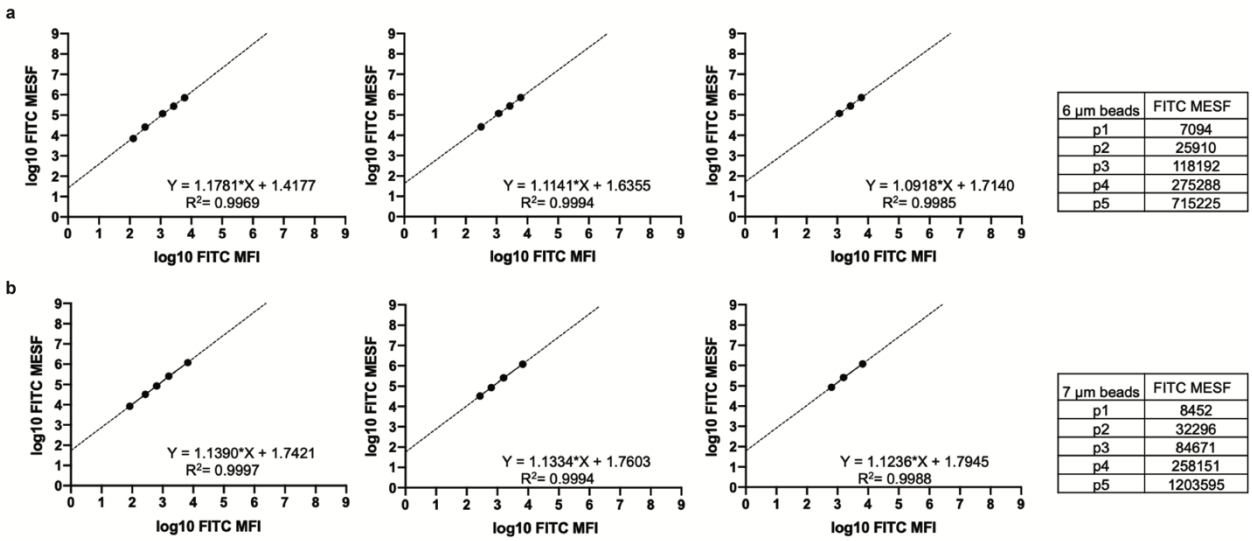
Figure 1. Comparison of the inclusion of different data points into linear regression analysis by using custom-made and commercial FITC MESF bead sets. (a) Least square linear regression analysis of custom-made 6 µm FITC MESF beads and **(b)** 7 µm Quantum™ FITC-5 MESF beads measured on the BD Influx. Provided FITC MESF for both sets (indicated in the figure) and measured FITC MFI values were transformed to log and plotted in a log-log fashion. Graphs including either five, four or three fluorescent populations (from left to right).

Figure 2. Evaluation of two different FITC MESF bead sets for the calibration of fluorescent intensities across three flow cytometer platforms. (a) Histogram overlays (axis in arbitrary units) of FITC fluorescent intensity peaks derived from the 6 µm (upper row) or the 2 µm (lower row)

573 FITC MESF beads **(b)** Table showing the median fluorescence intensity (MFI) statistic derived
574 from each of the fluorescent intensity peaks from dimmer to brighter being expressed in arbitrary
575 units as well as the assigned MESF values (right column). **(c)** Least square linear regression
576 analysis of 6 µm (black circles) and 2 µm (grey triangles) FITC MESF beads. Provided FITC
577 MESF and measured FITC MFI values were transformed to log and plotted in a log-log fashion for
578 the three platforms. **(d-g)** Bar graphs showing the expected absolute FITC MESF and the calculated
579 FITC MESF values for different unknowns as indicated in the figure. The absolute number of FITC
580 MESF was calculated for the indicated fluorescent intensity peak population and compared to the
581 expected value on the three flow cytometers. For comparison, we included a 20% variation from
582 the expected FITC MESF of the population as a reference point (green). **(h)** Table indicating the
583 expected and calculated FITC MESF values for each sample used in the analysis.
584
585 **Figure 3. MESF bead-based calibration of fluorescence signals from synthetic silica NPs. (a)**
586 Histogram overlay showing FITC fluorescence in arbitrary units (a.u.) (left), FITC MESF
587 calibrated axis based on the 6 µm (middle) and 2 µm FITC MESF beads (right) from the six
588 differently FITC-labeled 550 nm NP gated populations. **(b)** Table showing the MFI and CV as well
589 as the calculated FITC MESF values for each of the unknown populations with the percentage of
590 variation between the two calculated reference values. **(c)** Bar graph showing the calibrated FITC
591 MESF values for each of the six differently FITC-labeled populations based on the 6 µm (dark
592 green, generally lower numbers) and 2 µm (light green, generally higher numbers). Bars from 6
593 µm bead values and bars from 2 µm bead values are presented as overlays.
594
595 **Figure 4. MESF bead-based calibration of fluorescent signals from biological EV samples.** (a)
596 Analysis of EV samples by using a fluorescence threshold. Unstained EVs control (left), CFSE and

