ASSESSING THE VALIDITY AND RELIABILITY OF THE NEUROTOXICITY SCALE AND THE OPIOID CRAVING SCALE IN A SPECIAL POPULATION OF GENETICALLY DAMAGED PATIENTS

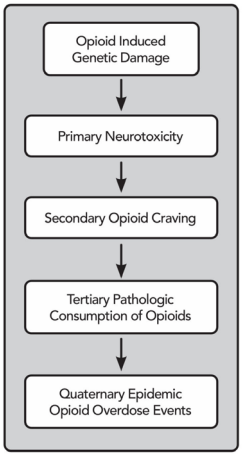
ABSTRACT

It remains our hypothesis that the genotoxicity due to repetitive opioid exposure ultimately results in a state of a neurotoxicity when opioid abstinence is attempted. This state of neurotoxicity is described by the victims as unbearable. According to the victims, the human body simply cannot endure this state of neurotoxicity. This neurotoxicity results in a secondary opioid craving as it is quickly learned that the full agonist opioids offer at least a temporary, but somewhat dangerous, respite from the horrors of the neurotoxicity. Subsequently, a tertiary aberrant opioid consumption associated with quaternary opioid overdose events manifests and overbears the life of the individual. The victims relate that their daily routine is now predominated by a near endless search for typically illegal opioids in a risky attempt to stave off the unbearable state of the neurotoxicity.

If any attempt at intervention into this deadly and all-consuming cycle is going to be made, a first step would include the ability to measure and quantify the levels of both the neurotoxicity and the opioid cravings. If these levels cannot be measured and quantified, management of the neurotoxicity and associated opioid cravings will remain elusive. For these reasons, we undertook the development of both a Neurotoxicity Scale and an Opioid Craving Scale. Knowing that the neurotoxicity associated with an opioid abstinence in the setting of the opioid-induced genotoxicity typically escalates in hours and the temporary respite from this neurotoxicity due to either a full or partial agonist opioid begins to manifest in minutes, we understood that we needed the scales to be able to respond to rapid changes in both the neurotoxicity and the opioid cravings. We examined currently existing scales but found none to be satisfactory for the requirements. We, therefore, undertook the task of designing and building our own two psychometric assessments, questionnaires, in order to be able to assess both the neurotoxicity and the opioid cravings along with any rapid changes in either clinical condition.

NEUROTOXICITY SCALE

1). I am yawning more than normal	0 1 2 3 4 5
2). My eyes are watering more than normal	0 1 2 3 4 5
3). My nose is running more than normal	0 1 2 3 4 5
4). I am having stomach cramps	0 1 2 3 4 5
5). I am vomiting	0 1 2 3 4 5
6). I have diarrhea	0 1 2 3 4 5
7). I am sweating more than normal	0 1 2 3 4 5
8). The hair on my body is standing on end	0 1 2 3 4 5
9). My heart is beating hard and fast	0 1 2 3 4 5
10). I feel anxious	0 1 2 3 4 5
11). I feel hot then cold	0 1 2 3 4 5
12). I have a tremor (shaking)	0 1 2 3 4 5
13). I feel like something bad is about to happen	0 1 2 3 4 5
14). I can't stand feeling this way	0 1 2 3 4 5



OPIOID CRAVING SCALE

1). If I had an opioid right now, I would take it.	0 1 2 3 4 5
2). I would not be able to stop myself from taking an opioid right now.	0 1 2 3 4 5
3). I would feel more in control of things if I could take an opioid right now.	0 1 2 3 4 5
4). Taking an opioid right now would make me feel better.	0 1 2 3 4 5
5). If I could take an opioid right now, I would feel less restless.	0 1 2 3 4 5
6). I am craving an opioid right now.	0 1 2 3 4 5
7). Using an opioid right now would make me feel better.	0 1 2 3 4 5

We understood that, in order to be effective, both scales, the Neurotoxicity Scale and the Opioid Craving Scale, would need to be valid and reliable. Specifically, we understood that these two questionnaires would need to be valid and reliable in the patient population for which they were intended, those individuals who had suffered from the genotoxicity due to repetitive opioid exposure. Validity is the extent to which a questionnaire measures what is intended to be measured 1. Reliability is the extent to which a questionnaire produces the same results on repeated trials under comparable circumstances 2. We found both the Neurotoxicity Scale and the Opioid Craving Scale to be very high in both validity and reliability. Psychometric Validity was evaluated by Construct Validity, Face Validity

and Content Validity and both questionnaires scored high on all three accounts. Psychometric Reliability was evaluated for Internal Consistency and Test-retest reliability. Again, both questionnaires scored well on all accounts.

Conclusion: Both the Neurotoxicity Scale and the Opioid Craving Scale are judged to be high in both Psychometric Validity and Psychometric Reliability. This analysis for Psychometric Validity and Psychometric Reliability was conducted within the targeted patient population of those individuals with laboratory proven hypermethylation within the OPRM1 gene and due to repetitive opioid exposure.

OPIOID CRAVING SCALE AT A GLANCE

Psychometric Validity
Construct Validity Strong (n = 15)
Content Validity Strong (n = 15)
Face Validity Strong (n = 8)

Psychometric Reliability
Internal Consistency Reliability
Cronbach's Alpha: Excellent (all values above 0.9)
Cronbach's Alpha with deleted variable: Excellent (all values above 0.9)
Pearson's (r): high degree of correlation (all values between 0.5 and 1.0)
Test-retest reliability
TOST Analysis with all variables deemed Equivalent

NEUROTOXICITY SCALE AT A GLANCE

Psychometric Validity
Construct Validity Strong (n = 15)
Content Validity Strong (n = 15)
Face Validity Strong (n = 8)

Psychometric Reliability
Internal Consistency Reliability
Cronbach's Alpha: Excellent (all values above 0.9)
Cronbach's Alpha with deleted variable: Excellent (all values above 0.9)
Pearson's (r): moderate to high degree of correlation (all values between 0.3 and 1.0)
Test-retest reliability
TOST Analysis with all variables deemed Equivalent

1). INTRODUCTION

When the emergence of fentanyl became obvious by 2015, the death rate only escalated. We understood that we were going to need an incubator model for the development of ideas and strategies in order to develop both the science and the treatment for the genotoxicity due to repetitive opioid exposure. The genotoxicity had not been yet recognized, much less prioritized for treatment. It was obvious to us that we were alone in our efforts. The model we developed was a decentralized model relying almost solely upon Smartphones and the existing telecommunications infrastructure. We did not require participants to have a Smartphone. But practically all participants did utilize a Smartphone. This reality of near Smartphone ubiquity gave rise to the opportunity to utilize true Ecological Momentary Assessment as our standard of research. We chose and developed further a software platform known as Formstacks for the majority of our data harvesting. Formstacks allowed us the opportunity to store our data in easy to query databases and with high levels of privacy and security for the individual participants. Our first IRB approved clinical trial was simply groundbreaking in that for the first time, a Catecholamine Toxicity was determined to be a plurality in opioid dependent participants experiencing a state of sickness due to a voluntary opioid abstinence. This Catecholamine Toxicity was of such a magnitude that an Emergency Halt was determined and the IRB was notified. 87% of the participants were found to have plasma levels consistent with a Catecholamine Toxicity. Catecholamine Toxicity is associated with cardiomyocyte necroptosis, Takotsubo Cardiomyopathy, and death 3,4,5. As a consequence of this Emergency Halt, we simply were not able to collect enough data to determine the reliability of either the Neurotoxicity Scale nor the Opioid Craving Scale during the short duration of our first clinical trial. That stated, we were able to collect adequate data for the validity analysis of the two questionnaires. And, most importantly, the validity analysis from the participants in the clinical

