Automated Quantification of Opioid Withdrawal in Neonatal Rat Pups Using Ethovision® XT Software

Hanna J. Kulbeth[a], **Saki Fukuda**[b], **Lisa K. Brents**[b,*]

[a]College of Medicine, University of Arkansas for Medical Sciences

[b]Department of Pharmacology and Toxicology, College of Medicine, University of Arkansas for Medical Sciences

Abstract

Chronic prenatal exposure to opioids often causes fetal opioid dependence that leads to neonatal opioid withdrawal syndrome (NOWS) shortly after delivery. Rat models of NOWS often require quantifying neonatal withdrawal behaviors using time-consuming, labor-intensive manual scoring methods. The goal of this study was to automate quantification of opioid withdrawal in neonatal rat pups. Accordingly, we used the animal behavior software Ethovision® XT to analyze archived videos of rat pups subjected to precipitated opioid withdrawal testing on postnatal day 0. We compared results obtained from Ethovision® XT with those previously obtained from manual scoring. Two endpoints reported by Ethovision® XT, Distance Moved (cm) and Movement Duration (s), had strong positive linear relationships with manually derived global withdrawal scores (GWS; $R^2 > .73$). Sensitivity and specificity of each endpoint to discriminate presence and absence of low-grade withdrawal were assessed by receiver operator characteristic curve analysis, which indicated that Distance Moved and Movement Duration had excellent accuracy (AUC > .90). Finally, we analyzed main and interaction effects of prenatal treatment (with vehicle or mu opioid receptor full agonists) and postnatal challenge (with saline or an opioid receptor antagonist) on each endpoint and determined they were similar for the manual and automated methods. These results show that Ethovision® XT software can reliably quantify opioid withdrawal in neonatal rat pups with non-inferiority to manual scoring even in videos that were not originally purposed and optimized for Ethovision® XT analysis. This faster and less labor-intensive method of analysis is expected to accelerate progress in preclinical studies of NOWS.

Keywords

Neonatal Opioid Withdrawal Syndrome (NOWS); rat model; Noldus Ethovision® XT; locomotor activity

---

[*]Corresponding author Postal address: Lisa Brents, PhD, Department of Pharmacology and Toxicology, 4301 West Markham Street, Slot 611, Little Rock, AR, 72205, LBRENTS@uams.edu.

Introduction

In the United States, the increased prevalence of opioid abuse during pregnancy has increased diagnoses of neonatal abstinence syndrome (NAS) by more than 5-fold since the early 2000s [1,2]. NAS caused by opioid exposure is also known as neonatal opioid withdrawal syndrome (NOWS) and is characterized in humans as a constellation of neurological, autonomic, and gastrointestinal dysfunctions that can present during the first few days of life and continue for several weeks [3]. Neurological problems include extreme irritability, tremors, seizures, and poor sleep [3]. Autonomic dysregulation causes excessive yawning, sweating, nasal stuffiness, and sneezing, and gastrointestinal problems include poor feeding, vomiting, and diarrhea[3]. NOWS can also increase respiratory rate and interfere with stable maintenance of body temperature [3]. Altogether, these problems may hinder growth and the ability to thrive [3]. Children with NOWS often require admission to a neonatal intensive care unit, and the aggregate hospital charges associated with care for NOWS patients has surpassed $1 billion [4].

Currently, there is a dearth of preclinical research investigating mechanisms and treatments of NOWS, which will be required to identify novel therapeutic targets and develop improved medications to mitigate this condition. Rodent models are useful for characterizing NOWS in the absence of common environmental confounders found in clinical populations, including polysubstance use, stress, poor prenatal care, and others[5–7]. Rodent models have also been used to assess the ontogeny of opioid dependence [8], to determine the involvement of each opioid receptor subtype in NOWS [9], to examine contributions of other neurotransmitter systems to the development of NOWS [10,11], and to test potential treatments for NOWS [12]. Generally, rodent models of NOWS consist of chronically exposing pregnant rats to opioids and observing their pups for signs of antagonist-precipitated opioid withdrawal on the day of their birth (postnatal day 0, PND 0)[13]. Opioid-dependent neonatal rat pups in withdrawal move considerably more than vehicle-treated controls [6,8]; therefore, withdrawal signs reported in these studies are primarily related to manually scored movements that include body stretches, body curls, and movements of the forelimbs, hindlimbs, and head[8,14]. Although our laboratory has used this method with high interrater reliability, it is highly time- and labor-intensive, which slows the progress of research.

To our knowledge, there are no reports of automated methods that quantify opioid withdrawal in rat pups during the early neonatal period. Because opioid withdrawal in neonatal rat pups presents as increased movement, we sought to develop and validate use of the animal behavior software Ethovision® XT to quantify neonatal rat pup movements from archived videos taken on PND 0 during precipitated opioid withdrawal testing. Our videos of the rat pups were originally intended for manual grading and therefore were not optimized for Ethovision® XT analysis at the time of testing. Regardless, we quantified pups' Distance Moved (in centimeters) and Movement Duration (in seconds) in these videos using Ethovision® XT software and compared the results to manually derived global withdrawal scores (GWS)[6] to confirm that the automated Ethovision® XT method quantifies withdrawal as well as or better than manual scoring. We then compared the sensitivity and specificity of manual and automated quantification to detect low-grade withdrawal by evaluating rat pup behavior on PND 0 following prenatal exposure to morphine, a positive control for

dependence, relative to behavior following prenatal exposure to vehicle. Last, we repeated group-level analyses previously performed with GWS using Distance Moved and Movement Duration to determine if similar conclusions are reached using manual and automated endpoints. An improved ability of automated compared to manual methods to detect and quantify neonatal movements associated with NOWS would be helpful in preclinical studies designed to compare the efficacy of treatments either administered prenatally or after birth. Indeed, the results of the present study suggest that the automated Ethovision® XT method, which is faster and less labor-intensive than manual scoring, provides a highly sensitive and specific method to quantify opioid withdrawal in neonatal rat pups.

Materials and Methods

Materials

Details of materials and methods are available in a previous publication [6], with the exception of the automated quantification of withdrawal signs, which is described in detail later in this section. Briefly, the National Institute on Drug Abuse (NIDA) Drug Supply Program (Bethesda, MD) provided norbuprenorphine (NorBUP) in free-base form. Morphine sulfate salt pentahydrate (morphine) and naltrexone hydrochloride (naltrexone) were purchased from Sigma-Aldrich (St. Louis, MO). NorBUP and morphine were dissolved in a vehicle consisting of dimethyl sulfoxide (DMSO; Fisher Scientific), polyethylene glycol-400 (PEG-400; Fisher Scientific), and sterile saline in a ratio of 1:2:1, respectively. Drugs were loaded into osmotic minipumps that continuously delivered 5 μL of contents per hour (0.120 mL/day) during the entirety of the dosing period (from gestational day 9 until delivery and euthanasia; Model 2ML2; Alzet, Cupertino, CA). Loaded minipumps were primed overnight in sterile saline at 37°C prior to subcutaneous implantation into pregnant rats. Naltrexone was dissolved in sterile saline prior to use in precipitated withdrawal testing.

Animal Care and Use

This study was performed in accordance with the Guide for Care and Use of Laboratory Animals. The University of Arkansas for Medical Sciences Institutional Care and Use Committee approved all uses of animals described in this study.

