UNITED STATES DISTRICT COURT
FOR THE EASTERN DISTRICT OF PENNSYLVANIA

IN RE: NATIONAL FOOTBALL LEAGUE PLAYERS' CONCUSSION INJURY LITIGATION	:	No. 2:12-md-02323-AB MDL No. 2323 **Hon. Anita B. Brody**
THIS DOCUMENT RELATES TO: APPEAL OF SETTLEMENT CLASS MEMBER [redacted] REGARDING DENIAL OF MONETARY AWARD	:	

INTRODUCTION

Mr. [redacted], a Retired NFL Football Player and Class Member under the Amended Class Action Settlement Agreement, appeals a denial of a Pre-Effective Date Diagnosis of Level 2 Neurocognitive Impairment. The procedural history of this claim is complex, including an extensive audit as well as three separate Appeals Advisory Panel reviews.

Notwithstanding this lengthy process, the record remains incomplete. As a result, I grant in part Mr. [redacted] Appeal, and remand to the Claims Administrator to obtain better documentation.

FACTUAL AND PROCEDURAL BACKGROUND

Mr. [redacted] first filed for benefits under the Settlement Agreement on April 26, 2017, based on a Pre-Effective Date Diagnosis of Level 2 Neurocognitive Impairment. Doc. 63132. On January 5, 2018, the Claims Administrator denied his Claim, along with 152 others, after an audit. That audit, and a subsequent ruling from the Special Masters, concluded that the neuropsychologist who performed the testing relied upon by the diagnosing neurologist failed to "meet the standard of care required for a Monetary Award." Doc. 151267. In adopting that decision, the Special Masters permitted affected Players to seek new evaluations through either the Baseline Assessment Program or from a Qualified MAF Physician. *Id.*

On March 20, 2018, Mr. [redacted] submitted a second Claim—this one—for Level 2 Neurocognitive Impairment. Doc. 162746. MAF neurologist Dr [redacted] certified the Diagnosis, with Dr. [redacted] providing a supporting neuropsychological evaluation. Doc.

162294. Dr. ██████ while concluding that Mr. ██████ performance validity level ought to be "considered optimal," still opined that:

> Validity scales indicated possible overreporting of psychiatric, somatic, and cognitive concerns, though his elevated report of these problems is consistent with his lengthy, documented history of distress in these areas. His level of endorsement on these validity scales, however, was not consistently high enough to definitively invalidate his clinical profile. Nonetheless, some caution is warranted in the interpretation of these scales.

Doc. 162285, at 13-14.

The Claims Administrator initially approved this second Claim. Doc. 163611. The NFL Parties then timely appealed, arguing that Mr. ██████ public affairs and activities were not consistent with a CDR score of 2, and that his functional decline was a byproduct of the physical and mental health issues that he experienced. Doc. 168864. On August 2, 2018, Special Master Pritchett denied the NFL Parties' Appeal. Doc. 181331. On September 10, 2018, the Claim was placed under audit, which ultimately yielded no adverse findings. Doc. 184718; Doc. 198203.

However, during this post-Appeal review process, the Claim was one of seventeen listed in the NFL Parties' request to stay particular claim payments until they were reviewed by the AAP. Doc. 189690. Special Master Pritchett denied the request for a stay. Doc. 189697. In upholding the Special Master's discretionary use of AAP Review, the District Court determined "that the claims subject to this appeal which are based on Qualifying Diagnoses by MAF Physicians would benefit from application of the revised Rules Governing Qualified MAF Physicians, adopted on April 11, 2019."[1]

Acting pursuant to the Court's order, an AAP Reviewer evaluated Mr. ██████ Claim and the accompanying medical evidence. This review, the first by the AAP for Mr. ██████ second-filed Claim, concluded that there was insufficient documentation to support Mr. ██████ Diagnosis of Level 2 Neurocognitive Impairment. Doc. 210482. Particularly, the Review emphasized what it saw as discrepancies between Mr. ██████ neurocognitive scores and the effects of physical pain and medication, as well as between his scores and observed behavior. *Id.*

> Dr. ████ records an MMSE score of 25/30 (mild impairment) and reports that the player is oriented with intact speech and language. Similarly, Dr. ██████ reports "rhythm and prosody of speech were within normal limits and conversational language was intact." At the same visit, the player achieves Level 2.0 impairment thresholds for language, but Dr. ██████ does not address these discrepancies in his rating of the Slick Criteria. The player has exceptionally low T-scores across multiple domains to the point of questionable validity and in the context of past

---

[1] Order Pursuant to Settlement Implementation Determination, In Re: National Football League Players' Concussion Injury Litigation, No. 2:12-md02323-AB (Apr. 12, 2019).