Sample ID	Average Promoter Hypermethylation	Average 25 CpGs Hypermethylation
5104	5.1	35.4
6628	5.8	37.0
6680	7.5	38.8
6765	8.0	37.9
7774	4.3	36.6
5826	19.8	46.7
6295	8.6	39.2
6753	8.7	39.8
6900	9.6	41.2
6997	9.2	40.4
7031	8.0	39.5
7099	10.1	41.0
7317	12.4	43.7
7341	8.3	39.7
7478	8.3	39.8
7488	8.8	40.3
7530	7.6	38.9
7532	13.0	37.5
29439	9.5	40.5
P-value	0.0349	0.0387
Average Control	6.1	37.1
Average Patients	10.1	40.6
Methylation Changes	-4.0	-3.4

HYPERMETHYLATION WITHIN THE OPRM1 GENE

The top five samples are from the opioid naive. The bottom 14 samples are from the opioid dependent

trial were derived from individuals with proven abnormalities within their DNA and due to repetitive opioid exposure. All 15 of the participants queried prior to the Emergency Halt had submitted a DNA sample for analysis. And every DNA sample successfully analyzed was found to have hypermethylation in the OPRM1 gene. This was judged to be of extreme importance. The two questionnaires were designed for use exclusively in a patient population that had suffered a genotoxicity due to repetitive opioid exposure. Therefore, it was deemed important that the validity and reliability analysis of the two questionnaires be done exclusively in the targeted patient population.

Having determined an Emergency Halt to the first clinical trial, we found ourselves in the position of never being able again to ask opioid dependent participants to voluntarily stop all opioids and allow themselves to become sick. It was judged that no amount of Informed Consent would ever be able to offset the risk of morbidity and mortality presented by widespread Catecholamine Toxicity. The risk to the participants was simply too high. This made further studies involving the state of the opioid related neurotoxicity problematic. We, therefore, understood that the only individuals that we were going to be able to include in a study were those individuals who went into the state of neurotoxicity due to an opioid abstinence before becoming participants in a study. Simply stated, we could not ethically ask human subjects to go into a state of neurotoxicity. The risk to the participant was too great. But, if an individual was experiencing the neurotoxicity at the time of initial contact, this individual could be included in a clinical trial without an ethical conflict. This was even more so true for a Retrospective Chart Review deemed by the IRB to be exempt from the necessity of an IRB approval. And this was the history of the data utilized in this manuscript for the statistical analysis for Psychometric Reliability of both of our questionnaires. As a final word on this subject, it should be stated that the same ethical considerations would prohibit the use of a placebo in the setting of an individual experiencing the opioid related neurotoxicity. Again, no amount of Informed Consent can offset the risk of morbidity and mortality presented by the clinical scenario that is Catecholamine Toxicity.

From the research perspective, we were fortunate in that the incubator model that we had chosen to build included on-boarding many patients who were sick at the initial contact with a neurotoxicity due to an opioid abstinence. Not only were we able to help these individuals get a proper diagnosis, we were further able to stabilize them with the antidote that is buprenorphine. In so doing, incredibly important data was being created and stored in the electronic medical records. This important medical data included the on-boarding and stabilization of thousands and thousands of individuals. This important medical data recorded each individual from the initial contact, often while the individual was in a state of the opioid related neurotoxicity, and through the initial stabilization with buprenorphine, and into a treatment plan of a chronic stabilization with buprenorphine. An Ecological Momentary Assessment (EMA) was conducted weekly. Included in this Ecological Momentary Assessment was both a Neurotoxicity Scale and an Opioid Craving Scale. Each EMA was worded to include the phrase "in the previous seven days…". Thus each EMA is interpreted to cover a seven day period. This made it feasible for us to rapidly accumulate a set of medical data that covered over 100,000 patient days. Knowing that critically important information with the ability to impact the national Opioid Crisis was gathered in the medical records, we approached the IRB for a verification that a Retrospective Chart Review, properly conducted, would be exempt from the requirement of an IRB approval. IRB exemption for the Retrospective Chart Review was obtained and the data was handled in a manner emphasizing the importance of patient confidentiality.

We knew that embedded within this data now accumulated was the information that we needed to complete the reliability analysis of both the questionnaires, the Neurotoxicity Scale and the Opioid Craving Scale. We had previously completed the Content Validity, Construct Validity, and Face Validity analysis on both questionnaires, the emphasis was now placed on data that would allow us to determine the Psychometric Reliability of both questionnaires. Fortunately, our results were highly queryable. We were able to rapidly identify individuals who had been suffering from the opioid related neurotoxicity upon initial contact with the telemedicine program that was serving also as a research incubator platform.

A series of meetings with the statistical consultants followed. The question was which set of data did the statistical consultants believe would result in a strong reliability analysis of the questionnaires. Ultimately, the statistical consultants focused on a group of 1053 patients, all of whom had been experiencing the opioid related neurotoxicity upon initial evaluation. All 1053 patients were treated with Buprenorphine and at a consistent dosage of 16 mgms a day. All 1053 patients were compliant and completed the weekly Ecological Momentary Assessment as required. As previously discussed, this weekly Ecological Momentary Assessment included both the Neurotoxicity Scale and the Opioid Craving Scale. For purposes of statistical analysis, three of these sets of questionnaires were analyzed for each of the 1053 patients: the initial set of questionnaires done while the patient was experiencing the sickness of the opioid related neurotoxicity; the set of questionnaires done during the second week; and the set of questionnaires done during the fourth week. This is the data that is presented in this manuscript. Patients who were not experiencing the opioid related neurotoxicity at the time of the initial evaluation were not included in this data. Patients who were not consistent with the buprenorphine at 16 mgms a day were not included in this data. Patients who were not compliant with the weekly Ecological Momentary Assessment were not included in this data.

2). LITERATURE REVIEW

Critical Test Theory (CTT) is the study of and attempt to reduce the errors inherent to the development and use of psychometric questionnaires. Critical Test Theory can be expressed mathematically with the formula:

$$X = T + E$$

wherein X is the observed score, T is the true score, and E is the error introduced by measurement process. It is apparent from this equation that Critical Test Theory begins with the assumption that an error rate is intrinsic to the process of measurement in a psychometric assessment. Obviously then, one would want to take the appropriate steps to minimize the errors within the measurement process. In the study of psychometrics, the intrinsic error rate in minimized by two factors: validity and reliability. As previously stated above, validity is simply the extent to which the questionnaire measures what it purports to measure. Validity typically relies upon opinion. This opinion as to validity is typically from both experts in the field and those with a personal knowledge of that which is purported to be measured. Reliability, again, as previously stated, the other half of the effort to minimize error inherent to psychometric measurement, is simply the extent to which the questionnaire produces the same results on repeated trials and under similar circumstances. The reliability of a questionnaire is typically presented as a statistical analysis of repeated questionnaires. In a highly simplified manner, validity pertains to the questions and reliability pertains to the score. And the analysis of our two psychometric questionnaires, the Neurotoxicity Scale and the Opioid Craving Scale becomes an understanding of their validity and reliability.