Thirty-six timed-pregnant Long-Evans rats (Charles River Laboratories, Wilmington, MA) arrived at the UAMS Division of Laboratory Animal Medicine on approximately gestational day (GD) 6 and were housed singly with environmental enrichment (e.g., cotton square nestlets and crinkled paper) and *ad libitum* access to food and water. On GD 9, primed osmotic minipumps containing vehicle or the mu opioid receptor full agonists morphine or NorBUP were subcutaneously implanted as instructed by the manufacturer to deliver the following: vehicle controls, 1:2:1 DMSO:PEG-400:saline, 0.120 mL/day (n = 8); the positive control morphine, 15 mg/kg/day (n = 3) or 20 mg/kg/day (n = 5); NorBUP, 0.3 mg/kg/day (n = 4), 1.0 mg/kg/day (n = 7), 3.0 mg/kg/day (n = 5), or 10 mg/kg/day (n = 4). Assignments were made by pseudo-randomization across 8 cohorts of dams (n = 36 dams in total) that were tested over the course of approximately 1 year (Supplemental Table 1). Other than cage changes, daily health checks, and weight measurements taken on weekdays, dams were left undisturbed for the remaining gestational period. Neonatal rat pups were separated

from dams and tested for opioid dependence between two and twelve hours after delivery as described in "Precipitated withdrawal testing" below. Dams were euthanized within 24 hours of pup testing.

Precipitated withdrawal testing

Precipitated withdrawal testing was adapted from previous reports[8,9,11]. Two to twelve hours after delivery, neonatal rat pups were separated from dams and tested for opioid dependence. For the first 15 litters (Supplemental Table 1), approximately half of each litter was randomized to be fostered to an untreated dam for a separate preliminary study examining dependence and withdrawal on PND 7. Of the remaining pups in each litter, at least one pup of each sex was randomly assigned to be challenged by intraperitoneal injection of either saline (10 mL/kg) or the opioid receptor antagonist naltrexone (1 or 10 mg/kg; 10 mL/kg) to precipitate opioid withdrawal. Pups were immediately videoed for 10 minutes after injection. Videos were processed and movements were scored both manually (to determine GWS; see "Manual scoring of withdrawal signs" section below) and by Ethovision® XT automated software (to determine Distance Moved and Movement Duration; See "Automated quantification of withdrawal" section below).

A total of 235 neonatal rat pups were subjected to precipitated withdrawal tests. For the present study, only pups that were exposed to vehicle (1:2:1 DMSO:PEG-400:saline, 0.120 mL/day), morphine (15 mg/kg/day), or NorBUP (0.3, 1.0, or 3.0 mg/kg/day) were included; pups exposed to morphine (20 mg/kg/day; n = 5) or NorBUP (10 mg/kg/day; n = 4) were excluded due to high rates of spontaneous abortions at these doses (40 and 75%, respectively)[6]. No spontaneous abortions were detected in the groups included in the present study. Additionally, one vehicle-exposed litter was excluded from the present study due to early use of non-standard testing chambers. Accordingly, 208 pups from 26 litters were used in the present study. See Table 1 for treatment group assignments.

For the three endpoints (GWS, Distance Moved, and Movement Duration), scores determined from same-sex littermates that received the same postnatal challenge were averaged to determine a single litter score for the sex. These scores were used to assess sex differences in our previous study[6]. Since sex-based scores were not significantly different[6], the male score and female score for each litter were averaged to determine a single sex-balanced score for each litter. This score was used in the grouped analyses of the previous and present studies[6]. Litters lacking a score for either sex were not included in these grouped analyses. Table 2 displays the number of litters per prenatal treatment and postnatal challenge groups with sex-balanced single scores for each endpoint.

Manual scoring of withdrawal signs

Manual scoring was adapted from an earlier study published by Jones and Barr[8]. Videos of the 10-minute testing sessions were edited using Shotcut video editing software (version 18.06.02; Meltytech, LLC) such that the first 10 seconds of each 30-second segment was retained and the remaining 20 seconds was discarded, resulting in a total of twenty 10-second segments per video. This editing allowed seamless "sampling" of pup's behaviors across the 10-minute testing period[8]. Graders that were blinded to pup sex and treatment

conditions recorded the presence (scored "1") or absence (scored "0") of each of the following 10 signs for each 10-second segment: body curls, body stretches, tremor, face wiping, foreleg movement, hindleg movement, head movement, wall climbing, locomotion, and quiet behavior (Table 2). All signs except quiet behavior are positive withdrawal signs that indicate presence of withdrawal. Quiet behavior is a negative withdrawal sign (i.e., absence of withdrawal) that is defined as "sedated appearance with no movement for at least 50% of the 10-second sampling time." Scores for quiet behavior were operationalized as positive by subtracting them from the maximum possible quiet behavior score of 20. Quiet behavior is included in the analysis to provide a finer gradation of withdrawal, particularly for pups that exhibited one or more of the positive signs during the sample but were motionless most of the sample. Scores for each of the 10 withdrawal signs were totaled for the twenty segments; therefore, the score for each sign ranged from 0 to 20. These scores were used to calculate the GWS for each pup by summing the scores for all positive withdrawal signs plus the quiet behavior score operationalized as a positive withdrawal score, making 200 the maximum possible GWS.

Automated quantification of withdrawal

Noldus Ethovision® XT version 8 was used to quantify neonatal pup locomotor activity in the videos taken during precipitated withdrawal testing on PND 0. The investigators using Ethovision® XT were blinded to each pup's treatment, sex, and GWS. Each video was edited using Shotcut video editing software to remove any footage extraneous to the 10-minute testing session. In contrast to the editing applied to the videos for manual grading, testing sessions were not edited into 10-second samples for Ethovision® XT analysis but instead were analyzed by the software as continuous 10-minute sessions. WM Converter Video/Audio software (version 5.1; All Alex, Inc) was used to convert edited videos from MP4 to MPEG1 format at a resolution of 1280 X 720 pixels. To quantify movement of the pups within the edited videos, the Ethovision® XT pre-defined template experiment for "Rodents- other" was selected and the sample rate was set at 5 samples per second. Initially, all possible outcome measurements were selected; subsequent analyses were carried out with "Distance Moved (cm)" and "Movement Duration (from center-point/moving, seconds)" only. Distance Moved is defined as the distance traveled by the center of the subject from one sample to the next sample (i.e., change in position every 0.2 seconds). Movement Duration gives the total amount of time the animal moved. These endpoints were selected because they provide continuous, positive measurements of general whole-body movement, which are easier to interpret than endpoints that were binary or provided negative measures of movement (e.g., time not moving). Thresholds for start and stop velocities for Movement Duration were set at 0.4 and 0.3 cm/sec, respectively, meaning that average velocity of movements exceeding 0.4 cm/sec triggered Ethovision® XT to begin measuring Movement Duration, and slowing below 0.3 cm/sec stopped the measuring of Movement Duration. Arena settings were calibrated for each test chamber (10.16 X 12.86 cm) (Figure 1). Under Detection Settings, sampling rate was set at 5 samples per second. "Gray scaling" was selected as the detection method, and detection settings were optimized for each video separately. The detection range was adjusted for each video such that pups were entirely selected (i.e., shaded in yellow) with minimal noise (shaded in orange) (Figure 2). Playback control and the Video Time slider were used to view the video with the selected detection

settings and ensure that pups were selected with minimal noise throughout the entire video; if they were not, the detection range and contour dilation were adjusted to improve detection as much as possible before beginning data acquisition. If any part of a pup moved into shadows such that its movements were not detected for more than 5 seconds, the pup was excluded (see Figures 2–3 for examples of shadows in testing chamber). To assess interrater reliability of choosing detection settings, two raters independently analyzed the videos using Ethovision® XT software. Additionally, in an effort to increase the number of pups that could be included in the analyses, Rater 2 attempted to analyze pups that Rater 1 excluded for moving into the shadows for more than 5 seconds by including pups that spent no more than 10 seconds in shadows.

