

UNITED STATES DISTRICT COURT
WESTERN DISTRICT OF NEW YORK

CRA HOLDINGS US, INC. AND SUBSIDIARIES,

Plaintiffs,

v.

UNITED STATES OF AMERICA,

Defendant.

**DECISION
and
ORDER**

15-CV-239W(F)

APPEARANCES:

ZERBE FINGERET FRANK & JADAV
Attorneys for Plaintiffs
JEFFERSON H. READ,
ROBERT G. WONISH, II, of Counsel
3009 Post Oak Boulevard, Suite 1700
Houston, Texas 77056

CAROLINE D. CIRAOLO
Principal Deputy Assistant Attorney General
Tax Division, U.S. Department of Justice
Attorney for Defendant
TERESA MARIE ABNEY
JAMES M. STRENDJORD, of Counsel
PO Box 55
Ben Franklin Station
Washington, DC 20044

In this tax refund action, Plaintiffs claimed a refund of \$419,924 for Plaintiffs' 2002 tax year and \$1,029,402 for Plaintiffs' 2003 tax year ("the tax years") based on Plaintiffs' assertion of a research and development ("R&D") tax credit, under 26 U.S.C. § 41, allegedly available to Plaintiffs for 6,100 remediation projects in which Plaintiffs engaged during the tax years. After filing income tax returns for the tax years, Plaintiffs, a national environmental engineering firm specializing in assisting clients in remediation projects involving environmental contamination, retained a consulting firm, Axiom

Custom Business Solutions (“Axiom”) to review its R&D activities and to produce a report that would allow Plaintiffs to file amended returns for tax years in which Plaintiffs would request a credit based on qualified research expenses (“QRE” or “QREs”) for the tax years, as defined by 26 U.S.C. § 41(a)(1)(A) and (b)(2)(B), resulting in a claim for a refund of taxes paid by Plaintiffs for the tax years. Based on the Axiom study of Plaintiffs’ R&D activities, Plaintiffs filed amended returns for the tax years requesting a refund. Following an audit, the Internal Revenue Service denied Plaintiffs’ refund request.

Included in such QREs are labor costs for conducting “qualified research” which must (1) “represent research and development costs in the experimental or laboratory sense,” (2) “for all purposes of discovering information which is technological in nature,” *i.e.*, information that relies “on principles of the physical or biological sciences, engineering, or computer science,” 26 C.F.R. § 1.41-4(a)(4), (3) the information discovered must be “useful in the development of a new or improved business component,” *i.e.*, a product, process, software, technique, formula, or invention the taxpayer intends to sell, lease, license, or use in its trade or business, 26 U.S.C. § 41(d)(2)(B), which is “new of different” from the taxpayer’s existing products, etc., see *Trinity Indus., Inc. v. United States*, 691 F.Supp.2d 688, 696 (N.D. Tex. 2010), *aff’d*, 757 F.3d 400, 412 (5th Cir. 2014), and (4) substantially all of the research activities must involve elements of “a process of experimentation,” 26 U.S. §§ 41(d)(1) & (3)(A), *i.e.*, “research and development costs in the experimental or laboratory sense” to “discover information that would eliminate uncertainty concerning the development or improvement of a product.” 26 C.F.R. § 1.174-2(a)(1). To qualify as a QRE, the related

research must precede commercial production, may not constitute an adaptation of an existing business component, may not result from research funded by other than the taxpayer, and must take place within the United States. 26 U.S.C. § 41(d)(4)(A), (B), (C), (F), (H). Taxpayers asserting an R&D tax credit are also required to retain the records necessary to show an entitlement to the credit. 26 C.F.R. § 1.41-2(d). The court's determination of the merits of a taxpayer's claim is *de novo* based on the facts presented to the court. See *Trinity Indus.*, 757 F.3d at 415. Tax credits are narrowly construed, *Shami v. Comm'r*, 741 F.3d 560, 567 (5th Cir. 2014), and a tax payer has the burden to demonstrate compliance with each of the statutory requirements. *United Stationers, Inc. v. United States*, 163 F.3d 440, 443 (7th Cir. 1998); *Research, Inc. v. United States*, 1995 WL 560140 at *3 (D.Minn. June 21, 1995). See also *Lenihan v. Comm'r*, 296 Fed.App'x. 160, 162 (2d Cir. Oct. 17, 2008) (disallowing deductions and capital loss carry over for which taxpayer produced no supporting evidence).