3). METHODS

As previously noted above, we understood as early as 2015 that an incubator format was going to be a requirement if we were going to develop the tools and methods that would be needed to impact the Opioid Crisis. The minimization of morbidity and mortality was the driving force. People were sick. People were dying. In 2015, the genotoxicity due to the opioids was not clinically recognized; the opioid related neurotoxicity was not clinically recognized; it was not clinically recognized that the opioid cravings experienced by those suffering from the opioid related neurotoxicity were due to the horror of the neurotoxicity as opposed to a desire "to get high" on the opioids. That was the prevailing belief at the time of our work. In a testimony to the pervasiveness and elasticity of Cognitive Bias, it was widely believed at this point in time that the Opioid Crisis is driven by a "brain disease" resulting in a desire by the individual to consume opioids "in order to get high"6. It was obvious to us that, in order to break through this Cognitive Bias and make an inroad into the morbidity and mortality associated with the Opioid Crisis, we were going to

need to have all three, the genotoxicity, the neurotoxicity, and the opioid cravings due to the neurotoxicity, clinically recognized. To us, this meant that we were going to need to be able to quantify all three: the genotoxicity, the neurotoxicity, and the opioid cravings. One of the three targets is an objective assessment of DNA methylation levels within the OPRM1 gene. Two of the three targets are a subjective assessment of exactly what an individual is experiencing when the horror of the neurotoxicity and opioid craving is upon them.

Quantification of the gentoxicity meant that we were going to need to be able to measure the level of methylation suffered by any one individual and within the OPRM1 gene. At this point in time, bisulfate conversion is the gold standard for the evaluation of levels of methylation. It is a somewhat costly and complex process. But we are fortunate that we have the technology. Development of the test strained our capacities. But, ultimately, we did develop an adequate test for the methylation levels of the OPRM1 gene due to repetitive opioid exposure. At the time of this writing, the DNA test we developed for opioid related methylation in the OPRM1 gene is undergoing a CLIA Certification process. This DNA test carry with it the ability to discern between a genotoxicity as opposed to a mental health disorder.

The objective testing of the DNA for methylation lent itself well to the laboratory world of isolation, magnification, and measurement. Such is the nature of objective testing. But the symptoms associated with the neurotoxicity and the level of opioid cravings were not objective measurements. These were purely in the subjective realm. Therefore, if we were going to develop the ability to measure either the level of neurotoxicity or level of opioid cravings experienced by any one individual, then these were going to require the use of psychometric questionnaires.

As noted above, we reviewed the existing questionnaires that had at least some overlap with our targeted construct. But no existing questionnaire was deemed adequate for our purposes. Furthermore, no existing questionnaire had undergone an analysis for validity and reliability specifically in the target population of those individuals who had suffered the methylation to their DNA due to repetitive opioid exposure. We, therefore, created our own questionnaires and with the intent to undergo an analysis for validity and reliability within the target population of individuals who had suffered a hypermethylation within their OPRM1 gene due to repetitive opioid exposure.

3.1). PSYCHOMETRIC VALIDITY

Psychometric Validity refers to the assessment's potential to evaluate what it claims to measure. If a test is valid, it means that the questions or test items are closely linked to the test's intended core focus 7. In the realm of research there are four commonly recognized types of research psychometric validity: Construct Validity, Content Validity, Face Validity, and Criterion Validity.

3.1.1). CONSTRUCT VALIDITY

Construct Validity is considered both the most important type of research validity and the most difficult to achieve. Construct Validity is how well a psychometric questionnaire measures the concept that it was designed to evaluate 8. In our specific circumstance, Construct Validity is simply that the Neurotoxicity Scale is accurately measuring the neurotoxicity and that the Opioid Craving Scale is accurately measuring the opioid cravings. We felt that only the participants who were experiencing the symptoms of the neurotoxicity and the opioid cravings could tell us whether or not our questionnaires were accurately reflecting the symptoms that they were experiencing and, and the severity of the symptoms, and, furthermore, when these specific symptoms had resolved. This information was captured in a series of interviews with the 15 participants in our first clinical trial. Each participant was questioned both while experiencing the neurotoxicity and the opioid cravings and two hours later after the initial dosage of buprenorphine when the neurotoxicity

and opioid cravings had greatly resolved.

3.1.2). CONTENT VALIDITY

Content Validity is an evaluation of how well a psychometric questionnaire covers all relevant parts of the concept that it was designed to evaluate 9. If one is to think of Construct Validity as an assessment of the accuracy of a test, then Content Validity is an assessment of the thoroughness of a test. Again, we felt that only the participants who were experiencing the symptoms of the neurotoxicity and the opioid cravings could tell us whether or not our questionnaires were thoroughly reflecting the symptoms that they were experiencing and, and the severity of the symptoms, and, furthermore, when these specific symptoms had resolved. This information was captured in a series of two interviews with the 15 participants in our first clinical trial. The participant was questioned both while experiencing the neurotoxicity and the opioid cravings and two hours later after the initial dosage of buprenorphine when the neurotoxicity and opioid cravings had greatly resolved.

3.1.3). FACE VALIDITY

Face Validity is the extent to which a psychometric questionnaire is subjectively viewed as covering the concept which it purports to measure 10. This subjective assessment of the psychometric questionnaire is one that is done by experts in the field. Face Validity is considered the weakest of the four research validities. Face Validity was a bit of a challenge for us as neither the opioid related neurotoxicity nor the secondary opioid cravings were widely recognized at the time of our analysis. We therefore were constrained to those experts who were working with us within our incubator platform. In all, eight experts were interviewed for our Face Validity.

3.1.4). CRITERION VALIDITY

Criterion Validity compares the psychometric questionnaire either to a future performance or to results from other well-established questionnaires. When the psychometric questionnaire is compared to a future performance, this is known as Predictive Validity. When the psychometric questionnaire is compared to an existing, well-established psychometric questionnaire, this is known as Concurrent Validity 11. Of the four types of research validities, we felt that Criterion Validity, both Predictive Validity and Concurrent Validity had the least to offer to the analysis of the validity of our two questionnaires as this is an emerging field of study. This emerging field of study is focused on a population of patients who have suffered a genetic damage due to repetitive opioid exposure. No other established questionnaire had undergone an analysis for validity and reliability specific to this subpopulation of patients. Therefore, there simply is not a psychometric tool that we felt was appropriate for a comparison. In conclusion, it was felt that an analysis by Criterion Validity was as likely to introduce bias and error as it was to be helpful.

3.2). PSYCHOMETRIC RELIABILITY

Psychometric Reliability refers to the ability of the assessment to produce similar results under consistent circumstances. If a test is reliable, it will produce similar results over time, repeated administrations, or similar conditions 12. There are a number of ways that reliability can be assessed. Typically, an assessment for reliability involves taking multiple measurements. Four types of Psychometric Reliability are commonly recognized in research: Parallel Forms of Reliability, Internal Consistency Reliability, Test-retest Reliability, and Inter-rater Reliability.