> conclusions of malingering, the submission of raw scores from the performance validity evaluation will be critical to any determination. Dr. ██████ identifies challenges to validity of symptom reporting on the MMPI, but then dismisses them. As with the performance validity testing, the absence of the actual scale scores on the MMPI make the validity of his interpretation questionable.

Id. The Reviewer also explained that "Dr. ████ does not document consideration of pain, fatigue, tinnitus, headache, chronic effects of depression and anxiety, or chronic opiate use at unknown doses in his formulation of the diagnosis." *Id.*

Relying on this analysis, the Claims Administrator denied the Claim on August 27, 2019. Doc. 213385. On October 13, 2019, Mr. █████ appealed. Doc. 215195.

On October 21, 2019, the Special Master remanded the Claim for re-review. Doc. 215943. An AAP Reviewer provided two additional analyses on January 28, 2020 and May 9, 2020, bookending Dr. ███████ April 22, 2020 submission of Mr. ██████ validity scores. Doc. 223709.

In the first, the Reviewer noted a number of concerns, including the impact of psychiatric disorders, insomnia, and chronic pain on Mr. ██████ neuropsychological assessment scores. Doc. 220222. The focus of that review was what the Reviewer saw as fatal flaws in Dr. █████ validity analysis:

> The neuropsychological report and other documentation contain multiple discrepancies raising concern about validity in terms of the Slick Criteria.
>
> Slick validity criteria are used in the Settlement to determine whether a retired player's test data is a valid reflection of his optimal level of neurocognitive functioning. This player shows a discrepancy between test data and observed behavior for the cognitive domain of language. This situation involves performance on two or more neuropsychological tests within a domain that are discrepant with observed level of cognitive function in a way that suggests an exaggeration or fabrication of dysfunction.
>
> This player is well educated and presents with no significant language disturbance in conversational speech during the clinical interview portion of the neuropsychological evaluation, but performs in the severely impaired range on formal tests of verbal fluency, auditory comprehension, and confrontation naming. In fact, despite severely impaired scores on all 3 tests within the language domain (BNT T-Score 30, Category Fluency T-Score 21, CIM T-Score 12), during the clinical interview Dr. █████ states, "Rhythm and prosody of speech were within normal limits and language was conversationally intact." Furthermore, Dr. ████ described the player as fully oriented, with intact speech and language. Dr. ████

> reports a MMSE score of 25/30; indicating only mild impairment, and not consistent with Level2.0 [sic] Neurocognitive Impairment.
>
> Furthermore, there are *discrepancies between test data and reliable collateral reports*. Discrepancies include the report of Dr. ██████ that the player continues to drive. Similarly, the Affidavit of ██████ describes the player as participating in "daily financial decisions and vetting certain investment proposals." The reported daily activities are inconsistent with Level 2.0 Neurocognitive Impairment.

Id. (internal citations omitted) (paragraph breaks added).

On re-reconsideration, the Reviewer maintained there was insufficient evidence to support a Level 2 Neurocognitive Impairment Claim. The Reviewer noted "glaring inconsistencies and discrepancies for Slick validity criteria," and especially "multiple discrepancies between test data and observed behavior." Doc. 224270. The Reviewer concluded that "[t]he additional tests [sic] scores for validity measures provided by Dr. ██████ actually increase concern for overall validity." *Id.* Regarding Mr. ██████ daily functioning, the Reviewer underlined that "[m]ultiple non-cognitive factors significantly limit the player's everyday functioning and were not adequate [sic] considered by Dr. ██████ which include chronic pain, fatigue, severe headaches, chronic effects of anxiety and depression, and chronic opiate use." *Id.* Finally, the Reviewer found that Mr. ██████ had failed multiple MMPI-2-RF validity scales, and that a score for Trail Making Test Part A was omitted. *Id.*

Relying on this revised analysis, the Claims Administrator again denied Mr. ██████ Claim. Doc. 224478. Mr. ██████ timely filed this Appeal. Doc. 225840.