In this case, Plaintiffs initially claimed an R&D tax credit for 6,100 R&D "projects" listed by Plaintiffs, Dkt. 29, Exh. 8, however, Plaintiffs later produced a revised list of its R&D projects at issue ("Plaintiffs' projects") which included 4,643 projects. Dkt. 35 at 5 n. 1. Following filing of the Case Management Order on May 5, 2016 (Dkt. 20), directed to whether sampling could be used to narrow the number of projects for purposes of litigating this case, and the parties' subsequent disagreement as to the feasibility of conducting discovery and trial based on a sample of Plaintiffs' projects, Defendant served, on July 6, 2016, Defendant's Second Set of Interrogatories particularly limited to Superseding Interrogatory No. 1, Dkt. 29, Exh. 5 ("Defendant's Superseding Interrogatory"), which requested Plaintiffs to respond to 13 questions seeking factual

information including the eight required elements for a QRE under § 41, upon which Plaintiffs' R&D tax credit claims are based, specifically, Plaintiffs' asserted QREs. In Plaintiffs' Amended Objections and Responses To Second Interrogatories (Dkt. 29, Exh. 10) served September 12, 2016, Plaintiffs objected to Defendant's Interrogatories 1(a)-(c) based on burdensomeness, lack of relevancy to the need to select a sample of the projects for purposes of discovery and trial, and referred Defendant to the list of Plaintiffs' projects as listed in the Axiom report previously produced to Defendant. As to Interrogatory No. 1(d), Plaintiffs stated all the projects' research activity was performed in the United States.

By papers filed September 19, 2016, Defendant moved to compel Plaintiffs' responses to Defendant's Superseding Interrogatory No. 1(Dkt. 29) ("Defendant's Motion") together with the Declaration of Teresa M. Abney, Trial Attorney, United States Department of Justice ("Defendant's Motion Exh. A"), Defendant's Brief in Support of Defendant's Motion (Exh. A) and Exhibits 1-13 In Support of Defendant's Motion ("Defendant's Motion Exh(s). __"). On October 3, 2016, Plaintiffs filed Plaintiffs' Memorandum In Opposition To Motion To Compel (Dkt. 31). Defendant filed, on October 11, 2016 its Reply To Plaintiffs' Opposition To Defendant's Motion (Dkt. 32) ("Defendant's Reply"). On October 12, 2016, Plaintiffs filed Plaintiffs' Motion For Case Management Order (Dkt. 33) ("Plaintiffs' Motion") together with Plaintiffs' Brief In Support of Plaintiffs' Motion attaching exhibits 1-10 ("Plaintiffs' Motion Exh(s). __"). On November 3, 2016, Defendant filed its Response In Opposition To Plaintiffs' Motion (Dkt. 35) ("Defendant's Response"). On November 10, 2016, Plaintiffs filed Plaintiffs' Reply To Defendant's Response Opposition To Plaintiffs' Motion (Dkt. 36) ("Plaintiffs'

Reply”). Oral argument on Defendant’s and Plaintiffs’ motions was conducted on November 30, 2016 (Dkt. 38) and decision reserved. On December 5, 2016, Defendant filed Defendant’s Supplemental Memorandum (Dkt. 39) (“Defendant’s Supplemental Memorandum”).¹ Thereafter, on December 14, 2016, Plaintiffs filed Plaintiffs’ Supplement To Plaintiffs’ Motion (Dkt. 40 (“Plaintiffs’ Supplement”); in accordance with the court’s request (Dkt. 41), Defendant’s Response to Plaintiffs’ Supplement was filed December 20, 2016 (Dkt. 42) (“Defendant’s Response to Plaintiffs’ Supplement”).