Quality control was performed at the conclusion of data acquisition to identify any potential samples that were detected incorrectly (i.e., shadows or changes in lighting detected as movement). Under Track Editor, auto-select settings for sample distance were set to 1.0 cm. This setting identified 0.2-second samples in which suspiciously large movements were detected (i.e., 1 cm/sample). Sections containing these suspicious samples were viewed using the integrated visualization function (Figure 3), and 0.2-second samples were omitted if movements in the sample were determined to be "false positives," or due to noise or artifacts. Some samples were also "missed" by the analysis due to high processor load. The number of "missed samples" are reported in combination with the number of samples omitted by raters during quality control as "percent missed samples." Exclusions were not made on the basis of percent missed samples.

Statistical analyses

All statistical analyses were performed using GraphPad Prism Version 8.2. Fisher's exact test was used to assess differences between included and excluded pups. Linear regression analyses were performed with data from individual pups to relate GWS with Distance Moved and Movement Duration. Pearson's correlation coefficient was determined to assess interrater reliability for Distance Moved and Movement Duration. Sensitivity and specificity of each endpoint to detect withdrawal was assessed using receiver operator characteristic (ROC) curve analyses. In these analyses, saline-challenged pups whose dams were treated with morphine (15 mg/kg/day) during gestation were classified as having low-grade opioid dependence and withdrawal, and saline-challenged pups whose dams were treated with vehicle were classified as not opioid dependent. These groups were selected based on findings in our previous study showing that saline-challenged morphine-exposed pups ("morphine + saline") had slightly, but insignificantly, elevated GWS relative to saline-challenged vehicle-exposed pups ("vehicle + saline")[6]. Because withdrawal signs were quite subtle in the morphine + saline pups, the sensitivity and specificity of a test to distinguish these pups from vehicle + saline pups demonstrates the value of the test for quantifying withdrawal. "Sensitivity" refers to the proportion of all morphine-exposed pups with a test score higher than an investigator-defined cutoff value ("true positive"), and "specificity" refers to the proportion of all vehicle-exposed pups with a test score lower than an investigator-defined cutoff ("true negative"). The ROC curve analysis calculates and plots sensitivity (y-axis) and 1-specificity (x-axis) at multiple cutoffs representing the range of potential test values. The area under the ROC curve (AUC) plot is summed to calculate a

single summary statistic that characterizes the performance of the test. An AUC value approximately 0.5 indicates that the test performs no better than chance; AUC values approaching 1 (the maximum score) indicate higher sensitivity and specificity.

For all other analyses, the litter was the statistical unit of analysis. Prenatal treatment was evaluated at 5 levels [vehicle, morphine (15 mg/kg/day), or NorBUP (0.3, 1.0, or 3.0 mg/kg/day)] and postnatal challenge was evaluated at 3 levels [saline or naltrexone (1 or 10 mg/kg)], yielding a total of 15 treatment groups (see Table 1 for sample sizes per treatment group). Main and interaction effects of prenatal treatment and postnatal challenge on GWS, Distance Moved, and Movement Duration were determined using two-way analysis of variance (ANOVA). Dunnett's multiple comparisons tests were conducted within each endpoint to determine if treatment group means differed from the mean of a control group that was prenatally exposed to vehicle and postnatally challenged with saline ("vehicle + saline"). All analyses were conducted at the 95% confidence level. Data are graphed as mean ± standard error of mean for each treatment group.

Results

Scores obtained by Raters 1 and 2 were highly correlated (Distance Moved: r = .8759, p < .0001; Movement Duration: r = .9239, p < .0001), indicating high interrater reliability. Movements of 166 pups (79.8 %) were successfully quantified by Rater 1 using Ethovision® XT; 42 pups (20.2%) were excluded from further analysis due to the presence of shadows in the arenas that interfered with testing. No prenatal treatment, postnatal challenge, or sex was overrepresented in the exclusions; however, pups of dams that were treated with the lowest dose of NorBUP (0.3 mg/kg/day) were underrepresented in the exclusions. Six percent (n = 2) of this group was excluded, whereas 23% (n = 40) of pups exposed to any other prenatal treatment were excluded (p = 0.02 by Fisher's exact test). Using less stringent exclusion criteria, Rater 2 successfully quantified movements of 179 pups (86.1%); 29 pups (13.9%) were excluded. The mean percent missed samples by Rater 1 and Rater 2 were 8.1% (95% confidence interval: 7.3 – 8.8%) and 0.08% (95% confidence interval: 0.05 – 0.10%), respectively. The interrater discrepancy in percent missed samples was likely due to differences in processor capability of the computers used by Raters 1 and 2.

To determine if Ethovision® XT can be used to quantify opioid withdrawal in the neonatal rat, we first assessed whether there was a linear relationship between manual GWS and automated endpoints obtained from Ethovision® XT. Across a wide range of GWS, Rater 1 determined there was a strong positive linear relationship between GWS and Distance Moved (y = 4.86x + 26.83; p < .0001; R^2 = .7691) and between GWS and Movement Duration (y = 5.36x − 12.68; p < .0001; R^2 = .8218) (Figure 4a). Rater 2 determined similar results (Distance Moved: y = 4.37x + 27.87; p < 0.0001; R^2 = .7357; Movement Duration: y = 4.62x − 9.66; p < .0001; R^2 = .7721) (Figure 4b).

We next calculated the sensitivity and specificity of each endpoint to detect low-grade opioid withdrawal associated with prenatal exposure to 15 mg/kg/day morphine, a positive control for NOWS in rat models[8,14]. We plotted the receiver operator characteristic (ROC) curve for each endpoint (Figure 5) and quantified the area under each ROC curve. The maximum

possible value for AUC is 1; AUCs close to 1 have high sensitivity and specificity. For Rater 1, AUC for Distance Moved was 0.8875 (95% CI: 0.7588 – 1.000) and for Movement Duration was 0.8938 (95% CI: 0.7718 – 1.000); for Rater 2, AUC for Distance Moved was 0.8693 (95% CI: 0.7326 – 1.000) and for Movement Duration was 0.8693 (95% CI: 0.7350 – 1.000). AUC for GWS was 0.7824 (95% CI: 0.5889 – 0.9759) for pups included in Rater 1's analyses of Distance Moved and Movement Duration (n = 16, vehicle-exposed pups from 7 litters; n = 10, morphine-exposed pups from 3 litters), and was 0.7955 (95% CI: 0.6125 – 0.9784) for pups included in Rater 2's analyses of Distance Moved and Movement Duration (n = 16, vehicle-exposed pups from 7 litters; n = 11, morphine-exposed pups from 3 litters). These results indicate that all three endpoints sensitively and specifically discriminated low-grade opioid withdrawal from normal pup behavior, and that no endpoint clearly performed better than the others.