DISCUSSION

This Opinion results from a review process stretching over two and a half years. Mr. ██████ Claim—his Second—has been granted, audited, appealed to the District Court, and twice remanded and denied. Various medical experts have disagreed whether he merits a Diagnosis of Level 2 Neurocognitive Impairment. His lawyers, the NFL Parties, and even Class Counsel have filed several rounds of expensive briefing. In short, something seems amiss.

A major bottleneck concerns whether Mr. ██████ test scores are valid measures of his neurocognitive impairment. His neurologist, later disqualified from the program, determined that the scores were valid. His treating neuropsychologist, a certified clinician approved to assist Qualified MAF Physicians, still maintains that opinion. But several AAP Reviewers, and now the Claims Administrator, have come to a different conclusion. Mr. ██████ case is not unique: multiple recent appeals have generated heated validity disputes.

It's thus worth starting with what a *good* process would have looked like. As the Special Masters have repeatedly emphasized, the Program's appeals system is "not intended to be an inquisitorial one, designed to nitpick the BAP and MAF physicians' clinical judgment. Rather, it

is built around layered deference to expertise."[2] The Special Masters defer to the decision of the Claims Administrator unless it is "clearly erroneous," while the Claims Administrator generally accepts the reasoned opinions of its pre-screened, expert, MAF and BAP Physicians.

This system rests on front-line clinicians following the Settlement's procedures: only by articulating their judgments with care and fastidiousness can the program be assured that claims are being paid appropriately. In the absence of well-articulated medical judgments, we get delay, excessive cost and a dreadful ratio of law to medicine. That is the story of this Appeal.

Respecting validity, clinicians' "articulated medical judgment" starts with the Settlement Agreement itself. As Exhibit 2 begins:

> Freestanding, embedded and regression based performance validity metrics will be administered to each Retired NFL Football Player during baseline and, if relevant, subsequent neuropsychological examinations. There will be at least seven performance validity metrics utilized during each assessment. The specific performance validity metrics utilized will not be released to the public in order to maintain the highest standards of assessment validity. The performance validity metrics employed will be rotated at intervals determined by the Appeals Advisory Panel in consultation with Co-Lead Class Counsel and Counsel for the NFL Parties.
>
> Each neuropsychological examiner must complete a checklist of validity criteria as set forth in Slick et al. 1999, and revised in 2013 (see below) for every Retired NFL Football Player examined in order to determine whether the Retired NFL Football Player's test data is a valid reflection of his optimal level of neurocognitive functioning.

The Exhibit then continues by describing nine validity criteria, derived from Slick's 2013 analysis. However, immediately after listing them, the Agreement makes clear that the *Slick* criteria are not dispositive:

> Notwithstanding a practitioner's determination of sufficient effort in accordance with the foregoing factors, a Retired NFL Football Player's failure on two or more effort tests may result in the Retired NFL Football Player's test results being subjected to independent review, or result in a need for supplemental testing of the Retired NFL Football Player.
>
> Note: Additional information relating to the evaluation of effort and performance validity will be provided in a clinician's interpretation guide.

The mandated *Slick* checklist, on which so much rests, results from work by neuropsychologists Daniel J. Slick, Elisbaeth M.S. Sherman, and Grant L. Iverson. Those experts

---

[2] Special Master Ruling on Pre-Diagnosis Evidence, at 6 (July 2, 2020), https://www.nflconcussionsettlement.com/–Docs/pre_diagnosis_evidence_sm.pdf.

aimed to develop criteria for discerning malingering—"false or grossly exaggerated physical or psychological symptoms that are voluntarily produced, motivated by external incentives"—in the context of neuropsychological evaluation.[3] They characterized their approach as one that "balance[d] specificity with flexibility":