In Defendant’s Motion, Defendant contends that Defendant did not, contrary to Plaintiffs’ assertion, agree to proceed with discovery and trial in this case based on a limited sample of Plaintiffs’ projects, *i.e.*, 10 cases to be selected by each side (“self-selected sample”), or any other sampling procedure and that before agreeing to any sampling procedure it would be necessary for Defendant to assess, by evaluation of Plaintiffs’ answers to Defendant’s Superseding Interrogatory, whether the case was amenable to a particular form of sampling. Dkt. 29 at 10; Dkt. 35 at 9-11. Plaintiffs objected to Defendant’s Superseding Interrogatory asserting the Interrogatory, which requested Plaintiffs, *inter alia*, identify the business component as defined by § 41(d)(2)(B) for each project for which Plaintiffs claims an R&D tax credit, the nature of the uncertainty requirement under § 41, the particular science, *e.g.*, biology, chemistry, involved in the related QRE, the process of experimentation used in connection with Plaintiffs’ projects, how the results of such required experimentation would be useful in Plaintiffs’ development of a new or improved business product, process, etc., and the

¹ A further oral argument was conducted on January 26, 2017 (Dkt. 45) limited to the question of why the sampling plan devised by the I.R.S. to facilitate the audit of Plaintiffs’ claims, *see* Dkt. 33-5, would not be suitable to provide a workable sample of Plaintiffs’ projects as the basis upon which to litigate the merits of Plaintiffs’ claims.

expenses and related amounts for which the client was billed and the project was performed, was premature absent an agreement to use a sample of Plaintiffs' projects as the basis for discovery, rather than requiring Plaintiffs to access the records relevant to all 6,100 projects. Dkt. 29, Exh. 10 at 1. Plaintiffs thus refused to answer the Superseding Interrogatory based on excessive burdensomeness, harassment and lack of relevance to the randomly selected "judgment sample,"¹⁰⁻¹² of Plaintiffs' projects, and referred Defendant to the original list of 6,100 projects and information, including the Axiom study, Plaintiffs previously provided to Defendant in digital form, Dkt. 29, Exh. 10 at 3-6, and agreed to supplement such responses to provide documents following selection of an acceptable sampling technique. *Id.* at 6.² As to Defendant's request for identification of Plaintiffs' business component relevant to each claim, Plaintiffs referred Defendant to the study prepared by Axiom and Plaintiffs' responses to "IDR requests,"³ which Plaintiffs state had been provided to Defendant and that such components included "designs, process, and techniques" employed by Plaintiffs "for each project" without further description. *Id.* at 4-5. As to Defendant's request for an explanation of the degree of uncertainty as to how the relevant business component and final design would be developed by Plaintiffs as required for a QRE, Plaintiffs stated that such uncertainty existed "at the outset" of each project which was "overcome" by Plaintiffs' employees, and that further details responsive to the Superseding Interrogatory would be forthcoming upon selection of the anticipated "judgment sample" from among Plaintiffs' projects. *Id.* at 5. In response to Defendant's request for a description of the scientific principles involved in the QREs for each project, Plaintiffs asserted

² The record is unclear as to whether Plaintiffs had previously limited Plaintiffs' claims to 4,643 projects, see Defendant's Motion Exhs. 8, 9, 10.

³ Such IDR requests are not defined in the record.