597 isotype-PE stained EVs (middle) and CFSE and CD9-PE stained EVs (right) dot plots showing
598 CFSE fluorescence Vs PE fluorescence in arbitrary units (a.u.) (upper row) or CFSE ERF Vs PE
599 MESF calibrated axis based on either the 6 µm (middle row) or 2 µm calibration beads (lower
600 row). The dashed line in each dot plot indicates the fluorescence threshold value used. Number of
601 events within the EV gate and MFI values for either CFSE or PE fluorescence (a.u.) are indicated
602 in the top row. CFSE ERF and PE MESF values are indicated based on the 6 µm (middle row) or
603 2 µm beads (lower row). **(b)** Table showing the ERF or MESF values obtained after calibration for
604 the CFSE and CD9-PE stained EVs and the percentage of variation between the use of the 6 µm or
605 2 µm bead sets. **(c)** Histogram overlays displaying fluorescence in arbitrary units from CFSE
606 stained EVs (blue) next to the 6 µm or 2 µm FITC-MESF bead set (green). **(d)** Histogram overlays
607 displaying fluorescence in arbitrary units from CD9-PE labeled EVs (purple) next to the 6 µm or
608 2 µm PE-MESF bead set (red).

609 **Figure 1**



6 µm beads	FITC MESF
p1	7094
p2	25910
p3	118192
p4	275288
p5	715225

7 µm beads	FITC MESF
p1	8452
p2	32296
p3	84671
p4	258151
p5	1203595

610

611

612

613

614

615

616

617

618

619

620

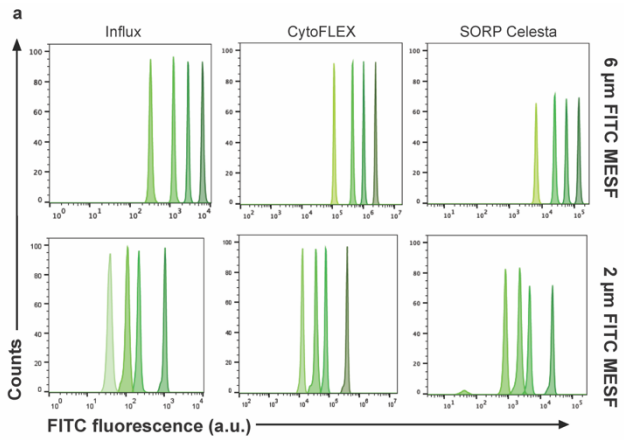
621

622

623

624

625 **Figure 2**



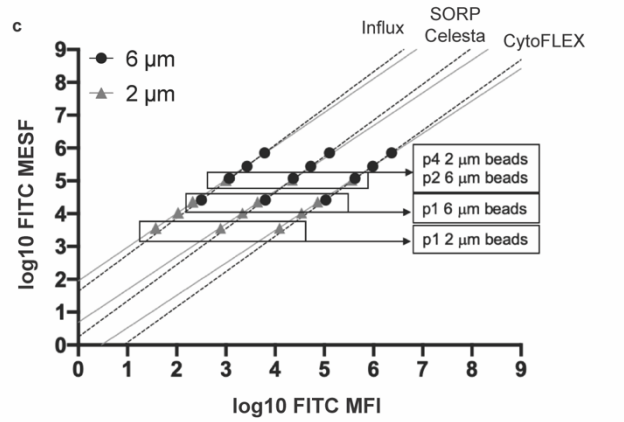
b

6 µm beads	MFI Influx	MFI CytoFLEX	MFI SORP Celesta	FITC MESF
p1	317	107336	6321	25910
p2	1173	421852	23322	118192
p3	2684	962928	53219	275288
p4	6111	2324148	127668	715225