> In particular, we attempted to include most possible sources of relevant data, and to consider the relative and cumulative weight of specific kinds of data. Nevertheless, every case is unique, and no set of criteria can cover every possible set of data and circumstances. Thus, the proposed diagnostic criteria are not intended for use in a reflexive or inflexible manner. The clinician must use the criteria in an integrative manner, recognizing that not all patients will be easily classified, and that in some instances there may be adequate justification to disagree with a diagnosis suggested by rigid application of the criteria to the available data.
>
> The appropriate approach is to treat malingering in the same manner as any other "disorder": as a diagnosis to be arrived at or rejected after a comprehensive evaluation. To conclude that a person is malingering, one must rule out the alternatives. A thorough consideration of differential diagnoses is required. Careful consideration of the consequences of diagnostic error is also required. Clinicians need to keep well in mind the limitations of assessment methodology and the cost of false positive errors. A "reasonable doubt" strategy should always be applied to decisions about the probability that a patient is malingering.
>
> Clinicians also need to be aware that diagnostic qualifiers such as possible or probable may be easier to use in research than clinical settings due to the difficulty of applying or conveying them in venues where discrete diagnoses are preferred (e.g., the courtroom). No criteria can be perfect, and diagnostic errors are bound to happen.[4]

The *Slick* criteria thus combine objective and subjective factors, and operate to guide and standardize a professional judgment—is the patient malingering?—on which reasonable minds may sometimes disagree. As their developers made clear, there is an inherent tension when such qualified judgments are shoehorned into binary legal outcomes.

The first *Slick* item is straightforward, though it has been disputed in recent appeals briefing. The Settlement directs clinicians to administer "performance validity embedded indicators or tests" and determine if those results are "suboptimal."[5] The "cutoffs for each test should be established based on empirical findings."[6] The Guide states:

---

[3] Daniel J. Slick & Elisabeth M.S. Sherman, *Differential Diagnosis of Malingering*, in MILD TRAUMATIC BRAIN INJURY: SYMPTOM VALIDITY ASSESSMENT AND MALINGERING 57 (Dominic A. Carone & Shane S. Bush eds., 2013); Daniel J. Slick, Elisabeth M.S. Sherman, & Grant L. Iverson, *Diagnostic Criteria for Malingered Neurocognitive Dysfunction: Proposed Standards for Clinical Practice and Research* 13 THE CLINICAL NEUROPSYCHOLOGIST 545, 551-58 (1999).

[4] Slick, Sherman, & Iverson, *supra* note 3, at 558 (paragraph breaks added).

[5] Settlement Agreement, Exhibit A-2, Section 2.

[6] As the relevant textbook states:

> Based on the Traumatic Brain Injury comparison group that is part of the ACS package, 2 of 5 test scores falling below either the 10th percentile of the Clinical Sample Base Rates, or 3 of 5 test scores falling below the 15th percentile would indicate high likelihood of invalid performance.[7]

In briefing here and elsewhere, counsel for the Class and for Claimants have argued that this Guide section means that if two of five ACS scores do *not* fall below the tenth percentile of the Clinical Sample, or three of five scores do not fall below the fifteenth percentile, the result is that those scores are neither invalid nor suboptimal under the Settlement. The NFL Parties, conversely, suggest that the ACS scores may be "suboptimal" even if they are not very likely to be invalid.[8] In short, the Claimants generally equate "invalid" with "suboptimal"; the NFL Parties argue that they have different meanings.

The Guide's "high likelihood of invalid performance" is ambiguously related to the Settlement's reference to "suboptimal" scoring. The Parties may consider revising the Guide. Until they do, I conclude that the most plausible reading is that the Guide provides a safe harbor for clinical judgment about the validity of the ACS scores. If the ACS validity scores do not indicate a "high likelihood of invalid performance," the Claims Administrator and its expert panel should defer if a clinician concludes that they are not "suboptimal" for the purposes of the first *Slick* criteria.[9]

By contrast, if the safe harbor is not satisfied, the clinician's judgment would lack foundation and could be set aside. And, because the Guide requires performance on the ACS tests to be evaluated against the other stand-alone measures, failures on the non-ACS instruments may