Defendant's request was vague and ambiguous but that for all Plaintiffs' projects, Plaintiffs nevertheless employed "principals [sic] of engineering, science, mathematics, physics . . . chemistry and biology." *Id.* at 5. As to Defendant's request that Plaintiffs describe the process of experimentation required by § 41 for the QRE, Plaintiffs stated such process was comprised of "a systematic iterative process" whereby technical issues and desired outcomes and potential solutions and alternatives were identified, *i.e.*, by trial and error, a methodology to analyze and test such potential solutions and, where such potential solutions were proven to be unavailing, repeating this process until a viable solution was identified. Dkt. 29, Exh. 6 to Defendant's Motion at 9. Plaintiffs also responded by stating that once the appropriate "judgment sample" was selected, Plaintiffs would "provide more specific detail on the technical challenges" mentioned by Plaintiffs and "how a process of experimentation was employed to resolve the technical challenges." *Id.* Similar responses were provided by Plaintiffs to Defendant's other Superseding Interrogatory requests including information as to Plaintiffs' projects' clients, managers, source of funding and expenses. As can be seen from the foregoing summarization of Plaintiffs' responses, Plaintiffs' responses were highly generalized, essentially parroting the statutory elements required for an R&D tax credit without providing sufficiently specific information from which Defendant could determine the factual parameters upon which Plaintiffs' refund claim for any of the 6,100 (or 4,643) projects is based, despite Plaintiffs' acknowledgment that such responsive information is in fact available to Plaintiffs and that Plaintiffs would make it available to Defendant but only as to an extremely small sample, 10-12 projects, to selected by the parties at a later time. See Dkt. 29, Defendant's Motion Exh. 10 at 5 ("Once a judgment sample list

of projects has been selected for discovery and to be tried, the project documents for each project . . . will be made available [to Defendants].”⁴ (underlining added⁵))

Defendant rejects this limitation on Defendant’s access to the facts underlying Plaintiffs’ claim as unworkable given the absence of any facts upon which it could be inferred that such limited self-selected samples would be sufficiently representative of the entirety of Plaintiffs’ projects upon which Plaintiffs’ refund claim is based. Dkt. 29, Exh. B at 13 (“Using test cases in this litigation is only appropriate if it will produce a valid estimate of [Plaintiffs’] qualified research expenses.”); Dkt. 35 at 11 (suggesting that if Plaintiffs had been forthcoming in segmenting the 6,100 projects into groupings with common characteristics such responsive details could have facilitated a stipulated sampling procedure). Defendant also argues Plaintiffs’ burdensomeness objection should be rejected as Plaintiffs have failed to sufficiently particularize, including by presenting affidavits from persons with direct knowledge of Plaintiffs’ record keeping system, demonstrating such excessive burdensomeness. Dkt. 29, Exh. B at 13-14 (citing *Swinton v. Livingston County*, 2016 WL 1056608, at *6 (W.D.N.Y. Mar. 17, 2016) (quoting *In re Priceline.com, Inc. Sec. Litig.*, 233 F.R.D. 83, 85 (D.Conn. 2005))); *Burns v. Imagine Films Entm’t, Inc.*, 164 F.R.D. 589, 593 (W.D.N.Y. 1996) (objections to interrogatories required to be specific with detailed explanations supporting an objection). Plaintiffs’ burdensomeness objection is further undercut by the fact that Plaintiffs are required to maintain records sufficient to support Plaintiffs’ tax credit and refund claims, a requirement Plaintiffs do not dispute. See 26 C.F.R. § 1.41-2(d); see also *Bayer Corp. and Subsidiaries v. United States*, 850 F.Supp.2d 522, 539 (W.D.Pa.

⁴ Unless indicated otherwise bracketed material has been added.

⁵ Unless indicated otherwise underlining has been added.

2012) (“*Bayer*”) (rejecting use of plaintiffs’ proposed restrictive sampling plan without particularizing nature of plaintiffs’ business components related to plaintiffs’ QREs as basis for plaintiffs’ R&D tax refund claim of approximately \$50 million so as to not give “a reward to [plaintiffs] for failing to keep evidence regarding research expenses in ‘sufficiently’ useable form and detail.”). Given Plaintiffs are seeking a refund of approximately \$1.499 million, for similar reasons Plaintiffs’ burdensomeness objection fails on its face.