3.2.1). PARALLEL FORMS OF RELIABILITY

Parallel Forms of Reliability refers to the reliability obtained by administering different versions of an assessment tool 13. As this is an emerging field of study and as this was an initial psychometric assessment, we chose not to administer different versions of the psychometric questionnaires. The concern was the possibility of introducing bias as this is a known and recognized risk of administering multiple versions of a psychometric questionnaire.

3.2.2). INTERNAL CONSISTENCY RELIABILITY

Internal Consistency is a measure of the reliability of a psychometric assessment. Internal Consistency reflects the extent to which items within the instrument measure various aspects of the same construct 14. A common measure of Internal Consistency Reliability is through the use of coefficients such as Cronbach's Alpha or Pearson's (r). Cronbach's Alpha is an assessment of how closely related a set of items are as a group 15. As we had a large sample size at 1053 sets of multiple completed questionnaires, we also felt comfortable adding Pearson (r) as a further statistical evaluation of the linear relationship among the variables.

3.2.3). TEST-RETEST RELIABILITY

Test-retest Reliability is the closeness of the agreement between the results of successive measurements of the same measure, under the same circumstances at separate points in time 16. Our participants were stabilized on buprenorphine when they completed the questionnaires at both the two week period and the four week period. These then were the two sets of questionnaires completed under the same circumstances and at separate points in time. As we had a large data set at n=1053, it was decided that a Two One Sided Test for Equivalency would be a preferred method.

3.2.4). INTER-RATER RELIABILITY

Inter-rated Reliability is the extent to which two or more raters agree. As our two psychometric questionnaires are scored on a 0-5 Likert Scale, Inter-rater Reliability is of less concern.

4). RESULTS

4.1). PSYCHOMETRIC VALIDITY

4.1.1). CONSTRUCT VALIDITY
(N = 15 PARTICIPANTS INTERVIEWED TWICE, RESULTS = STRONG CONSTRUCT VALIDITY)

As we are designing a psychometric analysis for an emerging field, we are unable to rely upon either the existing literature or the experts for guidance on whether or not our questionnaires were accurately capturing what the participant was experiencing. We were in a position wherein accuracy was to be found via direct communication with the individual patient both when experiencing the neurotoxicity and the opioid cravings and then again when the symptoms of both had greatly resolved. Thus, when our participants in our first clinical trial came into the office setting on Day #3 of a voluntary sickness due to a voluntary opioid abstinence, this gave us the opportunity for a direct discussion about whether or not the participants felt that the questionnaires adequately represented the symptoms that they were experiencing in the moment. Without any hesitation or difference of opinion, all 15 participants felt that both questionnaires were exceptional at capturing the sickness that they were experiencing. As each participant was then given buprenorphine at 16 mgms sublingual and interviewed a second time after completing the questionnaires

again, this gave us the opportunity to interview the same 15 participants now that they were no longer experiencing the symptoms of the neurotoxicity or the opioid cravings. And again, and without exception, all 15 participants voiced the opinion that the two questionnaires had done a very adequate job at describing the symptoms that they were no longer experiencing. When asked if there were any additional symptoms or questions that could have been added in order to increase the completeness or clarity of the questionnaires, no participant added or requested any additional questions to be added or symptoms to be covered.

Of note, it was not our intent to stop this process at 15 participants in total. But when the first catecholamine levels began to return from the lab, an Emergency Halt was determined and the IRB notified. We were fortunate to have gathered the information we had to that point. Based upon the interviews with the 15 participants we were able to complete, we feel confident in stating that the Construct Validity of both the Neurotoxicity Scale and the Opioid Craving Scale is excellent. According to the participants experiencing the symptoms, both questionnaires are measuring that which they purport to measure. It is important to note that all 15 participants submitted a DNA sample for analysis of hypermethylation in the OPRM1 gene. Every sample analyzed revealed a hypermethylation within the OPRM1 gene. We are thus assured that we interviewed the correct target population.

4.1.2). CONTENT VALIDITY
(N = 15 PARTICIPANTS INTERVIEWED TWICE, RESULTS = STRONG CONTENT VALIDITY)

Again, our evaluation of the Content Validity of the two questionnaires was a parallel process to our evaluation of the Construct Validity of the two questionnaires. We interviewed 15 participants both while they were experiencing the neurotoxicity and the opioid cravings and then again, two hours later after completing the questionnaires for a second time, this time after the initial dosage of the buprenorphine had resolved both the neurotoxicity and the opioid cravings. Just as the participants felt universally that both questionnaires had done an exceptional job at capturing the essence of the symptoms of the neurotoxicity and the opioid cravings, the participants also indicated that both questionnaires had been thorough in the process. Not a single participant voiced an opinion, complaint, or concern that any aspect of either the neurotoxicity or the opioid cravings had been overlooked or neglected. For these reasons we are confident in stating that the Content Validity of both questionnaires is high. Of note, as with the Construct Validity, we had not planned on stopping at 15 participants. But the Emergency Halt due to widespread and severe Catecholamine Toxicity was an unexpected occurrence. Again, DNA analysis for the hypermethylation in the OPRM1 gene confirmed that these subjects were representative of the target patient population.

4.1.3). FACE VALIDITY (N = 8 EXPERTS QUERIED, RESULTS = STRONG FACE VALIDITY)

As we stated previously, Face Validity is considered the weakest of the four validities under Critical Test Theory. But we did interview 8 experts in the field. There was a consensus among all eight experts that the questionnaires were covering the context as intended. There was no dissension in opinion. It is therefore concluded that both psychometric questionnaires have a strong Face Validity.

4.1.4). CRITERION VALIDITY (NOT PREFORMED)

Criterion Validity, whether as Predictive Validity or as Concurrent Validity, had very little to offer to our analysis as this is a newly emerging field of study and is specific to a closely defined population with genetic damage due to repetitive opioid exposure.

4.2). PSYCHOMETRIC RELIABILITY

4.2.1). PARALLEL FORMS OF RELIABILITY (NOT PREFORMED)

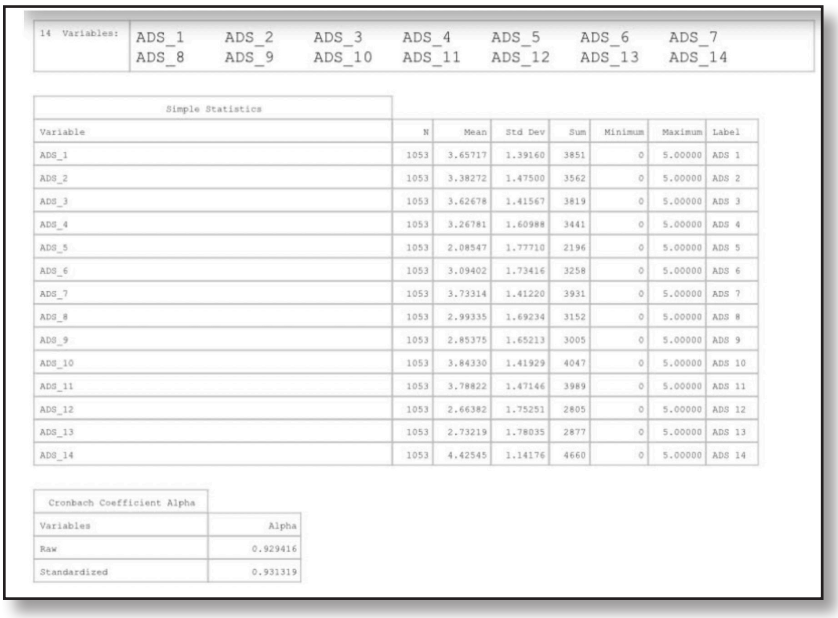
As previously stated, and so as not to introduce an unnecessary bias, we chose not to preform any Parallel Forms of Reliability in this newly emerging field of study.