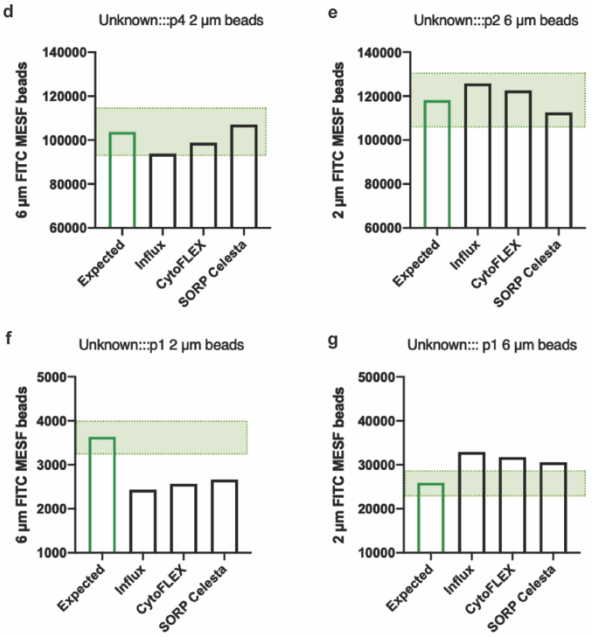
Brightness

2 µm beads	MFI Influx	MFI CytoFLEX	MFI SORP Celesta	FITC MESF
p1	37	12357	780	3634
p2	106	34544	2169	10520
p3	210	73054	4380	22841
p4	991	367458	22370	103706

Brightness



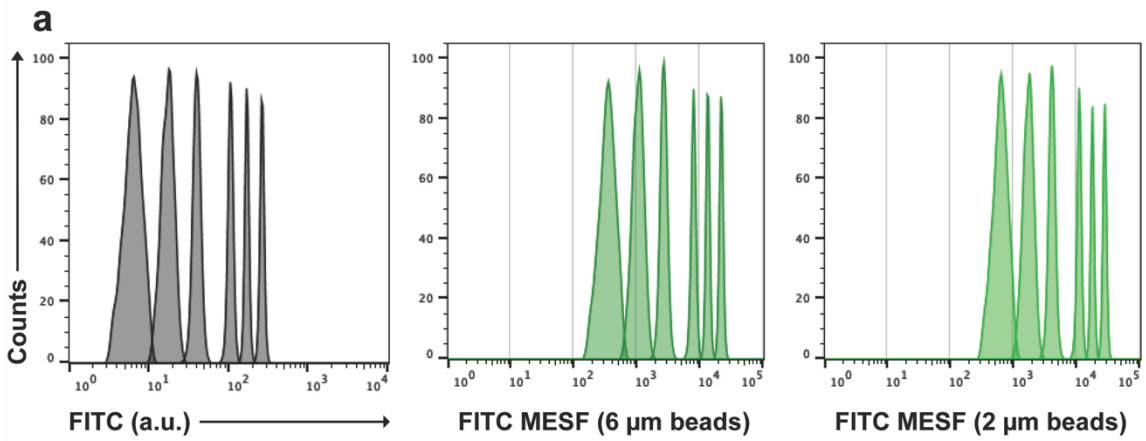
	6 µm FITC MESF beads	2 µm FITC MESF beads
Influx	Y = 1.1146*X + 1.6325 R^2 = 0.9995	Y = 1.0261*X + 1.951 R^2 = 0.9993
CytoFLEX	Y = 1.0762*X – 0.9947 R^2 = 0.9998	Y = 0.9869*X – 0.4632 R^2 = 0.9992
SORP Celesta	Y = 1.1004*X – 0.2432 R^2 = 0.9993	Y = 0.9989*X + 0.6884 R^2 = 0.9985



