Cite as: D. Di Carlo *et al.*, *Science* 10.1126/science.aav1429 (2019).

Comment on "Ghost cytometry"

Dino Di Carlo[1]*, Fumihito Arai[2], Keisuke Goda[3], Tony Jun Huang[4], Yu-Hwa Lo[5], Nao Nitta[3,6], Yasuyuki Ozeki[7], Kevin Tsia[8], Sotaro Uemura[9], Kenneth K. Y. Wong[8]

[1]Department of Bioengineering, University of California, Los Angeles, CA, USA. [2]Department of Micro-Nano Systems Engineering, Nagoya University, Nagoya, Japan. [3]Department of Chemistry, University of Tokyo, Tokyo, Japan. [4]Department of Mechanical Engineering and Materials Science, Duke University, Durham, NC, USA. [5]Department of Electrical and Computer Engineering, University of California, San Diego, CA, USA. [6]Japan Science and Technology Agency, Saitama, Japan. [7]Department of Electrical Engineering and Information Systems, University of Tokyo, Tokyo, Japan. [8]Department of Electrical and Electronic Engineering, University of Hong Kong, Hong Kong, China. [9]Department of Biological Sciences, University of Tokyo, Tokyo, Japan.

*Corresponding author. Email: dicarlo@ucla.edu

Ota *et al.* (Reports, 15 June 2018, p. 1246) report using pseudo-random optical masks and a spatial-temporal transformation to perform blur-free, high–frame rate imaging of cells in flow with a high signal-to-noise ratio. They also claim sorting at rates of 3000 cells per second, based on imaging data. The experiments conducted and results reported in their study are insufficient to support these conclusions.

Ota *et al.* (*1*) proposed an approach to perform image-based flow cytometry and cell sorting that has attracted substantial attention because high throughput (>3000 cells/s) and a high signal-to-noise ratio (SNR) were claimed. For example, on the basis of these assertions, the introductory commentary (*2*) referred to the system as an "ultrahigh-speed fluorescence imaging–activated cell sorter." Unfortunately, the data provided in the paper do not support these claims.

First, there is insufficient evidence that sorting is being performed on two-dimensional (2D) morphological information. A single image of a cell in flow is provided throughout the paper and there is no comparison to a validated image of the same cell, making it impossible to claim that this cell's image is accurately reconstructed or blur-free. Data in high-speed flow on standard beads of different sizes and intensities are needed to demonstrate whether the technique provides accurate morphological (e.g., size) and intensity information based on well-established validation techniques for flow cytometry.

Ghost cytometry (GC) uses 1D temporal information (i.e., waveform) for sorting, and the nonmorphological data present in these signals appear sufficient for cell classification in the cell mixtures that were evaluated. In fact, using 1D nonmorphological information (e.g., intensity and waveform length) available in an online data repository (Ota *et al.*, reference 32) yields a clear separation of the cell populations investigated in the paper (Fig. 1) and may be the feature that allows differentiation between cell populations in the sorting experiment. For the sorting experiment performed between MCF-7 cells and peripheral blood mononuclear cells (PBMCs), the size difference between these cells is quite large, and therefore the length and total intensity of the GC waveform are sufficient features for accurate sorting (Fig. 1A). Although Ota *et al.* indicate that MCF-7 and MIA PaCa-2 cells are of similar size without reporting data, previous reports indicate a significant difference in the diameters of these cells measured side-by-side (*3*). As expected on the basis of size differences, the waveform lengths and peak intensities for these cells also form distinct clusters (Fig. 1B). The waveform length is affected by other nonmorphological factors besides cell diameter, such as velocity differences. Cells can have differences in velocity resulting, for example, from differences in cell size or deformability that drive lateral migration in the microfluidic flow channel (*4*). This is likely in the sorting experiments, given the off-center hydrodynamic focusing design used (figure S7 of Ota *et al.*). The large distance between the hydrodynamic sheathing structure and sorting structure shown in their figure S7, combined with cell velocity differences, also has the potential to lead to incorrect sort timings and "sorting" based on the timing accuracy. Even a small diameter (a) difference, scaling with a^3, could be sufficient to lead to differences in lateral migration, downstream velocity, and accurate classification (*5*). One cannot consider sorting based on size-dependent migration differences as image-based sorting, just as one cannot consider sorting based on fluorescence peak width (*6*) or scattered light intensity (correlating with size or granularity) to be image-based sorting.