We used manual scoring in an earlier study to determine if prenatal treatment with the buprenorphine metabolite norbuprenorphine (NorBUP) caused dependence and withdrawal in neonatal pups [6]. In this report, a two-way ANOVA assessed the main and interaction effects of prenatal treatment and postnatal challenge on GWS, and Dunnett's multiple comparisons tests identified treatment groups with means that differed from that of the vehicle + saline control group. In the present study, we replicated these analyses with Distance Moved and Movement Duration as dependent variables to determine if these endpoints would yield similar results and conclusions as GWS obtained by manual scoring. Results of these analyses are shown in Figure 6 and Table 4. The previously reported data of GWS from all litters with a single sex-balance GWS are shown in Figure 6A with permission from the Journal of Pharmacology and Experimental Therapeutics[6]. This original analysis included pups that were excluded from the Ethovision® XT analyses in the present study due to shadows in the testing arenas. Because we wanted to compare the three endpoints in the same pups, we repeated the analyses of GWS, but included only the pups that were also included in the Ethovision® XT analysis by each respective rater (see Figures 6B and 6E). Similar results were obtained for GWS, Distance Moved, and Movement Duration for Raters 1 and 2 (Table 4, Figure 6). The main effects of prenatal treatment and postnatal challenge were significant (p <.0001) for all three endpoints and for both raters (Table 3). The interaction effect met criteria for significance (p < .05) in all analyses of GWS, for Distance Moved by Rater 1 and Movement Duration for Rater 2. The p-values for the interaction effects for Distance Moved (p = .0503) and Movement Duration (p = .0729) as determined by Rater 1 and Rater 2, respectively, were slightly greater than criteria for significance. In each model, the main and interaction effects of prenatal treatment and postnatal challenge combined explained a large majority of the variance in GWS (original: 77.60%; Rater 1: 91.36%; Rater 2: 82.65%), Movement Duration (Rater 1: 87.57%; Rater 2: 84.35%), and Distance Moved (Rater 1: 88.31%; Rater 2: 82.18%), which provides further evidence that all endpoints quantify withdrawal well. Means of the treatment groups showed highly similar patterns for each endpoint (Figure 6); pups exposed to morphine or high doses of NorBUP (1 or 3 mg/kg/day) and naltrexone (1 or 10 mg/kg) tended to exhibit higher scores for GWS, Distance Moved, and Movement Duration relative to the vehicle + saline group. Pups exposed to vehicle or a low dose of NorBUP (0.3 mg/kg/day) tended to exhibit no change relative to the vehicle + saline group for any endpoint. However, there were

discrepancies between the endpoints for the group that was exposed to 3 mg/kg/day NorBUP and challenged with the lower dose of naltrexone in Rater 1's analysis. Unlike manual grading, Ethovision® XT-derived scores for this group were not significantly greater than scores for the vehicle + saline group (Figure 6B-6D). However, this discrepancy was not present in Rater 2's analysis (Figure 6E-6G). Additionally, saline-challenged pups with prenatal exposure to 3 mg/kg/day NorBUP had GWS that were larger than vehicle + saline controls (Figure 6A), but this effect was diminished by excluding pups that were excluded from the Ethovision® XT analyses (Figure 6B and 6E), and was not determined by Rater 2 for Distance Moved (Figure 6C) and Movement Duration (Figure 6D), or by Rater 1 for Distance Moved (Figure 6F). The exclusions made in the automated quantification analyses (Table 1), which resulted in smaller group sizes for the Ethovision® XT endpoints, are a potential and likely source of these discrepancies. However, despite these exclusions and smaller group sizes, results of all analyses are strikingly similar and lead to the conclusion that prenatal morphine (15 mg/kg/day) and NorBUP (1 or 3 mg/kg/day) and postnatal challenge with naltrexone (1 or 10 mg/kg) increased GWS, Distance Moved, and Movement Duration.

Discussion

The purpose of the present study was to improve the quantification of opioid withdrawal in a rat model of NOWS. Previous studies evaluated withdrawal using manual grading, which involves observing and quantifying specific withdrawal signs, including curls, stretches, tremor, foreleg and hindleg movements[6,8]. This method is time consuming, labor intensive, and subjective in nature; variability can be introduced to the experiment due to grader fatigue and interrater differences. In contrast, the automated method validated in this study objectively scores total locomotor activity, a proxy for neonatal opioid withdrawal, in neonatal rat pups using the widely-used animal behavioral software Ethovision® XT[15].

The present study demonstrates that Ethovision® XT can reliably quantify opioid withdrawal in rat pups on PND 0. Strong linear relationships between GWS obtained manually and two automated endpoints acquired using Ethovision® XT, Distance Moved and Movement Duration, suggest that these endpoints and GWS are directly proportional and support the use of Ethovision® XT as a proxy for quantifying withdrawal. The high degree of interrater concordance infer that the automated method is rigorous, that scores can be reproduced reliably, and that one rater is sufficient for Ethovision® XT analysis (in contrast to manual scoring, which requires two blinded raters). Because prenatal opioid exposure and postnatal challenge with an opioid antagonist are known to cause withdrawal in fetal [8] and neonatal rat pups [6,16,17], our finding that prenatal exposure and postnatal challenge similarly predicted GWS, Distance Moved, and Movement Duration provides further evidence that the automated methods provide valid measurements of withdrawal. Furthermore, these results were obtained with fewer degrees of freedom using Ethovision® XT, suggesting that the new automated method may improve power and require fewer animals. As such, in addition to streamlining the analysis of withdrawal in neonatal rats, the automated method may further reduce costs associated with conducting the experiment (e.g., animal care and use, drugs, smaller sample size, etc). This improvement is likely due to differences in the resolution of data acquisition of each method. The automated method

captures continuous measures of locomotor activity by quantifying movements greater than 0.01 cm in 5 samples per second across the entire ten-minute testing session. In contrast, the manual grading method captures the presence or absence of ten discreet withdrawal signs only once per 10-second segment for twenty segments sampled over the 10-minute testing session. This new automated method gathers more data on locomotor activity relative to the manual grading method, enabling the automated method to elucidate even subtle differences in withdrawal.

To our knowledge, the present study is the first to determine the sensitivity and specificity of any method used to quantify opioid withdrawal in a rat model of NOWS. Sensitivity and specificity are important measures for defining the ability of a method to accurately discriminate between presence and absence of opioid withdrawal. The excellent performances of Distance Moved, Movement Duration, and GWS provide assurance that these measurements are trustworthy for quantifying withdrawal signs in neonatal rat pups. Care should be taken, however, when interpreting the results of these analyses. The most accurate, stringent interpretation is that Distance Moved, Movement Duration, and GWS have high sensitivity and specificity to detect opioid withdrawal equal to those caused by *in utero* exposure to 15 mg/kg/day morphine from GD 9 until delivery. The performance of Distance Moved and Movement Duration appear to be slightly better than GWS; however, the 95% confidence intervals for the three methods overlap, indicating no clear difference in performance.

Despite its advantages, the automated method has limitations. Twenty-nine of 208 pups were excluded from the Ethovision® XT analysis due to poor lighting conditions in the archived videos. Even small shadows in the testing chambers (Figure 3) can interfere with locomotor quantification. This can be especially problematic for pups in greater withdrawal, as they are more likely to move into shadows. Providing overhead lighting to eliminate shadows in the test chambers alleviates this problem. Although the present study examined performance of Ethovision® XT in archived videos, these methods can readily be applied to real-time acquisition by the Noldus system during withdrawal testing, further decreasing time- and labor-demands by eliminating the need to edit and convert videos prior to analysis by Ethovision® XT. One limitation of the Distance Moved endpoint is its ambiguity; it does not distinguish between a movement across the test chamber and moving in place, making this endpoint potentially difficult to interpret. For this reason, and because of its slightly better performance relative to Distance Moved, we prefer Movement Duration and believe it to be the superior endpoint. However, it should be noted that neither endpoint provides data on specific withdrawal signs (e.g., head movement, foreleg movement, etc.), which is a limitation of the method in cases that such a level of behavioral resolution is desired. One complication of neonatal opioid withdrawal testing on PND 0 in rat models is the lack of experimenter control over delivery times. While dams tended to deliver their litters during the afternoon, some deliveries occurred overnight when monitoring was difficult to accomplish without disrupting light-dark cycles. In these cases, the interval between birth and testing was uncertain, but no more than 12 hours. Having such a large range of post-birth testing intervals (2–12 hours) may affect results of withdrawal testing. However, in the present study, there were no differences in post-birth testing intervals among the treatment groups (Supplemental Figure 1), suggesting that this interval did not play a large role in the

