Appendix C

BRACH | EICHLER LLC

Riza I. Dagli
Direct Dial: 973-403-3103
Direct Fax: 973-618-5503
E-mail: rdagli@bracheichler.com

July 2, 2021

VIA EMAIL

[redacted]
Audit Supervisor
Office of the State Comptroller
Medicaid Fraud Division
PO Box 025
Trenton,, NJ 08625-0025

Re: Heart to Heart Health Care Services, LLC, d/b/a Heart to Heart Home Care

Dear Mr. [redacted]:

Our firm represents Heart to Heart Health Care Services, LLC (HTH). We are in receipt of the Medicaid Fraud Division's (MFD) Draft Audit Report letter dated June 15, 2021, wherein your agency alleges certain improper payments made to HTH, and which also offers an extrapolation methodology as a basis for a calculation of an overpayment. This letter is in response to that letter. All rights are reserved.

As you are aware, HTH provides medical and personal care services to the New Jersey community, primarily serving the complex needs of New Jersey's vulnerable Medicaid population. Since its inception, Heart to Heart has devoted itself to the highest standards of professionalism and customer care and satisfaction. The notion that HTH erroneously billed Medicaid to the staggering extent suggested by the Draft Audit Report is troubling, but more importantly, unbelievable and contravened by HTH's business practices and its own review of its records. Certainly, errors can happen in any business, and, considering the large amount of transactions by HTH, it is possible that some billing mistakes were made, and HTH will reimburse Medicaid for any billing mistakes, including the ones identified in the Draft Audit Report.

However, as a preliminary point, prior to a discussion of the specific allegations and error types identified in the Draft Audit Report, HTH objects to the method by which extrapolation was used in this instance. Extrapolation is a useful tool when applied correctly to identify trends and patterns in large volumes of data. However, for extrapolation to be effective, the representative sample must represent the larger pool of data. In other words, the representative sample must

contain all the attributes of the population in the same proportion that they exist in the population. Only then can the representative sample be used to generalize from merely a sample to the larger population. For the reasons discussed below, the representative sample chosen in the Draft Audit Report does not represent the larger population.

Poor Degree of Precision

Estimation methodologies using statistical sampling require analysts to weigh the estimate's uncertainty to determine whether the conclusions are useful for their desired purpose.[1] Several measures are useful when evaluating a study's uncertainty. *Precision* reflects the range of accuracy related to an estimated amount, while *confidence* is the degree of certainty that the sample correctly depicts the population. Together, confidence and precision yield the *confidence interval*, a range of values within which the true population value is estimated to fall.

In healthcare overpayment matters, precision levels from 5 to 10 percent are generally sought. However, the precision of MFD's analysis in this matter is significantly worse: 35 percent.[2] In addition to its overall precision, MFD also achieved extremely poor precision in each and every stratum. This is particularly problematic considering MFD's own stated objectives for achieving high precision in its sampling plan:

> *Used the 95% confidence, 5% precision (95/5) level or better in selecting sample sizes based on the examined values. Note: Selecting sample sizes at the 95/5 level does not guarantee each strata will achieve 5% precision. However, it does ensure the overall sample precision will be approximately 5% when estimating the total dollars in the universe.*[3]

Instead of achieving its own goal, the actual precision of MFD's analysis in this case was dramatically higher than 5%, yielding distinctly imprecise conclusions. This imprecision is highlighted by MFD's extremely large confidence interval (i.e., estimated range of overpayments) ranging from $1.5 to $3.3 million:

> *Using extrapolation, MFD can reasonably assert, with 90% confidence, that the true overpayment in the universe falls between $1,506,618 and $3,261,647 with the most likely overpayment amount (i.e. error point estimate) as $2,384,132.55.*[4]

In contrast to MFD's precision in this matter, most healthcare post-payment audits seek significantly lower (i.e., better) precision levels ranging from 5 to 10 percent, and RAT-STATS software (which MFD purportedly used) prepopulates with desired precision levels from 1 to 15

[1] United States, Internal Revenue Service, Bulletin 2007-23, Sampling Plan Standards, 2007.
[2] MFD Spreadsheet, Data Provider Copy.xlsx, Recovery Summary tab.
[3] MFD Spreadsheet, Data Provider Copy.xlsx, Sampling Plan tab. Emphasis added.
[4] MFD Spreadsheet, Data Provider Copy.xlsx, Recovery Summary tab. Emphasis added.

percent. Even guidance for OIG Corporate Integrity Agreements prescribes a maximum precision level of 25 percent.[5] The poor degree of precision in this case indicates a lack of technical rigor applied by MFD and a high degree of variability in MFD's analysis. It also indicates the inadequacy of the sample size chosen by MFD in this matter, since increasing sample size is generally the most effective technique for improving precision.