This concern that sorting is based on a separate nonmorphological factor is exacerbated by the fact that the MCF-7 cells and MIA PaCa-2 cells do not appear morphologically different. The authors highlight this point in the paper. Evidence to support sorting based on image/2D morphological information of the cells might come from a separate classification of the two populations of cells based on fluorescence microscopy images (*7*); however, such re-

sults were not presented. The authors instead hint that there may be information accessible only in the GC waveforms, but compressed sensing does not enable extraction of new information that is not already present in an image (*8*).

Second, on the basis of data provided by the authors, the claimed throughput values are not physically accurate. The authors claim that sorting experiments "were performed at a throughput rate of ~3000 cells/s." The authors also claim that waveform lengths, δ, of 100 μs and 300 μs correspond to throughputs, $N = \delta^{-1}$, of ~10,000 cells/s and ~3000 cells/s, respectively, for imaging and sorting, implying that waveforms for neighboring cells cannot overlap (Fig. 2, A and B). In order to realize these throughput values, flowing cells need to arrive at the optical GC analysis point and sorting point with perfectly uniform intervals. However, the arrival of flowing cells is random and is a Poisson process, resulting in nonuniform cell-cell intervals (*9*–*11*). In a Poisson process, the time interval T between successive events will follow a cumulative probability distribution function of the form $P(T < t) = 1 - \exp(-Nt)$, where N is the throughput (Fig. 2C). Therefore, assuming the claimed throughput of $N = \delta^{-1}$ would yield $P(T < \delta) \approx 64\%$. This means that 64% of events would have a shorter interval than the interval of the waveform, δ, leading to overlapping waveforms, which the authors imply would prevent analysis (Fig. 2D). This contradictory information leads us to ask: What was the actual "experimental" throughput for image-based sorting experiments? In figure 5 of Ota *et al.*, it appears that a total of 2000 cells were sorted per experiment; with a throughput of 3000 cells/s, this would be the result of less than 1 s of sorting.

Third, although no SNR value or minimum sensitivity is reported for the flowthrough implementation, analysis of the presented data shows that these values appear substantially lower than the scanning system. There is a SNR challenge at high flow rates with this technique. This is observed in figure 3D (ii) of Ota *et al.*, for example, where a significant noise level is seen in the baseline signal relative to the cell-induced signal, leading to a very low SNR. In addition, no temporal waveforms or reconstructed images are provided for the sorting experiments that make use of a polydimethylsiloxane microfluidic chip, which is expected to have higher autofluorescence (*12*). It is misleading to choose performance metrics from different systems (nonflow, flow without sorting, and flow with sorting) and combine these when reporting the overall performance of the system in the image-flow sorting implementation that is of the most general interest.

REFERENCES

1. S. Ota, R. Horisaki, Y. Kawamura, M. Ugawa, I. Sato, K. Hashimoto, R. Kamesawa, K. Setoyama, S. Yamaguchi, K. Fujiu, K. Waki, H. Noji, Ghost cytometry. *Science* **360**, 1246–1251 (2018). doi:10.1126/science.aan0096 Medline
2. V. Vinson, Seeing ghosts. *Science* **360**, 1198 (2018). http://science.sciencemag.org/content/360/6394/twis
3. S. C. Hur, N. K. Henderson-MacLennan, E. R. B. McCabe, D. Di Carlo, Deformability-based cell classification and enrichment using inertial microfluidics. *Lab Chip* **11**, 912–920 (2011). doi:10.1039/c0lc00595a Medline
4. K. D. Nyberg, S. L. Bruce, A. V. Nguyen, C. K. Chan, N. K. Gill, T.-H. Kim, E. K. Sloan, A. C. Rowat, Predicting cancer cell invasion by single-cell physical phenotyping. *Integr. Biol.* **10**, 218–231 (2018). doi:10.1039/C7IB00222J Medline
5. D. Di Carlo, J. F. Edd, K. J. Humphry, H. A. Stone, M. Toner, Particle segregation and dynamics in confined flows. *Phys. Rev. Lett.* **102**, 094503 (2009). doi:10.1103/PhysRevLett.102.094503 Medline
6. J. G. Monroe, J. C. Cambier, Sorting of B lymphoblasts based upon cell diameter provides cell populations enriched in different stages of cell cycle. *J. Immunol. Methods* **63**, 45–56 (1983). doi:10.1016/0022-1759(83)90208-9 Medline
7. C. Brasko, K. Smith, C. Molnar, N. Farago, L. Hegedus, A. Balind, T. Balassa, A. Szkalisity, F. Sukosd, K. Kocsis, B. Balint, L. Paavolainen, M. Z. Enyedi, I. Nagy, L. G. Puskas, L. Haracska, G. Tamas, P. Horvath, Intelligent image-based in situ single-cell isolation. *Nat. Commun.* **9**, 226 (2018). doi:10.1038/s41467-017-02628-4 Medline
8. E. J. Candès, J. K. Romberg, T. Tao, Stable signal recovery from incomplete and inaccurate measurements. *Commun. Pure Appl. Math.* **59**, 1207–1223 (2006). doi:10.1002/cpa.20124
9. T. Lindmo, D. C. Peters, R. G. Sweet, in *Flow Cytometry and Sorting* (Wiley-Liss, ed. 2, 1990), pp. 145–169.
10. J.-C. Baret, O. J. Miller, V. Taly, M. Ryckelynck, A. El-Harrak, L. Frenz, C. Rick, M. L. Samuels, J. B. Hutchison, J. J. Agresti, D. R. Link, D. A. Weitz, A. D. Griffiths, Fluorescence-activated droplet sorting (FADS): Efficient microfluidic cell sorting based on enzymatic activity. *Lab Chip* **9**, 1850–1858 (2009). doi:10.1039/b902504a Medline
11. D. J. Collins, A. Neild, A. deMello, A.-Q. Liu, Y. Ai, The Poisson distribution and beyond: Methods for microfluidic droplet production and single cell encapsulation. *Lab Chip* **15**, 3439–3459 (2015). doi:10.1039/C5LC00614G Medline
12. A. Piruska, I. Nikcevic, S. H. Lee, C. Ahn, W. R. Heineman, P. A. Limbach, C. J. Seliskar, The autofluorescence of plastic materials and chips measured under laser irradiation. *Lab Chip* **5**, 1348–1354 (2005). doi:10.1039/b508288a Medline