---

> Selecting a cut-off score for test interpretation requires a balance between sensitivity and specificity. Higher cut-off scores (i.e., closer to the mean) are more likely to identify those who have cognitive problems (improved sensitivity), but they are also more likely to include those who do not have cognitive problems (reduced specificity and increased false positives). On the other hand, lower cut-off scores (i.e., further away from the mean) are less likely to identify those with cognitive problems (reduced sensitivity and increased false negatives, especially in higher functioning people), but are also less likely to include those who do not have cognitive problems (improved specificity). Some clinicians may consider the balance of sensitivity and specificity when interpreting a single score, but fail to consider this balance when interpreting multiple scores.

Brian L. Brooks, Grant L. Iverson, & James A. Holdnack, *Understanding and Using Multivariate Base Rates with the WAIS-IV/WMS-IV* in WAIS-IV, WMS-IV, AND ACS: ADVANCED CLINICAL INTERPRETATION 75, 81 (James A. Holdnack, Lisa Whipple Drozdick, Lawrence G. Weiss, & Grant L. Iverson eds. 2013)

[7] Retired NFL Football Players' Baseline Assessment Program: Neuropsychologists Handbook (the Clinician's Interpretation Guide), at 8.

[8] In other briefing, the NFL Parties argue that while validity scores that meet thresholds that are set forth in the Guide represent a *per se* finding of a high likelihood of invalid performance, this does not mean that performance should automatically be determined to be valid. The NFL Parties insist that the BAP Handbook does not establish an inflexible standard. In doing so, they appear to argue that scores that are not necessarily invalid but do not clear a higher threshold may especially be factored into the *Slick* analysis.

[9] As the Guide points out, three of the ACS subtests lack normative data for individuals over 69. The Guide provides that for such claimants, other primary effort tests are the appropriate measures. Nothing in the text above thus speaks to the *Slick* 1 validity analysis for those older claimants.

trigger a finding that the *Slick 1* as a whole was suboptimal, even if the ACS test scores themselves were not.[10] In such circumstances, clinicians must form and articulate a holistic judgment.

This approach harmonizes the Guide's language with the Settlement's larger goal of providing a clear path to efficient and simple resolution of claims. And it does not end the inquiry: low, but not "suboptimal," scores may bear on the remaining *Slick* criteria.[11]

The remaining eight *Slick* criteria require the clinician to reconcile inconsistencies between medical, collateral, and observational evidence.[12] Properly undertaken, all "relevant aspects of an examinee's presentation and circumstances—such as the presence of psychosocial secondary gains that may be influencing examinee's behavior—should be noted as contributing or coexisting factors."[13] The Settlement explicitly states that the examiner "must complete a checklist of validity criteria as set forth in Slick et al. 1999, and revised in 2013 (see below) for every Retired NFL Football Player." And the Guide continues:

> After reviewing the embedded and stand-alone performance validity measures, the provider should confirm his or her conclusion by referencing the qualitative descriptions provided in the Slick Criteria, to determine if findings are inconsistent with known patterns of brain dysfunction, clinical observation, collateral reports, history, and/or known patterns of test performance before judging performance to be suboptimal.
>
> Instructions for judging each component point of the Slick Criteria is provided in Exhibit A, as well as an overall rating of the quality of the data collected.

Where a particular *Slick* criterion is inconsistent or discrepant, or criteria point in different directions, the examiner must thoroughly explain in writing why the testing was valid.[14] Indeed,

---

[10] As the agreement provides, *failure* on two or more effort tests (including non-ACS effort tests) may result in additional independent review. The relationship of "high likelihood of invalid performance" to "suboptimal" for the purposes of *Slick* 1 applies only to ACS tests; *failure* on non-ACS tests is well-defined in their relevant manuals.

[11] *See* Brooks, Iverson, & Holdnack, *supra* note 6, at 354 ("The a priori cut-off established by the clinician provides an indication that invalid performance is present. Further investigation of the obtained scores is necessary to identify probable malingering.").

