June 7, 2017

Via ECF
Hon. Analisa Torres
United States District Court
Southern District of New York
500 Pearl Street
New York, NY 10007

Re: *Floyd v. City of New York*, 08 Civ. 1034 (AT)

Dear Judge Torres:

On behalf of Plaintiffs in the above-captioned matter, we write in response to the Fifth Report of the Independent Monitor (ECF 554), which was filed with this Court on May 30, 2017. ("The Report"). Specifically, we write to bring to Your Honor's attention several concerns we have with the Report that we raised with the Monitor prior to May 30th, and which we feel strongly must be addressed in any future stop-and-frisk data analyses that will be used by the Monitor, the parties and the Court to assess the New York Police Department's (NYPD) compliance with the Court-ordered reforms in this case.

First, at several points the Report draws conclusions and presents a narrative about declining or disappearing racial disparities in the NYPD's stop-and-frisk data which are not supported by the results of the Monitor's own analyses of that data. For example, on page 12, when discussing the significant decrease in the number of reported stops in the 2013-15 period, the Report states that "[t]he decline in reported stops was primarily driven by steep decreases in the number of stops of Blacks and Hispanics relative to Whites," suggesting that stops of Blacks and Hispanics decreased at a higher rate than did stops of Whites. However, the numbers listed in Table 1 on page 7 of the Report clearly show that, during this same period, the numbers of recorded stops of all three groups actually declined by the same percentage, about 88%, so that Black and Hispanic pedestrians continue to comprise the vast majority, more than 80%, of the persons stopped each year by the NYPD.

The Report makes similarly misleading statements about racial disparities in post-stop outcomes. On page 6 the Report states that the disparities in weapons recovery rates between stops and searches of Blacks and stops and searches of Non-Hispanics "no longer existed" by 2015, even though Table 7 on page 40 indicates that, in 2015, the weapons hit rate for non-Hispanic stops was 35% higher than the hit rate for Black stops and the weapons hit rate for non-Hispanic searches was over 21% higher than the hit rate for Black searches. Similarly, on page 36 the Report states that there "was no evidence of racial disparities in [many of] the stop outcomes for Blacks and Hispanics in 2014 and 2015 relative to similarly situated Non-Hispanics" even though, as Table 6 on pages 36-37 indicates, the racial disparities for many of the stop outcomes in those two years were comparable and in some cases even numerically larger than the disparities present in 2013, which the Monitor characterized as "significant."

These latter two misleading statements about the absence of significant stop outcome and hit rate racial disparities are in turn a function of the Report's improperly exclusive reliance on statistical significance testing to determine whether racial disparities exist and are meaningful. However, statistical significance testing can only be used to determine the probability that a disparity observed in a random sample represents a real disparity in the population from which that sample is drawn; it says little about the magnitude of that disparity. *See Waisome v. Port. Auth. of N.Y. & N.J.*, 948 F.2d 1370, 1376, 1379 (2d Cir. 1991) (citing H Blalock, *Social Statistics* 163 (2d ed. 1972)). Moreover, because smaller data sets have lower statistical power, racial disparities in those data sets are less likely to be statistically significant even when the magnitude of such disparities is comparable to, or even greater than, statistically significant disparities found in larger data sets. Because, as the Report makes clear, the number of stops recorded in 2015 is a fraction of the number recorded in 2013, comparisons based on statistical significance, rather than actual disparity size, are not meaningful.

Thus, it is just not accurate to say that the racial disparities in post-stop outcomes or hit rates "did not exist, "disappeared" or "decreased" in 2014 and 2015. At most, one could say that the disparities present in 2014 and 2015 were more likely than the 2013 disparities to have occurred by chance. But even on that point, the Report's conclusion that many of the 2014 and 2015 post-stop outcome and hit rate disparities were not significant is undermined by the Report's use of an overly stringent $p<.01$ (1 time out of 100 probability of occurring by chance) standard for statistical significance, rather than the $p<.05$ (1 time out of 20) standard widely used by the Courts of this Circuit when empirically analyzing claims of racial discrimination in civil rights cases and by social scientists, including the Monitor's own expert, in published peer-reviewed statistical studies of racial disparities in police stop-and-frisk activity.[1] Indeed, Plaintiffs' expert's statistical analysis presented at trial, on which the Court based its liability finding for Plaintiffs' Equal Protection claim, employed the $p<.05$ (1 time out of 20) standard for statistical significance. *See* Pl. Trial Exs. 411, 412, 417; *Floyd* Liab. Op. (ECF 373) at 58-60.