The force of Plaintiffs’ further contentions with regard to the availability of some form of sampling as typically utilized in R&D tax credit refund litigation in recognition of the prohibitively extensive amount of potential trial time that would be imposed on the court if all of Plaintiffs’ projects were to require trial, Plaintiffs’ Brief, Dkt. 33, at 5, is significantly diluted by Plaintiffs’ recent and unexpected unilateral decision to reduce to 159 the number of projects for which Plaintiffs seeks R&D tax credits and a resultant refund. See Plaintiffs’ Supplement (Dkt. 40) at 1 (“Plaintiff [*sic*] submits an amended list of projects which Plaintiff [*sic*] claims as part of the calculation of qualified research expenses.”); Defendant’s Response to Plaintiffs’ Supplement (Dkt. 42) at 1 (“This amended list contained only 159 projects, a substantial reduction from the 6,100 projects originally claimed.”).⁶ Despite such a dramatic reduction in the Plaintiffs’ projects at issue, Plaintiffs nevertheless reiterate their request that discovery and trial be limited to 10-12 sample projects to be selected from among the 159 projects now relied

⁶ Although in its Supplement Plaintiffs state “will not seek to claim qualified research expenses from projects not included in the list of 159 projects attached as ‘Exhibit A,’” Plaintiffs have not filed an amended complaint consistent with this representation. Plaintiffs are therefore directed to file, promptly, an amended complaint consistent with Plaintiffs’ Supplement particularizing the tax years to which the 159 projects relate as well as the amount of QRE credit and tax refund Plaintiffs now seek based on the 159 projects correctly asserting the status, as Plaintiffs and Defendant, of the respective parties.

on by Plaintiffs in support of its claims, to be selected either randomly by the court or by the selection of five to six projects by each party. Plaintiffs' Supplement (Dkt. 40) at 2.

Unsurprisingly, Defendant rejects Plaintiffs' proposal, Defendant's Response to Plaintiffs' Supplement (Dkt. 42) at 2-3 (court should order "discovery . . . on all 159 claimed projects" as the results of such discovery may support disposition of the matter on summary judgment or Plaintiffs' entire refund claim (and the case) may be rendered moot by the loss of approximately two-thirds of Plaintiffs' original claimed QREs based on the 6,100 projects initially asserted by Plaintiffs as supporting Plaintiffs' original refund claims. See Defendant's Response to Plaintiffs' Supplement at 2 (calculating that reducing to 159 the relevant projects upon which Plaintiffs' claims are now based lowers to approximately \$546 thousand Plaintiffs' claimed tax refund from approximately \$1.4 million, resulting from a reduction in Plaintiffs' QREs to approximately \$5.5 million from \$14.5 million, see Defendant's Response to Plaintiffs' Supplement (Dkt. 42) at 2, thereby potentially "eliminat[ing] Plaintiff's [*sic*] claimed entitlement to a refund").

Defendant further notes that despite Plaintiffs' 97% reduction in the number of claimed projects, Plaintiffs have failed to sufficiently support that providing discovery as to all 159 projects would nevertheless impose an excessive amount of burdensomeness upon Plaintiffs. *Id.* at 5-6. Defendant also questions whether Plaintiffs had any good faith basis for its refund claim based on the original assertion of 6,100 projects given that Plaintiffs have acknowledged it failed to do so. *Id.* at 6 (citing Plaintiffs' Memorandum (Dkt. 33) at 11 (admitting Plaintiffs' QRE allegations were based on employee "interviews," a "review" of Plaintiffs' financial records, and "estimations" of expenses attributable to the projects by the Axiom study) or provide sufficient responses

to Defendant's Superseding Interrogatory seeking the factual bases for Plaintiffs' claims, particularly Plaintiffs' QREs, as required by § 41. *Id.* at 6. The court therefore turns to the question of whether discovery in this case should proceed, at this stage of the litigation, on the basis of a sample of the 159 projects as Plaintiffs now request or whether discovery should be directed against all 159 cases as Defendant argues.

