1 April 2021

Committee D02 on Petroleum Products, Liquid Fuels, and Lubricants
Subcommittee D02.040L on Gas Chromatography Methods

Research Report: D02-2016

Interlaboratory Study to Establish Precision Statements for ASTM D5580-21, Standard Test Method for Determination of Benzene, Toluene, Ethylbenzene, p/m-Xylene, o-Xylene, C9 and Heavier Aromatics, and Total Aromatics in Finished Gasoline by Gas Chromatography

Technical Contact:
Frank DiSanzo
Frank.p.disanzo@exxonmobil.com

Relative Bias Assessment of D5580 as an Alternative to D5769 for Determination of Benzene in Gasoline

Richard Stanley, Ph.D.
under contract to The American Petroleum Institute

ASTM International
100 Barr Harbor Drive
West Conshohocken, PA 19428-2959

Summary

This study was conducted to establish that ASTM Standard Method D5580[1] meets EPA requirements for use as an alternative method to determine the volume percent of benzene in gasoline. The study was based on data acquired by ASTM's Proficiency Testing Program, on 82 reformulated gasolines distributed to participating laboratories between 2004 and 2014. The fuels were tested by multiple laboratories, each lab using one or both of D5580 and ASTM Standard Practice D5769[2], the current EPA regulatory method. This report documents the analyses of these data in conformance with Standard Practice D6708[3], as required by EPA regulations promulgated in 2015[4].

This report is based on the data used for two earlier reports: RR:D02:1814, which documents the analyses supporting the qualification of D5580 as an alternative to D3606[5] (the regulatory method for benzene in gasoline in 2015, when it was written); and RR:D02:1816, for qualification of D5769 as an alternative to D3606. The performance reproducibilities demonstrated by the data were estimated in these earlier studies, for both D5580 and D5769, and were shown to be incompatible with the precisions published in the methods (which have not changed in the meantime, for either method,). The D6708 relative bias assessments of the 2015 reports did not include an evaluation of D5580 relative to D5769, which is needed now, as the regulatory method has been (or is about to be) changed to D5769.

The principal findings and conclusions of the study are as follows:

- The precision performance of both methods were significantly better than would have been expected based on the reproducibilities published in the methods. Averaged over the fuels of this study, the calculated reproducibilities were 45% and 34% smaller than the reproducibilities published in the active versions of the standards as of this writing: D5769-15 and D5580-15, respectively.

- The data meet the requirements for analysis by D6708, in that there were an adequate number of fuels, there was adequate variation among concentrations of the fuels, and there was sufficient correlation between the methods. Two relative bias evaluations were carried out. The first based only on RFG fuels, for all of which there were at least 17 non-outlying results available by both methods. The second evaluation included the RFG fuels together with the conventional motor gasoline fuels. None of the conventional fuels had been measured by 17 or more labs using D5769[6].

---

[1] D5580 *Standard Test Method for Determination of Benzene, Toluene, Ethylbenzene, p/m-Xylene, o-Xylene, C9 and Heavier Aromatics, and Total Aromatics in Finished Gasoline by Gas Chromatography*

[2] D5769 *Standard Test Method for Determination of Benzene, Toluene, and Total Aromatics in Finished Gasolines by Gas Chromatography/Mass Spectrometry*

[3] D6708, *Standard Practice for Statistical Assessment and Improvement of Expected Agreement Between Two Test Methods that Purport to Measure the Same Property of a Material*.

[4] Federal Register / Vol. 80, No. 33 / Thursday, February 19, 2015 / Rules and Regulations 9078 -9124

[5] *D3606 Standard Test Method for Determination of Benzene and Toluene in Finished Motor and Aviation Gasoline by Gas Chromatography*

[6] EPA regulations promulgated in 2015 required that in order to qualify a method as an alternative to the regulatory method (then D3606) for measuring in gasoline, there must be a successful D6708 bias evacuation of the relative

- For both evaluations, a small constant bias was observed between the methods. From the RFG-only evaluation, the correction equation was calculated as

$$C_{D5769\,predicted} = C_{D5580} - 0.003,$$

whereas from the RFG-MG combined evaluation, the bias correction was calculated to be only -0.002. Note that, in both cases, the bias corrections are smaller than the reporting resolutions of the methods. (Both are 0.01 vol%).

- From both evaluations, the variability of residuals was consistent with that to be expected due to measurement error alone: there were no discernable sample-specific biases. Consequently, the between-methods reproducibility[7] is the root mean square of the reproducibilities of the two methods.

- Average benzene concentrations of the 82 reformulated gasolines were from 0.06% to 1.71% when measured by D5769 and from 0.06% to 1.71% when measured by D5580. The RFG-only bias analysis is applicable to D5580 results between 0 and 1.84%.

- Average benzene concentrations of the 98 gasolines of the combined evaluation were from 0.06% to 2.20% when measured by D5769 and from 0.06% to 2.18% when measured by D5580. The combined fuels bias analysis is applicable to D5580 results between 0 and 2.34%.

These conclusions, when taken together, affirm that both relative bias evaluations meet the requirements of D6708, with passing category A3: bias correction with no sample-specific biases.

Statistical Report

The Data

API obtained from ASTM the following raw results collected from its Proficiency Test Program (PTP) participants: measurements of the volume percent benzene by both methods on 106 fuels distributed by the Reformulated Gasoline (RFG) program distributed between January, 2006 (RFG0601) and October, 2014 (RFG1410) and measurements on 17 fuels distributed by the Motor Gasoline (MG) program between April, 2009 and October, 2014. (The RFG program

bias between the two methods. An additional requirement was that for every fuel in the evaluation, for both methods, there had to be 17 or more non-outlying results available from different laboratories. This requirement may no longer be in force.

[7] From D6708: "*between methods reproducibility (R_{XY}), n*—a quantitative expression of the random error associated with the difference between two results obtained by different operators using different apparatus and applying the two methods *X* and *Y*, respectively, each obtaining a single result on an identical test sample, when the methods have been assessed and an appropriate bias-correction has been applied in accordance with this practice; it is defined as the 95 % confidence limit for the difference between two such single and independent results.