The Monitor appears to use the $p<.01$ (1 time out of 100) standard as a way to avoid so-called Type I errors or falsely positive findings of racial disparities that are not actually meaningful but instead occurred due to random chance. *See* Report at 33 n.11. However, because the Monitor analyzed data on all, rather than a random sample, of recorded NYPD stops in a given year, the risk of false positives is minimal; all observed racial disparities in the data are real. Instead, as illustrated by the above-mentioned stop outcome and hit-rate disparity results, claiming no racial disparities due to statistical non-significance under the $p<.01$ (less than 1 out of 100) standard risks committing Type II errors, or false negatives, which, given the current stage of this case and the Monitor's mandate, would cause much greater harm than would false positives. While false positive findings of racial disparities may result in continued monitoring of the NYPD's stop-and-frisk practices even after such practices are no longer racially biased, false

---

[1] *See, e.g.*, *Ottaviani v. State Univ. of N.Y.*, 875 F.2d. 365, 371-72 (2d Cir. 1989); *Waisome*, 948 F.2d at 1376; *Smith v. Xerox Corp.*, 196 F.3d 358, 366 (2d Cir. 1999); *Wright v. Stern*, 450 F.Supp.2d 335, 363 (S.D.N.Y. 2006); Geoffrey P. Alpert, John M. MacDonald, and Roger G. Dunham, "Police Suspicion and Discretionary Decision Making in Citizen Stops," 43 CRIMINOLOGY 407 (May 2005); Greg Ridgeway, *Analysis of Racial Disparities in the New York Police Department's Stop, Question and Frisk Practices* (RAND Corp. 2007); Amanda Geller and Jeffrey Fagan, "Pot as Pretext: Marijuana, Race, and the New Disorder in New York City Street Policing," 7 JOURNAL OF EMPIRICAL LEGAL STUDIES 591 (Dec. 2010).

negative findings of no racial disparities could bring a premature end to federal oversight of a police department that this Court has already found engaged in a decade-long, widespread practice of racially-biased stops and frisks before such practice has ceased, which would in turn result in continued violations of thousands of New Yorkers' constitutional rights. Thus, to the extent that statistical significance testing is an appropriate way to analyze certain of the racial disparities in the NYPD stop-and-frisk data, Plaintiffs ask that, going forward, all statistical tests conducted by the Monitor and the parties to assess the NYPD's compliance with the Court-ordered reforms employ the $p<.05$ (less than 1 out of 20) standard for statistical significance.

Finally, while the Report states at several points that racial disparities in stops, stop outcomes and hit rates, "diminish[ed]" or "improved" over time (*see* Report at 4, 5, 7), none of the statistical analyses conducted by the Monitor actually test for changes in disparities over time. There are fairly straightforward statistical tests for change over time, and they could be carried out on these data. Instead, in each of his analyses, the Monitor separately calculated and compared the levels of statistical significance (and not magnitude) of the racial disparities in 2013 stops, 2014 stops, and 2015 stops, without assessing whether and the extent to which the *size* of the disparities diminished over the course of that three-year period. Again, because larger samples have more statistical power (i.e., are more likely to yield statistical significance, all else being equal), comparisons between 2013 and later years when there are many fewer stops (i.e., lower statistical power) are confounded if the criterion is statistical significance. Moreover, the Monitor did not compare the disparities in the 2013-15 data to disparities in the data from the 2004-12 period during which, as the Court found, the NYPD had conducted racially-biased stops. Thus, because the degree to which racial disparities in the annual stop data diminish from those that existed at the time of trial in this case will likely be an important measure of the NYPD's compliance with the Court-ordered racial profiling-related reforms in this case, the Monitor should include a direct test of change in racial disparities over time in his future stop-and-frisk data analyses that encompasses data from 2012.

Thank you for your time and consideration.

Respectfully submitted,

\s\Darius Charney
Darius Charney
Britney Wilson
CENTER FOR CONSTITUTIONAL
RIGHTS

\s\ Jonathan C. Moore
Jonathan C. Moore
Joshua Moskovitz
BELDOCK, LEVINE & HOFFMAN, LLP

Attorneys for Plaintiffs