h

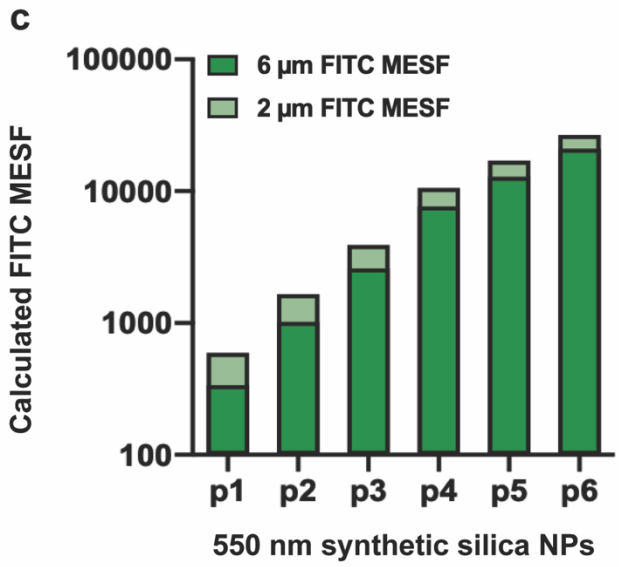
Used as calibrator	Used as unknown	Assigned FITC MESF Expected values	Calculated FITC MESF values Influx	CytoFLEX	SORP Celesta
6 µm FITC beads	p4:::2 µm FITC beads	103706	93745	98759	107046
2 µm FITC beads	p2:::6 µm FITC beads	118192	125680	122538	112554
6 µm FITC beads	p1:::2 µm FITC beads	3634	2432	2565	2665
2 µm FITC beads	p1:::6 µm FITC beads	25910	32904	31743	30550

Figure 3



b

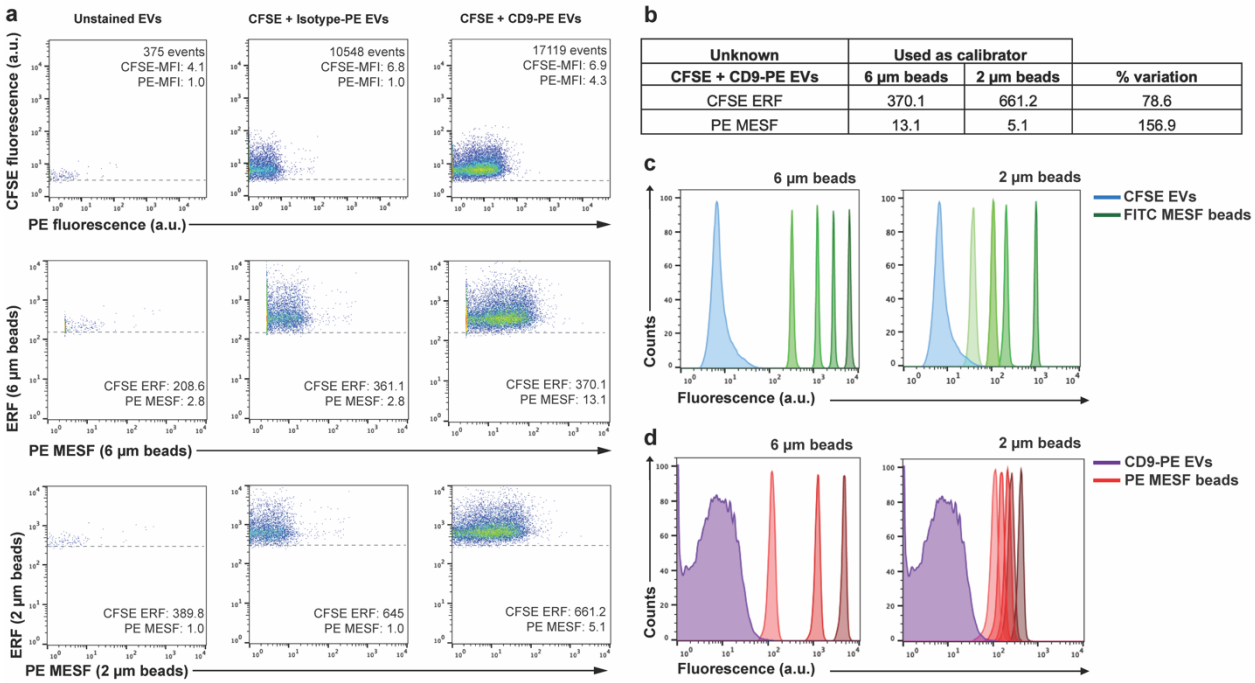
			Calculated FITC MESF values		
			Used as calibrator		
Unknown	MFI	CV	6 µm FITC beads	2 µm FITC beads	% variation
p1	6.35	25.53	337	595	76.5
p2	17.20	17.83	1022	1654	61.8
p3	39.60	10.89	2590	3894	50.3
p4	104.90	6.86	7671	10581	37.9
p5	167.00	5.12	12881	17050	32.4
p6	258.90	5.10	20995	26734	27.3



632

633

634 **Figure 4**



Unknown	Used as calibrator		
CFSE + CD9-PE EVs	6 µm beads	2 µm beads	% variation
CFSE ERF	370.1	661.2	78.6
PE MESF	13.1	5.1	156.9

635

636