The use of statistical sampling techniques to resolve questions of liability and damages in federal court during discovery and for trial in cases involving numerous parties or large quantities of documents is now well-established. *See Hallmark v. Cohen & Slamowitz*, 304 F.R.D. 165, 169 (W.D.N.Y. 2015) ("The use of sampling in connection with discovery in class actions as a practical and cost-effective means by which to obtain relevant discovery is well-recognized." (citing *Hinterberger v. Catholic Health Sys.*, 284 F.R.D. 94, 103 (W.D.N.Y. 2012) (citing MANUAL FOR COMPLEX LITIGATION, Fourth, § 21.14 (2009) (acknowledging use of sampling methodology for discovery in class actions may be appropriate))); *Schwab v. Philip Morris USA, Inc.*, 449 F.Supp.2d 999, 1244 (E.D.N.Y. 2006), *rev'd sub nom on other grds.*, *McLaughlin v. American Tobacco Co.*, 522 F.3d 215 (2d Cir. 2008) (citing authorities and cases). "[W]hen it is necessary to limit discovery, 'statistical sampling techniques [may be used] to measure whether the results of the discovery fairly represent what unrestricted discovery would have been expected to produce.'" *Id.* (quoting MANUAL FOR COMPLEX LITIGATION (FOURTH) § 11.422 (bracketed material in original)). "The use of acceptable sampling techniques in lieu of discovery and presentation of voluminous data from the entire population, may produce substantial savings in time and expense." *Id.* (quoting MANUAL FOR COMPLEX LITIGATION (FOURTH) § 11.493). Such sampling has been found

useful in the litigation of R&D tax cases including the extent to which a taxpayer has incurred QREs meeting the criteria of § 41. See, e.g., *United States v. McFerrin*, 570 F.3d 672, 679 (5th Cir. 2009) (“If McFerrin can show activities were “qualified research,” then the court should estimate the expenses associated with those activities.” (citing *Cohan v. Comm’r of Internal Revenue Service*, 39 F.2d 540 (2d Cir. 1930))); see also *Suder v. Comm’r of Internal Revenue*, 108 T.C.M. (CCH) 354, 2014 WL 4920724, at *6 (2014) (basing court’s rejection of respondent’s deficiency and penalty determination for two of petitioners’ tax years upon petitioner’s alleged excessive R&D tax credit claims based on the parties’ stipulated selection of 12 of the 76 projects upon which petitioners relied for computation of the claimed QREs related to each project and upon which petitioners’ returns were prepared); *Union Carbide Corp. and Subsidiaries v. Commissioner of Internal Revenue*, T.C.M. (CCH) 50, 2009 WL 605161, at *2 (2009) (deficiency for improper R&D tax credits resolved based on parties’ agreement to try case on five of largest of 106 claimed projects); *contra Bayer Corporation and Subsidiaries*, 850 F.Supp.2d at 546 (rejecting use, in R&D tax refund case, of plaintiffs’ proposed sampling technique where, in response to defendant’s interrogatory, plaintiff failed to particularize nature of plaintiff’s business components underlying plaintiff’s alleged QRE credits and citing cases).⁷ Despite the use of limited self-selected sample projects in Tax Court tax deficiency litigation, as Plaintiffs have proposed for this case, courts are reluctant to approve the use of such self-selected exemplars that could

⁷ The court notes that the I.R.S. formulated a statistical sample plan using, in part, random sampling as the “most practical and efficient examination method,” Dkt. 33-5 at 3, to conduct an audit of Plaintiffs’ R&D tax credit claims. See Dkt. 33-5. At the further oral argument conducted January 26, 2017, Defendant stated that because the I.R.S. sampling plan was based on groupings of Plaintiffs’ projects based on the amount of claimed QREs and not on the various types of remediation projects such as water and ground contamination represented in Plaintiffs’ QREs, it was unacceptable to Defendant; Plaintiffs indicated the I.R.S. sampling plan was not preferable but acceptable.