[12] While the first criterion relates to an assessment of scores on performance validity test (1), the remaining eight require the physician to assess the following: (2) the discrepancy between neuropsychological test performance and "accepted models of normal and abnormal central nervous system (CNS) function"; (3) the discrepancy between test performance and observed behavior; (4) the discrepancy between the test data and reliable collateral reports; (5) the discrepancy between the test data and documented background history; (6) the discrepancy between self-reported history and documented history; (7) the discrepancy between self-reported history and known patters of brain functioning; (8) the discrepancy between self-reported symptoms and behavioral observations; and (9) and the discrepancy between self-reported symptoms and information obtained from collateral informants. Settlement Agreement, Exhibit A-2, Section 2.

[13] Slick & Sherman, *supra* note 3, at 71.

[14] This is explicitly required by MAF Rule 20(d), i.e., that the clinician "must explain in writing, in the method prescribed by the Claims Administrator and to the satisfaction of the Claims Administrator, any deviation from the BAP diagnostic criteria." A deviation regarding validity occurs when:

even absent obvious discrepancies, *discussing each Slick criteria* is a strongly recommended best practice, indicated both in the Settlement and the Guide. Admittedly, providing a full analysis of each criteria has not always been the practice of either MAF or BAP physicians. But this case well-illustrates the mischief resulting from unadorned checklists.

After completely describing the criteria's application, the Settlement requires the clinician to use the articulated checklist to develop and then state a gestalt judgment as to validity. That opinion should demonstrate a thorough consideration of a claim's inconsistencies and instances of potential invalidity, and use the clinician's best articulated medical judgement to resolve any inconsistencies and dispel suggestions of invalidity.

Next (in this description of how the process *should* work), the Claims Administrator evaluates the submitted claim and determines whether it satisfies the Settlement's proscriptions. It does so guided by the AAP and its Consultants, whose job it is to make sure that claim adjudication follows the medicine.[15] As MAF Rule 23 states:

> Two AAP Members serve as the Claims Administrator's AAP Leadership Council to provide the Claims Administrator advice and assistance on any medical issues arising in the monitoring of the work of Qualified MAF Physicians. This includes review of specific claims or groups of claims at the Claims Administrator's request to determine compliance by Qualified MAF Physicians with the Settlement Agreement and these Rules and whether a Claim Package reflects and supports the Qualifying Diagnosis stated in the Diagnosing Physician Certification form.

Respecting validity, the Parties contemplated that there might be times when the AAP will disagree with the examining clinician, particularly when their records are incomplete. In those cases, the Parties have jointly written guidance to the AAP about what it is to do:

> The Parties agree that, when the *Slick* criteria are a factor in connection with the denial of a given claim, the AAP/AAPC are to note it as a factor in the denial explanation and include a thorough review of the *Slick* criteria, including the results of the performance validity testing that was administered as part of the test battery, with reasoning or explanation with respect to how and/or why the *Slick* criteria factored into the determination.[16]

> The Retired NFL Football Player was assigned scores or results indicating that the player failed two or more of the embedded and/or stand-alone performance validity measures in the neuropsychological test battery and/or where the application of the clinical criteria for assessing performance validity under Slick et al. otherwise indicates that the test data may not be a valid reflection of his optimal level of neurocognitive functioning.

[15] For Pre-Effective Date claims outside of the BAP/MAF, for example, AAP review is mandatory. Settlement Agreement, Section 6.4(b). Otherwise, the Settlement empowers the review of the Claims Administrator in Section 6.4(b), and the AAP in Section 8.6.

[16] NFL Concussion Settlement Program: Guidance to the AAP and AAPC, at 2.

Here, again, the Parties have wisely focused on the importance of written, thorough explanations. Unlike the examining clinicians, the AAP is not necessarily required to describe each *Slick* item. Rather, the AAP must discuss those criteria that "factored into the determination" of a denial. If there is disagreement, the analysis must be thorough.

Putting together the clinicians' duty to provide a complete *Slick* analysis with the AAP's role of making sure that they follow an appropriate process leads naturally to a system of review, mostly consistent with past practice but which I now make explicit.

The AAP should defer to a clinician's *Slick*-criteria-based validity analysis when it results from reasoning completely articulated in contemporaneous reports, unless the analysis is clearly erroneous. Conversely, when clinicians fail to articulate their judgment through complete *Slick* analyses, the AAP may thoroughly and independently assure themselves the criteria do not indicate invalid testing.