4.2.2). INTERNAL CONSISTENCY RELIABILITY

As these are new questionnaires in an emerging field of study, we put emphasis on the importance of the Internal Consistency Reliability. We wanted a particularly strong Cronbach's Alpha analysis, if possible. This included also a separate Cronbach's Alpha analysis with each singular item deleted. Finally, for the sake of completeness, we undertook a Pearson's (r) analysis for the evaluation of linearity.

4.2.2.1). INTERNAL CONSISTENCY RELIABILITY- NEUROTOXICITY SCALE

Cronbach's Alpha and Cronbach's Alpha with each deleted variable:



14 Variables: ADS_1 ADS_2 ADS_3 ADS_4 ADS_5 ADS_6 ADS_7 ADS_8 ADS_9 ADS_10 ADS_11 ADS_12 ADS_13 ADS_14

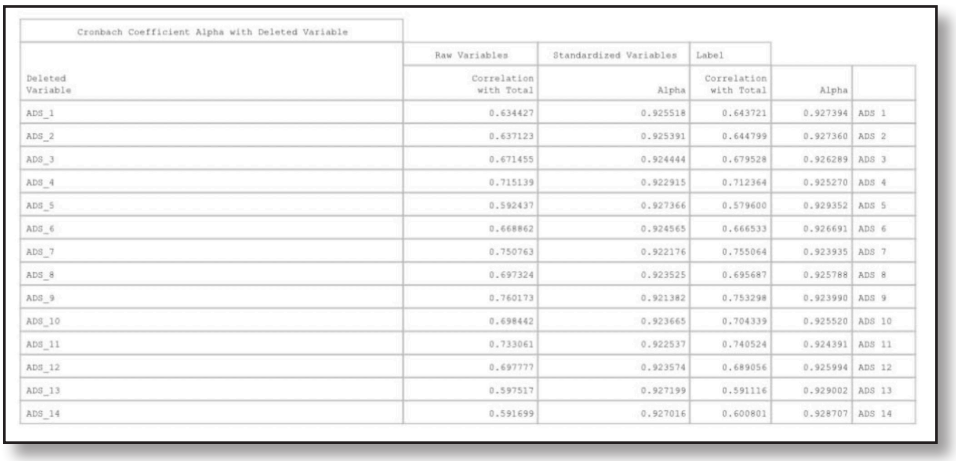
Simple Statistics

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum	Label
ADS_1	1053	3.65717	1.39160	3851	0	5.00000	ADS 1
ADS_2	1053	3.38272	1.47500	3562	0	5.00000	ADS 2
ADS_3	1053	3.62678	1.41567	3819	0	5.00000	ADS 3
ADS_4	1053	3.26781	1.60988	3441	0	5.00000	ADS 4
ADS_5	1053	2.08547	1.77710	2196	0	5.00000	ADS 5
ADS_6	1053	3.09402	1.73416	3258	0	5.00000	ADS 6
ADS_7	1053	3.73314	1.41220	3931	0	5.00000	ADS 7
ADS_8	1053	2.99335	1.69234	3152	0	5.00000	ADS 8
ADS_9	1053	2.85375	1.65213	3005	0	5.00000	ADS 9
ADS_10	1053	3.84330	1.41929	4047	0	5.00000	ADS 10
ADS_11	1053	3.78822	1.47146	3989	0	5.00000	ADS 11
ADS_12	1053	2.66382	1.75251	2805	0	5.00000	ADS 12
ADS_13	1053	2.73219	1.78035	2877	0	5.00000	ADS 13
ADS_14	1053	4.42545	1.14176	4660	0	5.00000	ADS 14

Cronbach Coefficient Alpha

Variables	Alpha
Raw	0.929416
Standardized	0.931319

Analysis of the Neurotoxicity Scale for Internal Consistency Reliability via Cronbach's Alpha shows that the Alpha scores fall in the category above 0.9, generally considered as excellent internal consistency for the variables within the Neurotoxicity Scale.

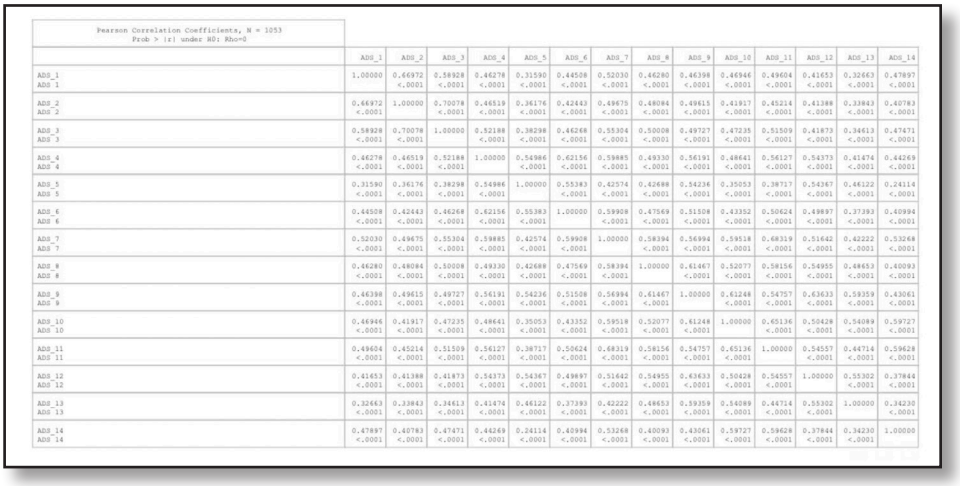


Cronbach Coefficient Alpha with Deleted Variable

Deleted Variable	Raw Variables: Correlation with Total	Raw Variables: Alpha	Standardized Variables: Correlation with Total	Standardized Variables: Alpha	Label
ADS_1	0.634427	0.925518	0.643721	0.927394	ADS 1
ADS_2	0.637123	0.925391	0.644799	0.927360	ADS 2
ADS_3	0.671455	0.924444	0.679528	0.926289	ADS 3
ADS_4	0.715139	0.922915	0.712364	0.925270	ADS 4
ADS_5	0.592437	0.927366	0.579600	0.929352	ADS 5
ADS_6	0.668862	0.924565	0.666533	0.926691	ADS 6
ADS_7	0.750763	0.922176	0.755064	0.923935	ADS 7
ADS_8	0.697324	0.923525	0.695687	0.925788	ADS 8
ADS_9	0.760173	0.921382	0.753298	0.923990	ADS 9
ADS_10	0.698442	0.923665	0.704339	0.925520	ADS 10
ADS_11	0.733061	0.922537	0.740524	0.924391	ADS 11
ADS_12	0.697777	0.923574	0.689056	0.925994	ADS 12
ADS_13	0.597517	0.927199	0.591116	0.929002	ADS 13
ADS_14	0.591699	0.927016	0.600801	0.928707	ADS 14

Pearson's (r) analysis:

As the majority of the coefficient values fall in the high degree of correlation (between 0.5 and 1.0) or the moderate degree of correlation (between 0.3 and 0.49), it can be concluded that a moderate to high degree of correlation exists among the variables within the Neurotoxicity Scale.