Improper Use of Point-Estimate

In reaching its conclusions regarding Heart to Heart's extrapolated overpayment amount, MFD based its overpayment demand on the point-estimate, stating the following:

> *Using extrapolation, MFD can reasonably assert, with 90% confidence, that the true overpayment in the universe falls between $1,506,618 and $3,261,647 with the most likely overpayment amount (i.e. error point estimate) as $2,384,132.55.*[6]

Here, MFD incorrectly contends that the point estimate is the most likely amount of overpayment. This characterization is untrue, and it suggests a limited understanding of probability theory. Selecting the point-estimate (or any value in a confidence interval) is not a probabilistic statement, and no value that lies within the confidence interval is *more likely* than another to be the *true* overpayment value. The point-estimate is simply the convenient midpoint of the confidence interval and is therefore anticipated to over-assess the disallowance almost half of the time. This distinction becomes more significant as the level of imprecision in a particular analysis grows, since the confidence interval grows wider with increased imprecision and over-assessments may be even greater.

In cases of poor precision such as this, the point-estimate is not the preferred estimate. Instead, the lower-limit of the 90 percent confidence interval is preferred in cases where adequate precision is not achieved. For example, CMS prefers the use of the lower-limit "in most cases" in post-payment audits since it "allows a reasonable recovery without requiring the tight precision that might be needed to support a demand for the point-estimate."[7] Similarly, the OIG's Statistical Sampling Toolkit for MFCUs states "When the precision is poor, the uncertainty in the sample can often be managed through the use of alternate estimates such as the lower limit of a confidence interval."[8]

[5] U.S. Department of Health and Human Services, Office of the Inspector General, Corporate Integrity Agreement FAQs, CIA Claim Reviews. Available at https://bit.ly/2MertiD

[6] MFD June 15, 2021, Draft Audit Report.

[7] Medicare Program Integrity Manual, 8.4.5.1.

[8] U.S. Department of Health and Human Services, Office of the Inspector General, Statistical Sampling: A Toolkit for MFCUs, September 2018.

In this matter, the lower-limit of the 90% confidence interval is $1,506,618 using MFDs own calculations and without considering any of Heart to Heart's other arguments.[9]

Lack of Scientific Rigor in Sample Size Determination**:**

MFD's sample size of 118 claims was determined without sufficient scientific rigor and RAT-STATS, a statistical sampling software, was used improperly leading to a non-representative sample selection and insufficient levels of statistical precision. In accordance with the CMS' MPIM, one of the "major" steps of statistical sampling involves "Performing the appropriate assessment(s) to determine whether the sample size is appropriate for the statistical analyses used."[10]

Despite the well-known risk of selecting a sample size that is too small to achieve valid results, MFD adopted a sample size of only 118 claims to estimate overpayments for a population totaling 296,374 claims (i.e., a sample of less than 0.040 percent). Had they carefully considered an appropriate sample size; they would have concluded that a much larger sample would be required to reach sufficiently precise conclusions in this matter.

MFD's stated reason for choosing a sample size of 118 was reliance on RAT-STATS and its stratified sample size calculation module. MFD demonstrated that the calculation of sample size is determined using three variables: (1) desired confidence, (2) desired precision, and (3) standard deviation (i.e., variance). However, in its own analysis, MFD misapplied these variables. MFD calculated sample size using the standard deviation of the irrelevant claim payment amounts, as opposed to the more-appropriate standard deviation of the overpayment amounts (i.e., the actual variable of interest). MFD had relevant overpayment data from its probe sample, however seemingly failed to consider their own analysis and instead relied on less relevant payment data.

In fact, MFD's probe sample provides meaningful data that should have been used in MFD's sample size calculations. Failing to do so ignores a basic purpose of probe samples – collecting initial data to make better-informed decisions about the sample design (including sample size). OIG and CMS specifically address the role of probe samples in developing sampling analysis and determining sample size. Also, RAT-STATS' Unrestricted Sample Size module, which MFD failed to use, specifically utilizes the probe sample when determining sample size.