Next, the Claims Administrator will normally accede to the judgment of the AAP Reviewer. And review on Appeal (in this ideal procedure) ought to be simple. The appellant bears a heavy burden of offering "clear and convincing evidence" that the Claims Administrator was wrong.[17] "Clear and convincing evidence" means that it is "highly probable or reasonably certain" that there was an error.[18]

Now, let us turn to this particular claim and consider how it contrasts with the ideal.

As I earlier described, Dr. [REDACTED] administered the neuropsychological testing on which Dr. [REDACTED] relied. In his report, Dr. [REDACTED] (without originally providing scores) described each of the embedded validity test scores as "optimal." He then put a simple checkmark next to "no" for each of the *Slick* criteria, while discussing none. His general analysis, consisting of one short paragraph, concluded that the "classification of performance validity is considered optimal and the current evaluation is considered to be a valid representation of Mr. [REDACTED] cognitive ability." Doc. 162285. Later, when discussing the MMPI-2-RF, Dr. [REDACTED] admitted the possibility of overreporting, but dismissed the concern due to what he considered to be Mr. [REDACTED] "heightened emotional distress and preoccupation with physical health." *Id.*

As Mr. [REDACTED] exceeded the safe harbor for ACS scores, and did not fail any of the other performance validity tests, it would have been reasonable for Dr. [REDACTED] to have concluded that the first *Slick* criteria was "not suboptimal." He did not explicitly make such a judgment with reference to the scoring. However, with the benefit of hindsight and the relevant scores, the Claims Administrator could have, and apparently did, conclude that Slick 1 was satisfied.

As for the remaining eight criteria, Dr. [REDACTED] analysis is, simply, not good enough. As the AAP describes, inconsistencies in Mr. [REDACTED] record implicate multiple *Slick* criteria:

---

[17] *See* Order Appointing Special Masters, at 5.

[18] *In re Fosamax Alendronate Sodium Prods. Liab. Litig.*, 852 F.3d 268, 285-86 (3d Cir. 2017).

- Mr. ██████ assessed communication skills during the assessment are discrepant with the severe impairment indicated by his Language domain scores. This implicates the third criterion—test results versus observed behavior.
- Mr. ██████ engagement in activities such as driving and assessing financial investment proposals contrasted with testing indicating Level 2 Neurocognitive Impairment. This implicates the fourth criterion—discrepancies between test data and reliable collateral reports.
- Moreover, comorbidity might exist between his neurocognitive impairment and an array of additional health disorders, including severe, untreated psychiatric disorders, insomnia, and chronic headaches. This implicates the sixth criterion—discrepancies between self-reported history and documented history.
- Mr. ██████ MMPI-2-RF scores reveal "possible over-reporting of psychiatric, somatic, and cognitive concerns." The MMPI-2-RF scores also implicate the additional criteria – i.e., whether self-reported symptoms comport with known patterns of brain functioning and behavioral observations.[19]

It is true that now, on Appeal, both Dr. ████ and Dr. █████ offer to explain pieces of this puzzle. Dr. █████ states that, for instance, low normative scores "do not indicate that someone is *necessarily* mute or otherwise unable to communicate. Therefore, there was no clear discrepancy between performance on language measures and his ability to communicate his thoughts and, thus, no need to indicate a [*Slick*] discrepancy." Doc. 225623, at 16-17. But "necessarily" is a word that indicates a grudging and adversarial practice respecting what are intended to be medical criteria. And a letter offered in support of an Appeal is no substitute for a medical judgment created contemporaneously with an exam, in a setting where expressing such a judgment, and not post hoc advocacy, is the order of the day.

Here, the AAP Reviewers have repeatedly expressed concern with Dr. ██████ judgment, but (at least as I read the file) much of their frustration relates to his original failure to fully describe and synthesize discrepant findings. There are many red flags in this file, and Dr. ██████ report mostly ignores them. By offering such a sparse report, which did not articulate the grounds for his judgment in the face of countervailing evidence, Dr. █████ undermined the AAP's ability to defer to his judgment.