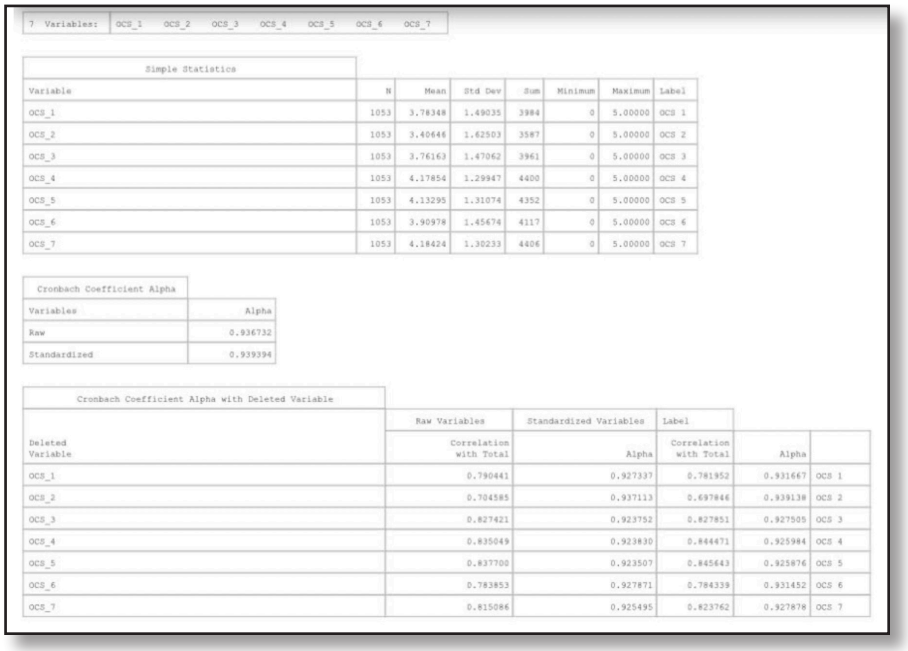
Pearson Correlation Coefficients, N = 1053
Prob > |r| under H0: Rho=0

	ADS_1	ADS_2	ADS_3	ADS_4	ADS_5	ADS_6	ADS_7	ADS_8	ADS_9	ADS_10	ADS_11	ADS_12	ADS_13	ADS_14
ADS_1	1.00000	0.66972 <.0001	0.58928 <.0001	0.46278 <.0001	0.31590 <.0001	0.44508 <.0001	0.52030 <.0001	0.46280 <.0001	0.46398 <.0001	0.46946 <.0001	0.49604 <.0001	0.41653 <.0001	0.32663 <.0001	0.47897 <.0001
ADS_2	0.66972 <.0001	1.00000	0.70078 <.0001	0.46519 <.0001	0.36176 <.0001	0.42443 <.0001	0.49675 <.0001	0.48084 <.0001	0.49615 <.0001	0.41917 <.0001	0.45214 <.0001	0.41388 <.0001	0.33843 <.0001	0.40783 <.0001
ADS_3	0.58928 <.0001	0.70078 <.0001	1.00000	0.52188 <.0001	0.38298 <.0001	0.46268 <.0001	0.55304 <.0001	0.50008 <.0001	0.49727 <.0001	0.47235 <.0001	0.51509 <.0001	0.41873 <.0001	0.34613 <.0001	0.47471 <.0001
ADS_4	0.46278 <.0001	0.46519 <.0001	0.52188 <.0001	1.00000	0.54986 <.0001	0.62156 <.0001	0.59885 <.0001	0.49330 <.0001	0.56191 <.0001	0.48641 <.0001	0.56127 <.0001	0.54373 <.0001	0.41474 <.0001	0.44269 <.0001
ADS_5	0.31590 <.0001	0.36176 <.0001	0.38298 <.0001	0.54986 <.0001	1.00000	0.55383 <.0001	0.42574 <.0001	0.42698 <.0001	0.54236 <.0001	0.35053 <.0001	0.38717 <.0001	0.54367 <.0001	0.46122 <.0001	0.24114 <.0001
ADS_6	0.44508 <.0001	0.42443 <.0001	0.46268 <.0001	0.62156 <.0001	0.55383 <.0001	1.00000	0.59908 <.0001	0.47569 <.0001	0.51508 <.0001	0.43352 <.0001	0.50624 <.0001	0.49897 <.0001	0.37393 <.0001	0.40994 <.0001
ADS_7	0.52030 <.0001	0.49675 <.0001	0.55304 <.0001	0.59885 <.0001	0.42574 <.0001	0.59908 <.0001	1.00000	0.58394 <.0001	0.56994 <.0001	0.59518 <.0001	0.68319 <.0001	0.51642 <.0001	0.42222 <.0001	0.53268 <.0001
ADS_8	0.46280 <.0001	0.48084 <.0001	0.50008 <.0001	0.49330 <.0001	0.42688 <.0001	0.47569 <.0001	0.58394 <.0001	1.00000	0.61467 <.0001	0.52077 <.0001	0.58156 <.0001	0.54955 <.0001	0.48653 <.0001	0.40093 <.0001
ADS_9	0.46398 <.0001	0.49615 <.0001	0.49727 <.0001	0.56191 <.0001	0.54236 <.0001	0.51508 <.0001	0.56994 <.0001	0.61467 <.0001	1.00000	0.61248 <.0001	0.54757 <.0001	0.63633 <.0001	0.59359 <.0001	0.43061 <.0001
ADS_10	0.46946 <.0001	0.41917 <.0001	0.47235 <.0001	0.48641 <.0001	0.35053 <.0001	0.43352 <.0001	0.59518 <.0001	0.52077 <.0001	0.61248 <.0001	1.00000	0.65136 <.0001	0.50428 <.0001	0.54089 <.0001	0.59727 <.0001
ADS_11	0.49604 <.0001	0.45214 <.0001	0.51509 <.0001	0.56127 <.0001	0.38717 <.0001	0.50624 <.0001	0.68319 <.0001	0.58156 <.0001	0.54757 <.0001	0.65136 <.0001	1.00000	0.54557 <.0001	0.44714 <.0001	0.59628 <.0001
ADS_12	0.41653 <.0001	0.41388 <.0001	0.41873 <.0001	0.54373 <.0001	0.54367 <.0001	0.49897 <.0001	0.51642 <.0001	0.54955 <.0001	0.63633 <.0001	0.50428 <.0001	0.54557 <.0001	1.00000	0.55302 <.0001	0.37844 <.0001
ADS_13	0.32663 <.0001	0.33843 <.0001	0.34613 <.0001	0.41474 <.0001	0.46122 <.0001	0.37393 <.0001	0.42222 <.0001	0.48653 <.0001	0.59359 <.0001	0.54089 <.0001	0.44714 <.0001	0.55302 <.0001	1.00000	0.34230 <.0001
ADS_14	0.47897 <.0001	0.40783 <.0001	0.47471 <.0001	0.44269 <.0001	0.24114 <.0001	0.40994 <.0001	0.53268 <.0001	0.40093 <.0001	0.43061 <.0001	0.59727 <.0001	0.59628 <.0001	0.37844 <.0001	0.34230 <.0001	1.00000