Notwithstanding MFD's failure to consider its own probe sample, Heart to Heart recalculated an appropriate sample size by properly using RAT-STATS. Using MFD's own determinations for its probe sample of 46 claims, its own stratification criteria, and its own stated criteria for determining sample size (i.e., 95% confidence and 5% precision) **an appropriate stratified sample size would require a random selection of over 8,839 claims** (i.e.,

---

[9] MFD Spreadsheet, Data Provider Copy.xlsx, Recovery Summary tab.
[10] Medicare Program Integrity Manual, 8.4.1.3 (5).

approximately 3 percent of the total universe).[11] Even when using the most aggressive values of confidence and precision available in RAT-STATS (i.e., 80% confidence and 15% precision) the calculated sample size for Heart to Heart's universe would be a **minimum sample size of 435 claims**.[12] Had MFD chosen an adequately sized sample, many of the issues described in this document (i.e. representativeness, precision, etc.) would have likely been avoided.

Lack of Sample Representativeness

This dramatic difference in sample size is not merely a theoretical issue. In a universe with high variability (i.e., heterogeneity) small samples risk failing to adequately capture subsets or characteristics of the universe, thereby misrepresenting an extrapolated estimate. In fact, that is precisely what occurred in this case. Even if MFD's limited sample size was determined to be technically sound, the sample of claims that was actually selected is not adequately representative of the universe from which it was chosen. Since characteristics of a sample will be used to infer characteristics of the broader population, a sample must be reasonably representative of the population to permit a valid extrapolation. If the sample chosen is not representative of the population, inferences about the population may be irreparably biased and invalid. Although selecting a sample randomly is anticipated to lead to a representative sample, it is not guaranteed, particularly when small samples are selected (such as this case).

Nonetheless, MFD provided no evidence that it adequately addressed the representativeness of its own sample in this matter. More importantly, a diligent review of MFD's chosen sample instead suggests it is not representative of the population of claims at issue, and therefore insufficient for the purposes of making inferences (i.e., extrapolation) about the distinctly heterogeneous population. Had MFD properly selected a larger sample, it likely would have selected and examined many more of these ignored claims leading to a more representative and reliable sample.

Extrapolation is Likely Impermissible**:**

In its audit letter, MFD evaluated a sample of 118 claims and identified a Claim Error Rate (i.e., the percentage of claims with any measurable deficiency) to be 16.1 percent. More meaningfully, MFD identified a Net Financial Error Rate in the sample (i.e., the percentage of payment amounts found in error) to be 12.95 percent.[13] Even if Heart to Heart's arguments disputing these identified errors were ignored, MFD's calculated error rates are insufficient to allow extrapolation in similar matters.

---

[11] Heart to Heart Re-Calculation of Stratified Sample Size with RAT-STATS.pdf.
[12] Heart to Heart Re-Calculation of Stratified Sample Size with RAT-STATS.pdf.
[13] MFD Draft Audit Report, dated June 15, 2021, page 1.

Specifically, CMS authorities have ruled that error rates must exceed 50% in order to permit extrapolation, and extrapolations based on smaller error rates have been excluded in CMS administrative hearings citing "the Provider error rate is below the threshold of 50% required to justify extrapolation."[14] In fact, CMS states in its Medicare Program Integrity Manual ("MPIM") guidance on statistical sampling that "For purposes of extrapolation, a sustained or high level of payment error shall be determined to exist through a variety of means, including, but not limited to: high error rate determinations by the contractor or by other medical reviews (i.e., greater than or equal to 50 percent from a previous pre- or post-payment review)."[15]

In this matter, MFD has presented no evidence that Heart to Heart's error rate was sustained over any period of time, and based upon similar CMS decisions, Heart to Heart's error rate is also not "high" as contemplated by CMS. Consequently, extrapolation is likely impermissible for the purpose to estimating overpayments in this matter.

Error Theories:

A. HTH failed to verify the professional certification of an HHA.

In the sample of 118, MFD found that one instance where an HHA did not have a current certification, because the HHA's temporary certification had expired. By extrapolation, MFD is stating that 1 out of 118 claims, or 0.85% of all claims by HTH involved an expired temporary certification. Said another way, MFD is claiming there were 2,512 instances within the pool of 296,374 claims where the claim was invalid due to an HHA's temporary certification being expired. It is doubtful that HTH employed any individuals with an expired temporary license, other than this particular HHA, so it is not appropriate to expand this single unusual occurrence into a pattern or trend.