The Claims Administrator has tried to sort this out across many reviews. But that appraisal built scaffolds on soft ground. The result is that the AAP (and the Claims Administrator) risk seeming to substitute their judgment for that of an examining physician, even though the truth is that the AAP has done exactly what it is supposed to do: making sure that the clinicians follow the medicine.

---

[19] *See* Special Master Ruling on Validity Testing and Cause of Functional Impairment (Aug. 19, 2020), https://www.nflconcussionsettlement.com/Docs/testing_impairment_sm.pdf. The original filed opinion mistakenly referred to these as *Slick* criteria eight and nine. In fact, they are criteria seven and eight, as an amended opinion clarified.

Given this analysis, there are two potential outcomes. The first is a denial. Appellants must demonstrate clearly and convincingly that the Claims Administrator erred. Mr. [REDACTED] has not made that showing here. The underlying reports do not offer sufficient evidence that Mr. [REDACTED] tests were valid measures. The Claims Administrator was justified in rejecting the Claim on that basis, among others.

However, I am hesitant to deny a Claim that has been so often delayed through no obvious fault of the Claimant himself. This is not a review process that any reasonable observer could look to with pride. It has been halting, tortuous, and expensive. Requiring Mr. [REDACTED] to start again from scratch is something the program should avoid if possible. A remand, though it is (somewhat unbelievably) the third for just this Claim, offers the possibility of a more efficient and equitable path.

Therefore, under Rule 24, I re(re)-remand this Claim to the Claims Administrator. I direct the Claims Administrator to ask Dr. [REDACTED] for a new analysis of Mr. [REDACTED] test validity. A conforming assessment at this point must describe why Dr. [REDACTED] believes that *each Slick* criterion is not a concern. He should pay particular and evenhanded attention to the discrepancies indicated by the AAP Reviewers.[20] And he must come to a fresh holistic judgment, as the Settlement requires, noting where possible why particular discrepancies ought not compel a finding that the testing was invalid.

If and when Dr. [REDACTED] completes such a *Slick* analysis, I direct the Claims Administrator to ask the AAP Leadership Council to review the file. As I have indicated, ordinarily deference will be due to clinicians' articulated medical judgments regarding the *Slick* criteria and their import. In this case, given the procedural history and Dr. [REDACTED] seemingly superficial previous analysis of the *Slick* criteria, the AAP Review may be unusually searching.

However, the Leadership Council should also bear in mind (in this case and others) the Settlement's structure of layered deference, and its need to avoid re-litigating each clinical judgment in later AAP Review. The Panel ought generally to police the process, not its outcomes.[21]

I fully expect that the Claims Administrator will defer to the results of this process. If necessary, either party may then appeal, raising such issues as remain.[22]

---

[20] Dr. [REDACTED] has partly addressed these issues in a letter attached to Mr. [REDACTED] Appeal. Doc. 225623. However, an appropriate *Slick* checklist analysis merits a more thorough discussion, especially in regards to Mr. [REDACTED] conversational nature during his neuropsychological assessment, as well as of the impact of his psychiatric condition. It is also the case that Dr. [REDACTED] has not had the benefit of this Opinion's statement of what is expected from our MAF Physicians, which the Claims Administrator will convey to him.

[21] Of course, the goal—shared by MAF, BAP, and AAP physicians alike—is to follow the medicine and to pay deserving claims. And, respecting particular claims, *de novo* review is explicitly called for by the Agreement itself.

[22] Even if the Council was to conclude that Mr. [REDACTED] testing was reliable, the Claims Administrator might still have chosen to deny the Claim (again) if the CDR analysis did not demonstrate functional impairment. The relevant Denial Notice appears to conclude that a Level 1.5, but not Level 2.0, Diagnosis would be indicated given the state of the record. The Claims Administrator may, on remand, come to the same or a different conclusion . My review for clear error will follow. The NFL Parties' arguments on this, and all other grounds, are of course preserved.

CONCLUSION

For the reasons given, Mr. ██████ Claim is remanded to permit Dr. ██████ to further provide analysis and documentation.

Date: October 21, 2020

____
David A. Hoffman, Special Master