4.2.2.2). INTERNAL CONSISTENCY RELIABILITY - OPIOID CRAVING SCALE

Cronbach's Alpha and Cronbach's Alpha with each deleted variable:

Analysis of the Opioid Craving Scale for Internal Consistency Reliability via Cronbach's Alpha shows that the Alpha scores fall in the category above 0.9, generally considered as excellent internal consistency for the variables within the Opioid Craving Scale.



7 Variables: OCS_1 OCS_2 OCS_3 OCS_4 OCS_5 OCS_6 OCS_7

Simple Statistics

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum	Label
OCS_1	1053	3.78348	1.49035	3984	0	5.00000	OCS 1
OCS_2	1053	3.40646	1.62503	3587	0	5.00000	OCS 2
OCS_3	1053	3.76163	1.47062	3961	0	5.00000	OCS 3
OCS_4	1053	4.17854	1.29947	4400	0	5.00000	OCS 4
OCS_5	1053	4.13295	1.31074	4352	0	5.00000	OCS 5
OCS_6	1053	3.90978	1.45674	4117	0	5.00000	OCS 6
OCS_7	1053	4.18424	1.30233	4406	0	5.00000	OCS 7

Cronbach Coefficient Alpha

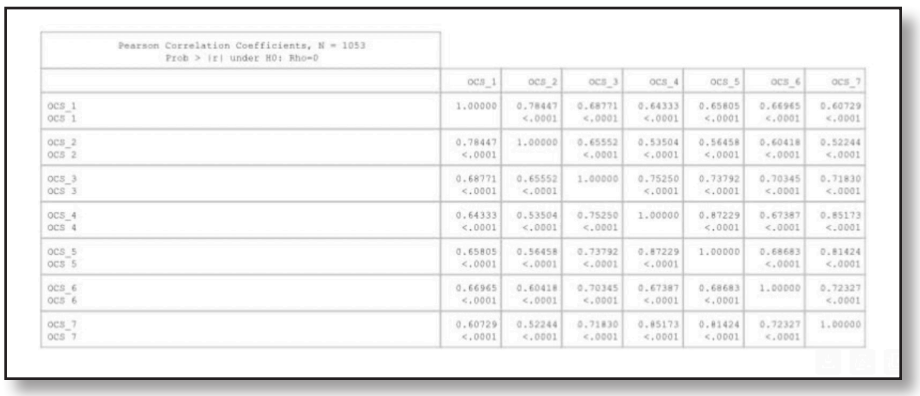
Variables	Alpha
Raw	0.936732
Standardized	0.939394

Cronbach Coefficient Alpha with Deleted Variable

Deleted Variable	Raw Variables: Correlation with Total	Raw Variables: Alpha	Standardized Variables: Correlation with Total	Standardized Variables: Alpha	Label
OCS_1	0.790441	0.927337	0.781952	0.931667	OCS 1
OCS_2	0.704585	0.937113	0.697846	0.939138	OCS 2
OCS_3	0.827421	0.923752	0.827851	0.927505	OCS 3
OCS_4	0.835049	0.923830	0.844471	0.925984	OCS 4
OCS_5	0.837700	0.923507	0.845643	0.925876	OCS 5
OCS_6	0.783853	0.927871	0.784339	0.931452	OCS 6
OCS_7	0.815086	0.925495	0.823762	0.927878	OCS 7

Pearson's (r) analysis:

As every single of the coefficient values fell in the high degree of correlation (between 0.5 and 1.0), it can be concluded that a high degree of correlation exists among the variables within the Opioid Craving Scale



Pearson Correlation Coefficients, N = 1053
Prob > |r| under H0: Rho=0

	OCS_1	OCS_2	OCS_3	OCS_4	OCS_5	OCS_6	OCS_7
OCS_1	1.00000	0.78447 <.0001	0.68771 <.0001	0.64333 <.0001	0.65805 <.0001	0.66965 <.0001	0.60729 <.0001
OCS_2	0.78447 <.0001	1.00000	0.65552 <.0001	0.53504 <.0001	0.56458 <.0001	0.60418 <.0001	0.52244 <.0001
OCS_3	0.68771 <.0001	0.65552 <.0001	1.00000	0.75250 <.0001	0.73792 <.0001	0.70345 <.0001	0.71830 <.0001
OCS_4	0.64333 <.0001	0.53504 <.0001	0.75250 <.0001	1.00000	0.87229 <.0001	0.67387 <.0001	0.85173 <.0001
OCS_5	0.65805 <.0001	0.56458 <.0001	0.73792 <.0001	0.87229 <.0001	1.00000	0.68683 <.0001	0.81424 <.0001
OCS_6	0.66965 <.0001	0.60418 <.0001	0.70345 <.0001	0.67387 <.0001	0.68683 <.0001	1.00000	0.72327 <.0001
OCS_7	0.60729 <.0001	0.52244 <.0001	0.71830 <.0001	0.85173 <.0001	0.81424 <.0001	0.72327 <.0001	1.00000

4.2.3). TEST - RETEST RELIABILITY

To test for reliability between the 2-week and 4-week timepoints, paired t-tests for equivalence of means were conducted using a null hypothesis lower equivalence limit of less than or equal to -0.5 and upper equivalence limit of greater than or equal to 0.5. The alternative hypothesis was the mean difference was between -0.5 and 0.5 on each variable on the Neurotoxicity Scale (NS) as well as the Opioid Craving Scale (OCS). Each t-test tested:

Where represents the mean of the response at 4-weeks minus the response at 2-weeks. The sample mean difference, a 95% confidence interval, and corresponding p-value were computed for each equivalence test. Equivalence bounds of -0.5 and 0.5 were chosen, i.e., the null hypothesis would be rejected if the 95% confidence interval was contained within these bounds. For each variable on the NS and OCS scales, the data suggest ($p < 0.0001$) that the responses at 4-weeks vs. 2-weeks are equivalent with 95% confidence. This suggests that once the subjects were given Buprenorphine, their responses remained stable at the 2-week timepoint through the 4-week timepoint.

The above method does have limitations for this dataset. Specifically, the data used in this study was ordinal and sparse. Therefore, we cannot assume continuity or normality. That said, the sample size (n = 1053) was large enough that these assumptions are less of a concern.