B. HTH Failed to perform timely in-home evaluations.

MFD alleged 14 instances where a claim fell outside a range within 60-days of an in-home evaluation. MFD is alleging violations of *N.J.A.C.* 10:60-3.5(a)(2) and *N.J.A.C.* 13:45B-14.9(g), which require periodic in-home evaluations to be sure the POC (plan of care) is correct. MFD is not alleging that the POC should have been revised in those 14 instances, or that the care provided to the patient was deficient in some way, or that at in-home evaluation would have resulted in different services or a different POC. In those 14 cases, the patients in question received the same care before and after the 60-day in-home evaluation was performed, albeit later than required. As MFD is aware, in-home evaluations are sometimes scheduled but canceled by the patient or the patient's guardian, thereby making 60-day visits not as timely as would be preferred. But it is important to note that tardiness of the 60-day evaluations did not change the POC. Appropriate

---

[14] QIC redetermination decision, dated June 1, 2017.

[15] Medicare Program Integrity Manual, 8.4.1.4.

services were provided, and the services were properly paid. However, assuming the claims should not have been paid because they fell outside the 60-day period, it is not appropriate to suggest that 11.9% (14/118) of all 296,374 claims fall outside the 60-day period. A more accurate methodology would have been to determine how many claims in the sample of 118 fall outside the 60-day period (in this case 14), and of those claims, how many days were not in compliance for each patient. For the 14 instance identified by MFD, HTH looked at how many days were billed before a subsequent 60-day evaluation occurred. The results are below:

[redacted]	114
[redacted]	188
[redacted]	58
[redacted]	196
[redacted]	178
[redacted]	64
[redacted]	26
[redacted]	10
[redacted]	186
[redacted]	76
[redacted]	42
[redacted]	245
[redacted]	3
[redacted]	55

The total of days in the above chart is 1,441. As noted by MFD, the remaining 104 claims reviewed by MFD did not fall outside the 60-day window. Roughly speaking, MFD's sample of 118 claims represents 118 patients, or 43,070 days of service in one year (365 x 118). Within these there were 1,441 days outside the 60-day period. The ratio of 1,441/43,070 is 3.3%. In other words, only 3.3% of the total days of service related to the representative sample in one year fell outside the 60-day evaluation period. Moreover, at the time of the audit, the total number of days outside of the 60-day period was even less, only 564 days, resulting in an extrapolated rate of 1.3%. Although HTH is not conceding that extrapolation is appropriat, HTH would offer either 1.3% or 3.3% as more accurate than the methodology used by MFD. MFD's methodology erroneously fails to consider the number of days outside compliance, but rather focuses on instances of non-compliance, which give a skewed number.

To further illustrate the inaccuracies of MFD's extrapolation in connection with the 60-day evaluations, we looked closer at the 14 instances identified by MFD. In the Draft Audit Report, MFD claims that 11.9% of all claims are outside the 60-day evaluation period. However, even the 14 instances identified by OSC were only 564 days out of compliance at the time of audit. Out of

5,114 possible days for the 14 patients (365 x 14), 564 were not in compliance, resulting in a percentage of 11.0%. In other words, even ignoring the 104 *good* claims, the 14 *bad* claims themselves have a lower error rate (11.0%) than the extrapolated rate (11.9%)! Clearly, this demonstrates that MFD's extrapolation methodology is flawed.

C. HTH Billed for Unsubstantiated Service

In the sample of 118, MFD found three (3) instances where HTH could not find a time-sheet. There is no allegation that the claim is otherwise improper or that the service was not performed. HTH believes its claims are valid, even if a timesheet was not located at the time of audit. Three (3) is too few to form the basis of a trend or pattern from which an extrapolation can be made.

D. HTH Failed to Prepare a POC Prior to Initiating Service

In the sample of 118, MFD found that an instance where HTH failed to prepare a POC prior to initiating service. By extrapolation, MFD is stating that 1 out of 118 claims, or 0.85% of all claims by HTH involved a failure to prepare a POC prior to service. Said another way, MFD is claiming there were 2,512 instances within the pool of 296,374 claims where a POC was not prepared prior to service. HTH would respectfully submit that a single occurrence cannot form the basis of a trend or pattern from which an extrapolation can be made.

E. HTH Improperly Billed PCS while Beneficiaries Were Inpatient in a Hospital

HTH reserves its rights in connection with challenging this allegation upon additional investigation. HTH does not presently concede that the applicable hospital records, as opposed to records of HTH, are more reliable.

CONCLUSION

HTH disputes the findings in the Draft Audit Report. Nonetheless, HTH recognizes that billing mistakes can, and do, occur. As always, HTH is willing to continue its discussions with MFD to achieve a resolution of MFD's concerns and resolve this matter.

Very truly yours,

Riza I. Dagli

RID:amp