create the possibility that such self-selection will not result in a sample representative of all claims at issue in the case. See *In re Chevron U.S.A.*, 109 F.3d 1016, 1020-21 (5th Cir. 1997) (disapproving use of 30 of plaintiffs' 3,000 cases, self-selected by the parties, to extrapolate defendant's toxic tort liability for the class as this procedure did not support that such limited sample would be representative of all of plaintiffs' claims in the case and that such failure to assure such representativeness also violated due process). "Typically, a finding [of liability] must be based on competent, scientific, statistical evidence that identifies the variables involved and that provides a sample of sufficient size so as to permit a finding that there is a sufficient level of confidence that the results obtained reflect results that would be obtained from trials of the whole." *Id.* at 1020.

Here, similar to the dispute in *Bayer*, Defendant refuses to agree to any form of sampling even of the reduced number (159) of Plaintiffs' projects drawn from the original 6,100 projects which Plaintiffs now claim form the basis of Plaintiffs' refund claim without benefit of some indication that Plaintiffs can or cannot establish there is any factual basis for Plaintiffs' claims, for example, the exact nature of Plaintiffs' asserted business components, new products, techniques, etc., to which any of Plaintiffs' alleged QREs relate as required by § 41. To permit Plaintiffs to proceed based on a self-selected sample of 10-12 cases as Plaintiffs propose, Dkt. 40, Plaintiffs' Supplement, at 2, to be randomly selected by the court or self-selected by the parties, will not allow Defendant to understand which, if any, of the remaining cases may or may not meet the requirements of § 41, and would require Defendant to litigate the merits of the case based on a sample which, depending on the actual characteristics of the 148-

150 remaining cases with respect to QRE requirements, may well be too small to be fairly representative of the population of 159 projects, albeit a significantly reduced number. In particular, “[a] sample is said to be *representative* with respect to a variable [or relevant characteristic of the population to be sampled] if its relative distribution in the sample is equal to its relative distribution in the population.” Jake Bethlehem, *APPLIED SURVEY METHODS – A STATISTICAL PERSPECTIVE*, Wiley (2009) (italics in original) at 24. As relevant, it is also accepted that “[b]igger samples give estimates that are more precise,” and where the population to be “sampled is heterogeneous, random error⁸ will be large; a larger sample will be needed to offset the heterogeneity.” *REFERENCE MANUAL ON SCIENTIFIC EVIDENCE (THIRD ED.)*, Federal Judicial Center, at 246. Moreover, “[a] pilot sample may be useful to estimate heterogeneity and determine the final sample size.” *Id.*

In the instant case, given that there are at least eight relevant, *i.e.*, required, factors for any valid QRE in accordance with § 41, it is likely that among the population of 159 projects, a significant number will not meet all applicable requirements thereby rendering the population sufficiently heterogeneous with regard to the relevant variables, *i.e.*, § 41 requirements, and requiring a larger sample in order to reasonably assure the resultant sample is reliably representative. More specifically, the apparent flaw in Plaintiffs’ self-sampling proposal is that, absent some indication as to the QRE variations among the 159 projects put forth by Plaintiffs, it is not possible to determine whether the population of 159 projects is amenable to the use of any sampling technique for purposes of discovery and trial, including use of a random sample. Thus,

⁸ The calculated “impact of random chance on study results.” *REFERENCE MANUAL ON SCIENTIFIC EVIDENCE* at 240.

in the absence of an affidavit from one qualified to advise the court on a competent sampling technique for this case, the court is required to make its own judgment to assure any sample is reliably representative. See *Hallmark v. Cohen & Slamowitz*, 307 F.R.D. 102, 106 (W.D.N.Y. Jan. 8, 2015), *aff'd sub nom, Hallmark v. Cohen & Slamowitz, LLP*, 2015 WL 4483959 (W.D.N.Y. July 22, 2015) (“*Hallmark*”) (failure by defendant to proffer technical support for defendant’s sampling proposal leads court to impose 10% sample of disputed documents, standard credit card contract terms and conditions to discern clauses supporting defendant’s contractual authorization defense).