results. Little is known about morphine and NorBUP pharmacokinetics in the neonatal rat or about the time course of withdrawal in the rat pup following prenatal exposure to morphine or NorBUP. As such, it is possible that pups with longer post-birth testing intervals experienced spontaneous withdrawal prior to the precipitated withdrawal testing, which could conceivably affect withdrawal during testing. Additionally, peak withdrawal caused by morphine and NorBUP may differ due to potential differences in pharmacokinetics. Another factor that should be considered is the possibility of effects due to stress experienced by dams early in gestation during shipment and re-housing. If the dams in our study indeed experienced stress that exacerbated dependence in their pups, then their pups could have exhibited higher levels of withdrawal than pups of unstressed dams (e.g., those bred in-house). Finally, the method described is valid for rat pups on PND 0. Because of the rapid changes in size and motor activity of neonatal rat pups during the first days of life, one should not assume this method can be used to quantify opioid withdrawal in rat pups beyond PND 0 or in species other than rat.

Conclusions

Studies of opioid withdrawal behavior in neonatal rat models suffer from the lack of an automated approach and require labor intensive and subjective characterization of several individual withdrawal signs. Here, we described development and validation of an improved method using Noldus EthoVision® XT software to quantify opioid withdrawal signs in neonatal rat pups on PND 0 that is rapid and exhibits high sensitivity and specificity for capturing movement parameters predictive of opioid withdrawal in neonates. This work will enable laboratories exploring withdrawal in neonatal rats to more rapidly progress toward finding more effective treatments for this condition. This work may also be adapted to facilitate research focused on other influencers of neonatal motor behavior, including non-opioid drugs that are associated with dependence and withdrawal, including selective serotonin reuptake inhibitors[18], benzodiazepines[19], and ethanol[20].

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments:

References

1. Winkelman TNA, Villapiano N, Kozhimannil KB, Davis MM, Patrick SW. Incidence and costs of neonatal abstinence syndrome among infants with medicaid: 2004–2014. Pediatrics. 2018;141(4):10.1542/peds.20173520. doi: e20173520 [pii].

2. Patrick SW, Schumacher RE, Benneyworth BD, Krans EE, McAllister JM, Davis MM. Neonatal abstinence syndrome and associated health care expenditures: United states, 2000–2009. JAMA. 2012;307(18):1934–1940. doi: 10.1001/jama.2012.3951 [doi]. [PubMed: 22546608]

3. Jansson LM, Velez M, Harrow C. The opioid-exposed newborn: Assessment and pharmacologic management. J Opioid Manag. 2009;5(1):47–55. [PubMed: 19344048]

4. Patrick SW, Davis MM, Lehman CU, Cooper WO. Increasing incidence and geographic distribution of neonatal abstinence syndrome: United states 2009 to 2012. J Perinatol. 2015. doi: 10.1038/jp.2015.36 [doi].

5. Barr GA, Zmitrovich A, Hamowy AS, Liu PY, Wang S, Hutchings DE. Neonatal withdrawal following pre- and postnatal exposure to methadone in the rat. Pharmacol Biochem Behav. 1998;60(1):97–104. doi: S0091-3057(97)00596-0 [pii]. [PubMed: 9610930]

6. Griffin BA, Caperton CO, Russell LN, et al. In utero exposure to norbuprenorphine, a major metabolite of buprenorphine, induces fetal opioid dependence and leads to neonatal opioid withdrawal syndrome. J Pharmacol Exp Ther. 2019. doi: jpet.118.254219 [pii].

7. Wallin CM, Bowen SE, Roberge CL, Richardson LM, Brummelte S. Gestational buprenorphine exposure: Effects on pregnancy, development, neonatal opioid withdrawal syndrome, and behavior in a translational rodent model. Drug Alcohol Depend. 2019;205:107625. doi: S0376-8716(19)30402-8 [pii].

8. Jones KL, Barr GA. Opiate withdrawal in the fetal rat: A behavioral profile. Pharmacol Biochem Behav. 2000;66(2):419–424. doi: S0091-3057(00)00209-4 [pii]. [PubMed: 10880699]

9. McPhie AA, Barr GA. The role of opioid receptors in morphine withdrawal in the infant rat. Brain Res Dev Brain Res. 2000;124(1–2):73–80. doi: S0165380600001024 [pii]. [PubMed: 11113514]

10. Zhu H, Barr GA. Naltrexone-precipitated morphine withdrawal in infant rat is attenuated by acute administration of NOS inhibitors but not NMDA receptor antagonists. Psychopharmacology (Berl). 2000;150(3):325–336. [PubMed: 10923761]

11. Little PJ, Price RR, Hinton RK, Kuhn CM. Role of noradrenergic hyperactivity in neonatal opiate abstinence. Drug Alcohol Depend. 1996;41(1):47–54. doi: 0376871696012367 [pii]. [PubMed: 8793309]

12. Ceger P, Kuhn CM. Opiate withdrawal in the neonatal rat: Relationship to duration of treatment and naloxone dose. Psychopharmacology (Berl). 2000;150(3):253–259. [PubMed: 10923752]

13. Byrnes EM, Vassoler FM. Modeling prenatal opioid exposure in animals: Current findings and future directions. Front Neuroendocrinol. 2018;51:1–13. doi: S0091-3022(17)30058-4 [pii]. [PubMed: 28965857]

14. Jones KL, Barr GA. Ontogeny of morphine withdrawal in the rat. Behav Neurosci. 1995;109(6):1189–1198. [PubMed: 8748967]

15. Noldus LP, Spink AJ, Tegelenbosch RA. EthoVision: A versatile video tracking system for automation of behavioral experiments. Behav Res Methods Instrum Comput. 2001;33(3):398–414. doi: 10.3758/bf03195394 [doi]. [PubMed: 11591072]

16. Tao PL, Yeh GC, Su CH, Wu YH. Co-administration of dextromethorphan during pregnancy and throughout lactation significantly decreases the adverse effects associated with chronic morphine administration in rat offspring. Life Sci. 2001;69(20):2439–2450. doi: S0024-3205(01)01316-9 [pii]. [PubMed: 11681630]

17. Robinson SE, Wallace MJ. Effect of perinatal buprenorphine exposure on development in the rat. J Pharmacol Exp Ther. 2001;298(2):797–804. [PubMed: 11454944]

18. Koren G, Nordeng H. Antidepressant use during pregnancy: The benefit-risk ratio. Am J Obstet Gynecol. 2012;207(3):157–163. doi: 10.1016/j.ajog.2012.02.009 [doi]. [PubMed: 22425404]

19. Iqbal MM, Sobhan T, Ryals T. Effects of commonly used benzodiazepines on the fetus, the neonate, and the nursing infant. Psychiatr Serv. 2002;53(1):39–49. doi: 10.1176/appi.ps.53.1.39 [doi]. [PubMed: 11773648]

20. Robe LB, Gromisch DS, Iosub S. Symptoms of neonatal ethanol withdrawal. Curr Alcohol. 1981;8:485–493. [PubMed: 7343193]

Highlights

- First to quantify neonatal opioid withdrawal in rats with Ethovision XT®.