ACKNOWLEDGMENTS

Competing interests: N.N. is the president of, and holds shares in, K.K. CYBO (Tokyo, Japan). K.G. is a shareholder in K.K. CYBO and Cupido Inc.

20 August 2018; accepted 28 March 2019
Published online 19 April 2019
10.1126/science.aav1429

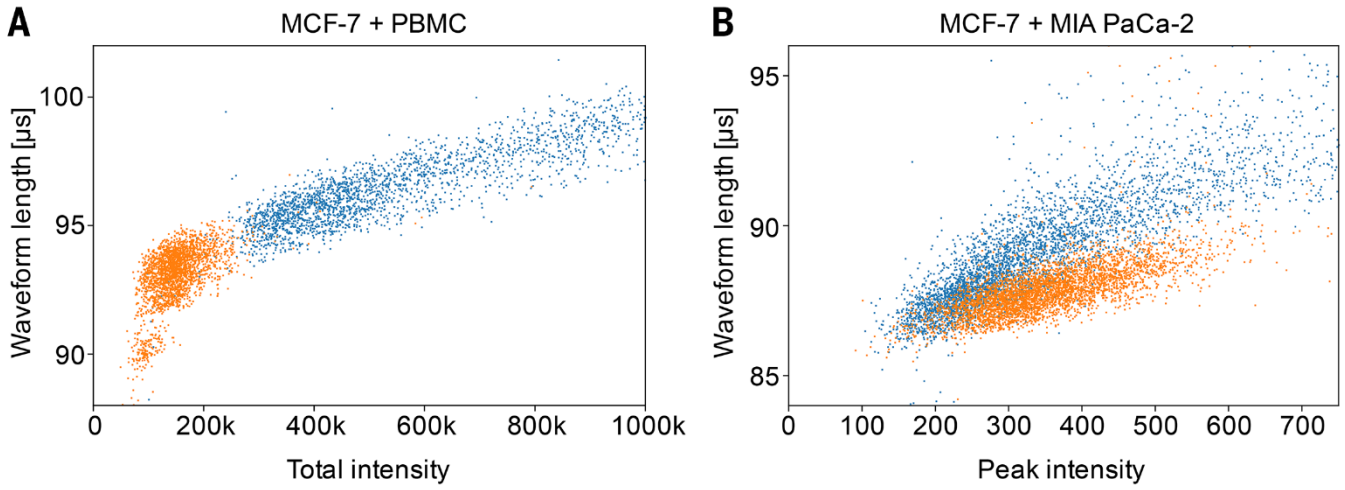

Fig. 1. Raw data from Ota *et al.*'s paper indicate that nonmorphological features of waveforms (total or peak intensity and waveform length) cluster cell types distinctly. (**A**) Scatterplots of MCF-7 cells (blue) and PBMCs (orange) in which the total intensity is plotted against the waveform length of the fixable green (FG) signal. (**B**) Scatterplots of MCF-7 cells (blue) and MIA PaCa-2 cells (orange) in which the peak intensity is plotted against the waveform length of the FG signal. Both (A) and (B) show that the two different cell populations of interest cluster separately.