In contrast to *Hallmark*, which involved a population of more than 1,000 documents – standard preprinted credit card form contracts – with a high degree of likely homogeneity, here the exact nature and details of the 159 projects, *i.e.*, the types of asserted remediation involved, is presently unknown as Exhibit A to Plaintiffs’ Supplement indicates only that all 159 projects involved remediation of various contaminations resulting in a “new process of remediation” not otherwise specified, alleged to comprise the requisite business component, along with identification of the project manager and project location. Nor is this representation by Plaintiffs submitted under oath. Significantly, Plaintiffs’ labeling of the projects as involving remediation provides no indication as to whether the other prerequisites for a QRE, such as the nature of experimentation, perceived uncertainties, why the process is “new,” compared to Plaintiffs’ previous “processes,” and applied scientific principles, are present in any of the 159 projects. As a safeguard against possible reliance on an unrepresentative sample it is therefore necessary to enlarge the size of the random sample of the 159 projects, as a “pilot” sample, to provide preliminary insights as to whether discovery

(and possible trial) of Plaintiffs' claims can be confidently predicated based on any sample of the 159 projects, or whether full discovery and trial must be conducted as to all 159 projects in order to provide "assurance" to the court against awarding any relief to Plaintiffs amounting to "unguided largesse." *Bayer Corp. and Subsidiaries*, 850 F.Supp.2d at 542 n. 40 (quoting *Williams v. United States*, 245 F.2d 559, 560 (5th Cir. 1957)).

Thus, in the instant case, to determine whether the case is amenable to use of a sampling technique such as random⁹ sampling, either by stipulation or one imposed by the court, the parties shall within 10 days of this Decision and Order randomly select 40 projects (approximately 25%) of the Plaintiffs' 159 projects as a pilot sample. The parties shall, prior to taking the sample, establish a protocol which assures that all project files to be sampled are marked with a permanent identification number and indexed by project name and identification number so as to prevent any possible manipulation of the sample project files after the sample is completed by substituting files preferred by Plaintiffs. Defendant shall have reasonable access to Plaintiffs' project files to assure the integrity of the sampling procedure. The Plaintiffs shall then serve answers to Defendant's Superseding Interrogatory for each project identified by the random sample within 30 days thereafter. Defendant's review of Plaintiffs' answers shall be completed within 30 days of the service of Plaintiffs' answers; Defendant shall advise Plaintiffs in writing when Defendant's review is complete. The parties, within 10 days after notice of completion of Defendant's review of Plaintiffs' responses and after evaluating the results of the pilot sample, shall jointly as a proposed stipulation or

⁹ "A random sampling is one in which every item in the population is equally likely to be chosen." Alan Graham, *STATISTICS – A COMPLETE INTRODUCTION*, Hodder & Stroughton (2013) at 129.

individually advise the court whether some, if any, form of sampling upon which to proceed with this case, including discovery and further proceedings, acceptable to the parties, together with proposed scheduling orders governing such further proceedings. The parties may also request a conference with the court to discuss the scheduling of further proceedings following completion and review of the results of the pilot sample and Plaintiffs' answers to the Superseding Interrogatory.

CONCLUSION

Based on the foregoing, Defendant's motion (Dkt. 29) is GRANTED in part and DENIED in part; Plaintiffs' motion (Dkt. 33) is DENIED.

SO ORDERED.

/s/ Leslie G. Foschio

LESLIE G. FOSCHIO
UNITED STATES MAGISTRATE JUDGE

Dated: January 26, 2017
Buffalo, New York