4.2.3.1). RESULTS FROM NEUROTOXICITY SCALE:

N	Mean	Std Dev	Std Err	Minimum	Maximum
1053	-0.0199	0.1399	0.00431	-1.0000	0

Mean	95% CL Mean		Std Dev	95% CL Std Dev	
-0.0199	-0.0284	-0.0115	0.1399	0.1341	0.1461

TOST Level 0.05 Equivalence Analysis

Mean	Lower Bound		90% CL Mean		Upper Bound	Assessment	
-0.0199	-0.5	<	-0.0270	-0.0128	<	0.5	Equivalent

Test	Null	DF	t Value	P-Value
Upper	-0.5	1052	111.37	<.0001
Lower	0.5	1052	-120.63	<.0001
Overall				<.0001

Variable: ADS2

N	Mean	Std Dev	Std Err	Minimum	Maximum
1053	-0.00285	0.0533	0.00164	-1.0000	0

Mean	95% CL Mean		Std Dev	95% CL Std Dev	
-0.00285	-0.00607	0.000376	0.0533	0.0511	0.0557

TOST Level 0.05 Equivalence Analysis

Mean	Lower Bound		90% CL Mean		Upper Bound	Assessment	
-0.00285	-0.5	<	-0.00555	-0.00014	<	0.5	

It is noted that in each and every variable within both the Neurotoxicity Scale and the Opioid Craving Scale that the results of the Two One Sided Test for Equivalency was that the results were Equivalent.

5). DISCUSSION

As stated above, Psychometric Validity is the analysis of whether or not the psychometric assessment is measuring what it purports to measure. This includes whether or not any special population is the target of the analysis. In our instance, we wished to develop two separate questionnaires: one to evaluate the level of neurotoxicity among those genetically damaged by the opioids; and one to evaluate the level of opioid craving among those genetically damaged by the opioids. Specifically, we wanted to utilize these two questionnaires within the population of those individuals whom had been genetically damaged by the opioids. Therefore, in our validity analysis of these two questionnaires, we interviewed exclusively individuals whom had been determined by laboratory evaluation to have the hypermethylation within the OPRM1 gene. From a common sense perspective, it would not make sense to ask individuals who had never experienced the opioid related neurotoxicity to give an opinion as to whether or not our questionnaire was capturing the essence of the opioid related neurotoxicity. Likewise, it would not make sense to ask individuals who had never experienced the neurotoxicity related opioid cravings to give an opinion as to whether or not our questionnaire was capturing the essence of the opioid cravings. By limiting our validity analysis only to those individuals with laboratory proven damage to their DNA and as a result of repetitive opioid exposure, we were, in essence, assuring that our validity analysis would be accurate.

When the question was limited to the correct genetically damaged patient population, the opinions given by the participants was that both questionnaires were doing an excellent job of measuring exactly what they were purporting to measure. Both questionnaires were found to have a very strong Content Validity and Construct Validity. Furthermore, both questionnaires were found to have a very strong Face Validity. Criterion Validity was less of a concern for us as this is an emerging field of study.

Again, as stated above, Psychometric Reliability refers to the assessments ability to produce similar results under consistent circumstances. Both of the questionnaires were determined to have exceedingly high Internal Consistency with all Cronbach's Alpha scores well above 0.90. Cronbach's Alpha scores in this range are judged as indicative of very high Internal Consistency. This statement remains true even with each variable deleted for further evaluation. Under a Pearson's (r) evaluation, all variables for the Neurotoxicity Scale fell between 0.3 and 0.49 indicating that a moderate to high degree of correlation exists among the variables within the Neurotoxicity Scale. The Pearson's (r) evaluation for the Opioid Craving Scale showed an even higher level of correlation exists among the variables with all variables having a coefficient between 0.5 and 1.0 indicating that a high degree of correlation exists among the variables within the Opioid Craving Scale. Paired t-test for equivalence was performed for each variable in both the Neurotoxicity Scale and the Opioid Craving Scale. Each variable in both scales returned with strong equivalency when analyzed for both the two week time frame and the four week time frame.

The Neurotoxicity Scale has exceptional Psychometric Validity and Psychometric Reliability for use within the patient population of those individuals whom have experienced the genotoxicity due to repetitive opioid exposure. The Opioid Craving Scale has exceptional Psychometric Validity and Psychometric Reliability for use within the patient population of those individuals whom have experienced the genotoxicity due to repetitive opioid exposure.

BIBLIOGRAPHY

1. Atikuzzaman, Md; Ahmed, S.M. Zabed, (2023),. What is a validated test instrument or survey? - Ask Us!, Journal of Academic Librarianship, Vol 49, Issue 1
https://doi.org/10.1016/j.acalib.2022.102623

2. Wu, Penglong, Cai, Mingqu, Liu, Kinbau, Want, Xuejun, (2021) Catecholamine Surges Cause Cardiomyocyte Necroptosis via a RIPK1–RIPK3-Dependent Pathway in Mice, Sec. Cardiovascular Therapeutics, Vol 8
https://doi.org/10.3389/fcvm.2021.740839

3. Pelliccia, F., Kaski, J., Crea, F., & Camici, P., (2017) Pathophysiology of Takotsubo Syndrome, Circulation. 135:2426–2441
https://doi.org/10.1161/CIRCULATIONAHA.116.027121

4. Lu, W.-H., Chen, H.-H, Chen, B-H, Lee, J.-C, Lai, C.-C, Li, C.-H, & Tseng, C.-J, (2020) Norepinephrine Leads to More Cardiopulmonary Toxicities than Epinephrine by Catecholamine Overdose in Rats, Toxics, 16:8(3)69
https://doi.org/10.3390/toxics8030069

5. Volkow, N., Koob, G., & McLellan, T. (2016) Neurobiologic Advances from the Brain Disease Model of Addiction, N Engl J Med, 374:363-371
https://www.nejm.org/doi/full/10.1056/nejmra1511480

6. Professional Testing, Inc., Test Topics
https://www.proftesting.com/test_topics/test_quality_validity.php#:~:text=The%20term%20validity%20refers%20to,a%20specific%20job%20or%20occupation

7. Linder, J., (2023) Construct Validity in Psychology | Definition, Types & Examples, Study.com
https://study.com/academy/lesson/construct-validity-in-psychology-definition-examples-quiz.html

8. Nikolopoulou, K., (2022) What Is Content Validity? | Definition & Examples, Scribbr
https://www.scribbr.com/methodology/content-validity/

9. Holden, R. (2010) Face validity, Wikipedia
https://en.wikipedia.org/wiki/Face_validity

10. Criterion Related Validity - an overview, ScienceDirect
https://www.sciencedirect.com/topics/nursing-and-health-professions/criterion-related-validity

11. Psychometric Reliability Meaning and Definition | Define psychometric reliability , Mercer| Mettl Glossary
https://mettl.com/glossary/p/psychometric-reliability/

12. Types of Reliability - Research Methods Knowledge Base, Conjointly
https://conjointly.com/kb/types-of-reliability/

13. Revicki, D., (2014) Internal Consistency Reliability, Encyclopedia of Quality of Life and Well-Being Research
https://link.springer.com/referencework/10.1007/978-94-007-0753-5

14. What does Cronbach's alpha mean? | SPSS FAQ
https://stats.oarc.ucla.edu/spss/faq/what-does-cronbachs-alpha-mean/

15. Wikipedia, Repeatability
https://en.wikipedia.org/wiki/Repeatability