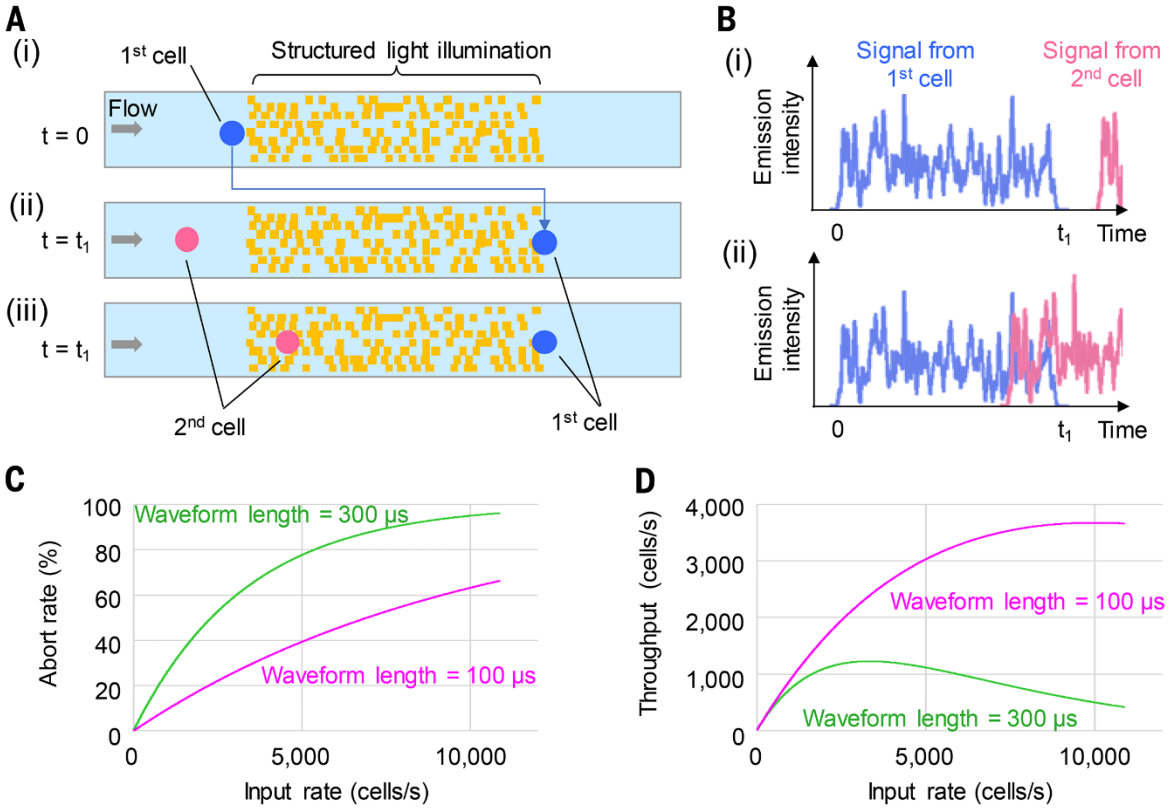

Fig. 2. Throughput limit caused by overlapping waveforms. (A) Schematic of two consecutive cells detected by structured light illumination. We assume that the first cell enters the illumination at $t = 0$ (i) and exits at $t = t_1$ (i.e., waveform length = t_1) followed by the second cell, which can be either outside (ii) or inside (iii) the illumination at $t = t_1$. (B) Schematic of waveforms obtained from the first and second cells, when the second cell is outside (i) or inside (ii) the illumination when the first cell exits the illumination. In the latter case, the event needs to be aborted, in the same manner as a "doublet" in conventional fluorescence-activated cell sorting. (C) Calculated abort rate for various input rates (the number of cells flowing in the channel per second) when cells flow with random intervals. Green and magenta curves show the abort rates when the waveform length is 300 µs and 100 µs, respectively. (D) Throughput of events, obtained from the abort rate and the input rate in (C). The peaks of the curves indicate the maximum achievable throughputs, which are 1226 cells/s and 3679 cells/s for the waveform lengths of 300 µs and 100 µs, respectively. The throughput value claimed contradicts these theoretical limits.

Comment on "Ghost cytometry"

Dino Di Carlo, Fumihito Arai, Keisuke Goda, Tony Jun Huang, Yu-Hwa Lo, Nao Nitta, Yasuyuki Ozeki, Kevin Tsia, Sotaro Uemura and Kenneth K. Y. Wong

Science **364** (6437), eaav1429.
DOI: 10.1126/science.aav1429