- Results from Ethovision XT® and manual scoring are highly correlated.

- Ethovision XT® method has high interrater reliability.

- Ethovision XT® method has high sensitivity and specificity for withdrawal.
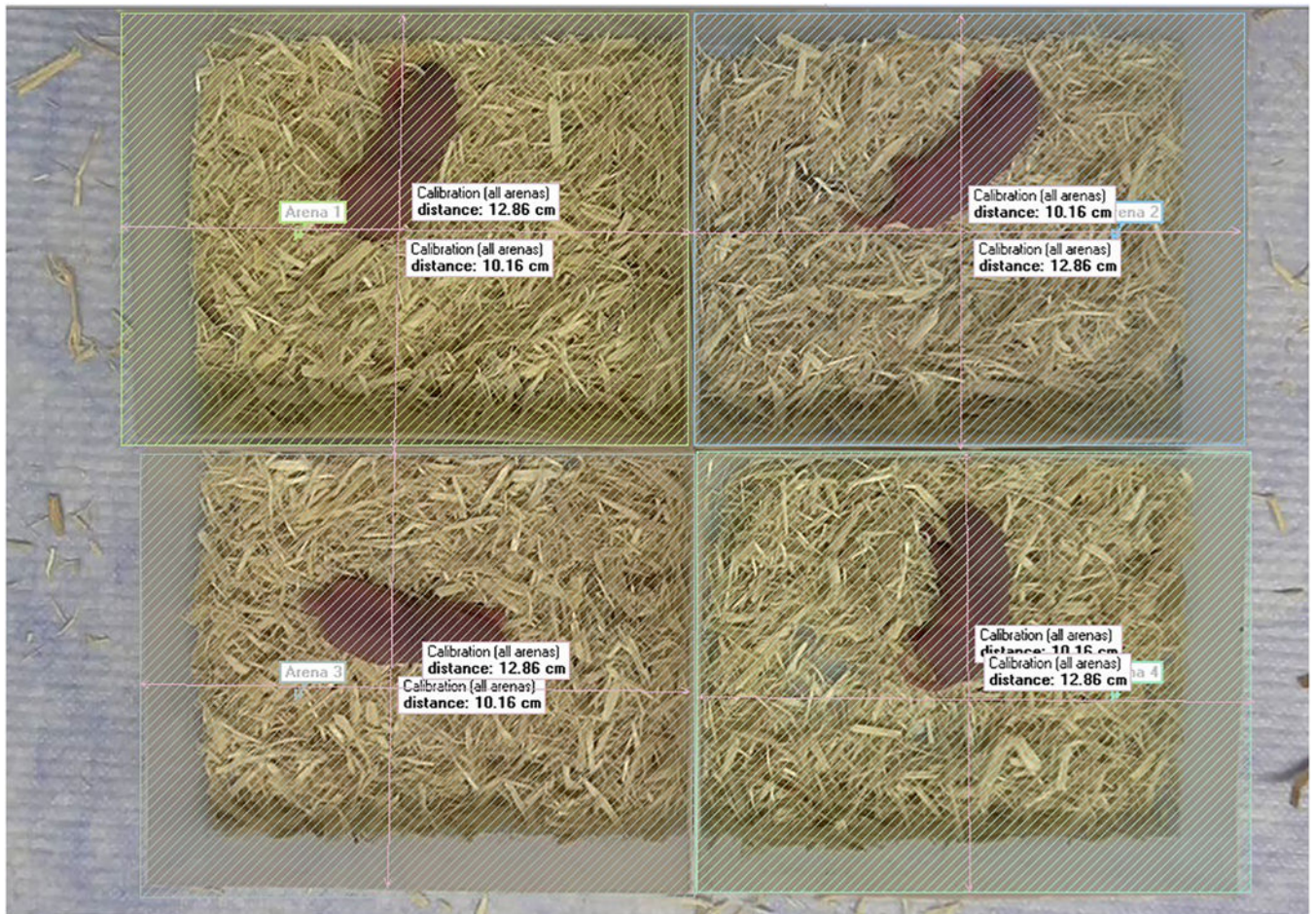
Figure 1. Demarcation and calibration of test arenas.
Using options under the Arena Setting menu, the test chambers (noted as "arenas" in the figure) were demarcated as shown (slanted lines). The horizontal and vertical lines indicate the length and width of each test chamber and were used to calibrate the image to the actual size of each chamber (12.86 X 10.16 cm).
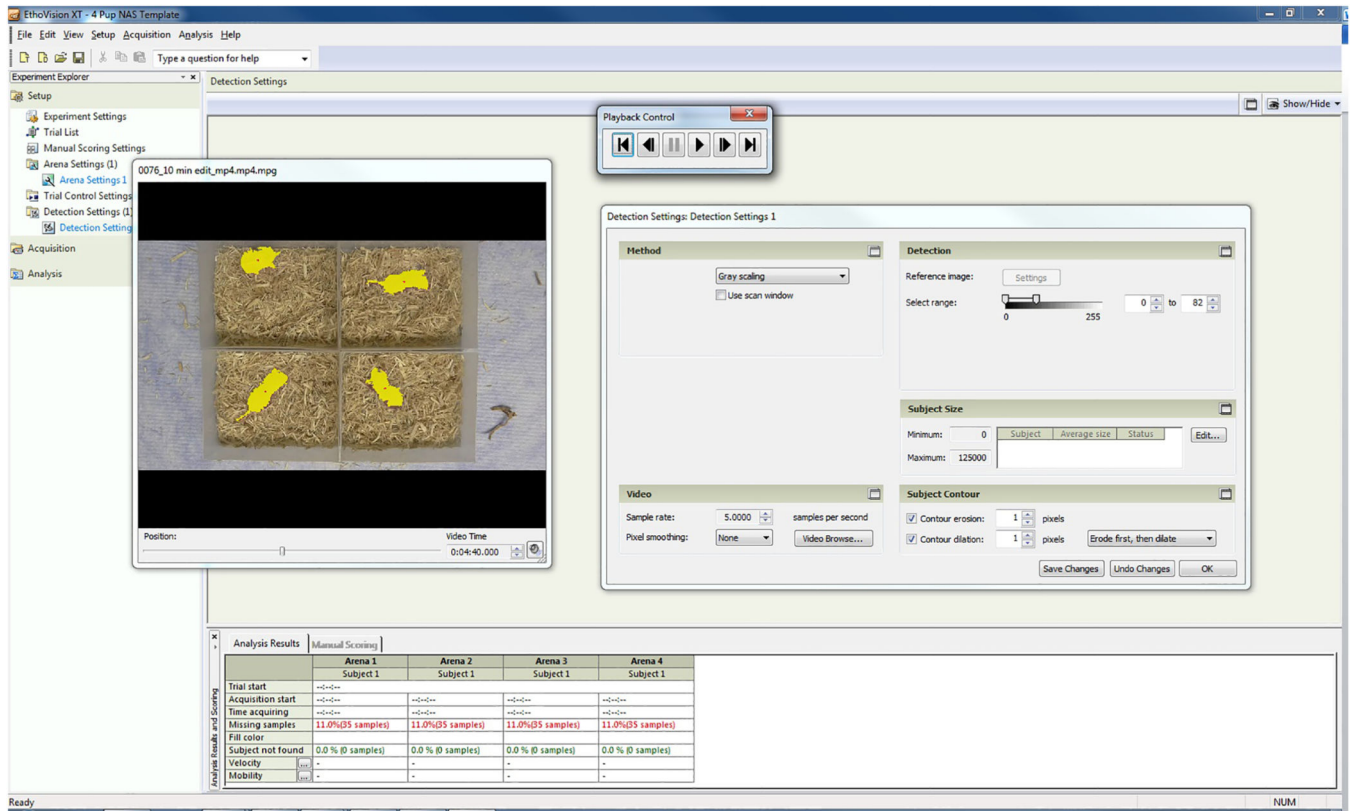
Figure 2. Example of detection settings used to quantify pup movements.

The gray scaling range was adjusted using the slider until pups were completely covered with yellow shading with little or no orange shading in the test arenas. Red dots represent the center point of each pup.
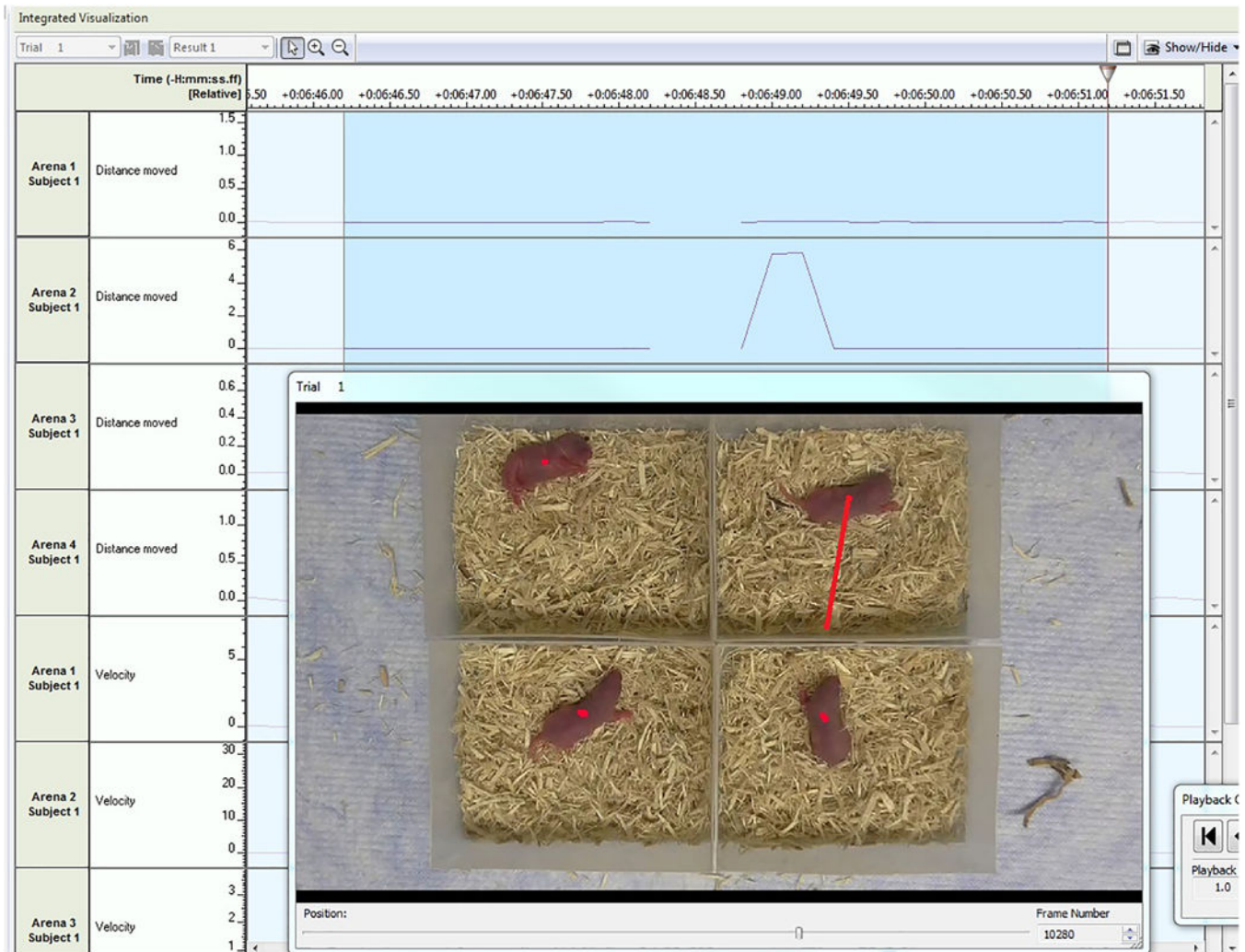
Figure 3. Quality control of inappropriately detected movements (i.e., false positives and artifacts).

In the frame shown, movements greater than 1 cm per sample were detected in arena 2 and identified using auto-select settings. The red line in the video frame extending from the pup's center point to shadowed area of the test arena indicates an inappropriate detection caused by shadow artifacts. This 0.2 second sample, and others like it, was omitted.
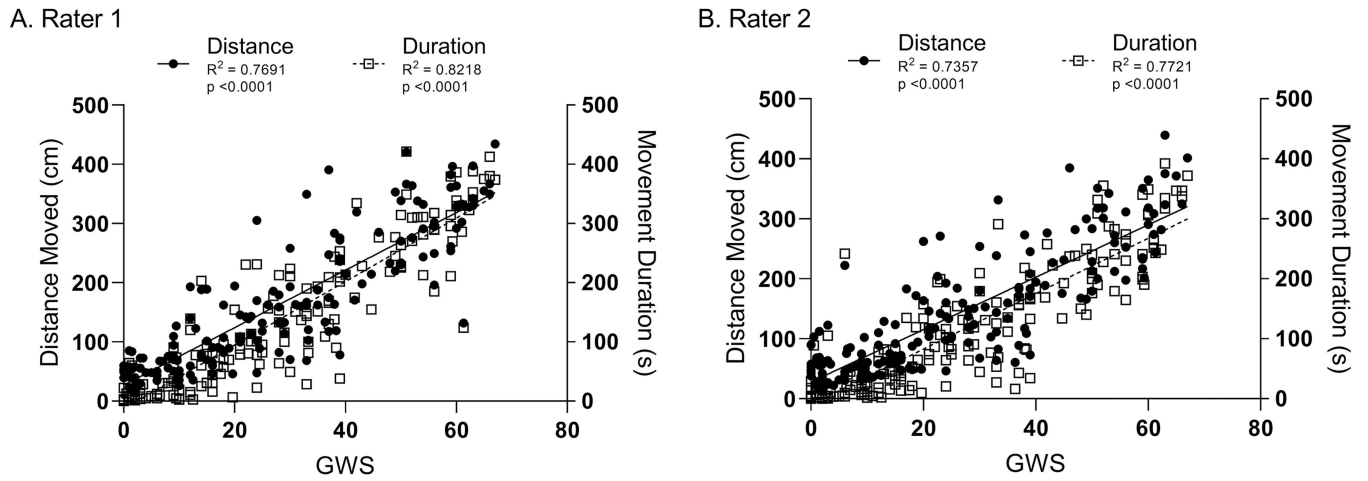
Figure 4. Distance Moved and Movement Duration have strong positive linear relationships with global withdrawal score (GWS).

Symbols represent individual pup's scores for Distance Moved (filled circles; left y-axis) and Movement Duration (open squares; right y-axis) as functions of global withdrawal score (GWS; x-axis), as measured by Rater 1 (A) and Rater 2 (B). Solid and dashed lines illustrate the linear regression of Distance and Movement, respectively, as functions of global withdrawal score.
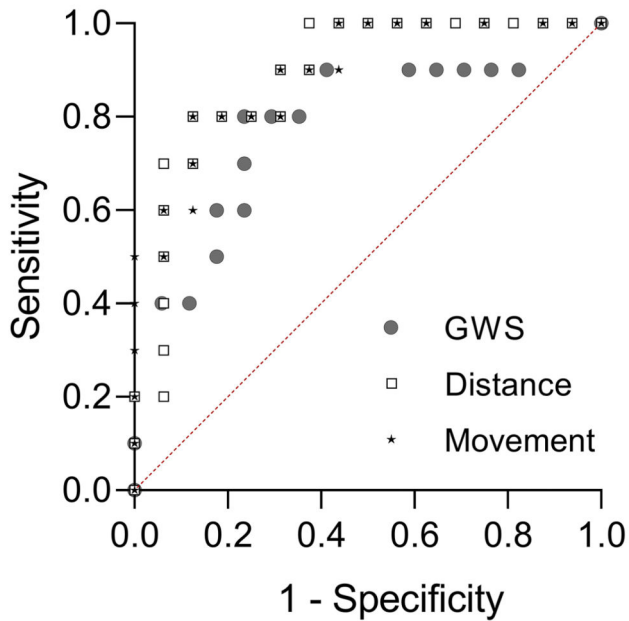
Figure 5. Receiver operator characteristic (ROC) curves indicate Distance Moved and Movement Duration clearly discriminate prenatal morphine and vehicle exposures.
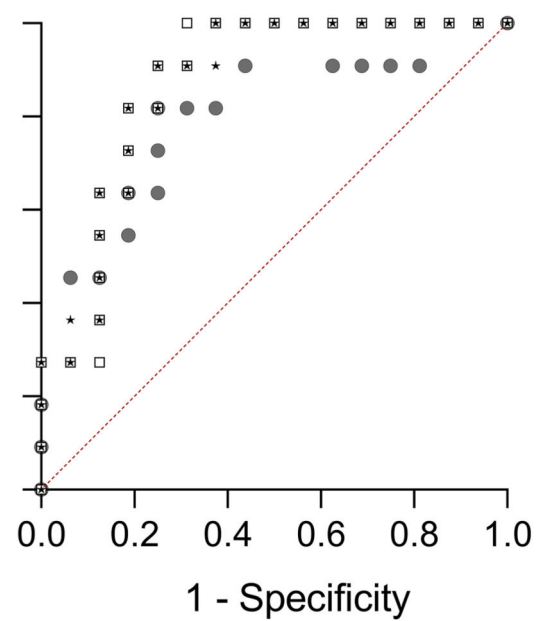Data points represent sensitivity (true positive rate, y-axis) plotted against 1-specificity (false positive rate, x-axis) of GWS (grey circles), Distance Moved ("Distance," open squares), and Movement Duration ("Movement," stars) to discriminate morphine-exposed pups (n = 10 for Rater 1 and n = 11 for Rater 2) and vehicle-exposed pups (n = 16) at multiple cutoff values for each endpoint. Values were determined by Rater 1 (A) and Rater 2 (B). Cutoff values were automatically determined by the GraphPad Prism function for ROC curve analysis. Dotted red line provides a reference of curves that are no better than chance at discriminating morphine and vehicle exposure. All ROC curves shown have p < .0001 for discriminating morphine and vehicle exposure.
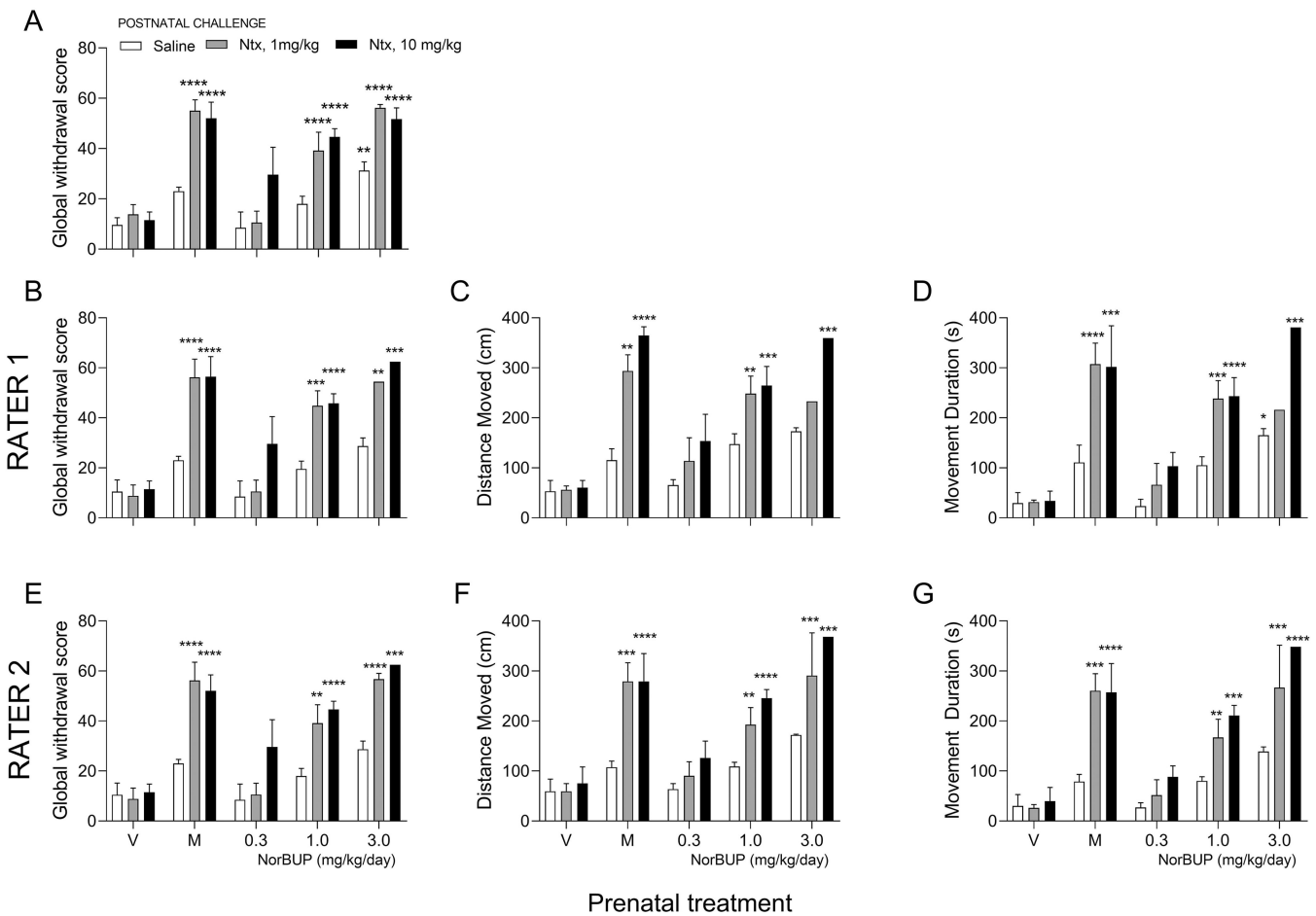
Figure 6. Manual grading and automated quantification methods closely agree in analyses of group means.

Panels represent analyses performed using GWS (A, B, E), Distance Moved (C, F), and Movement Duration (D, G) as endpoints. Scores for these endpoints are plotted on the y-axes, and prenatal treatment with vehicle (V), morphine (M, 15 mg/kg/day), or NorBUP (0.3, 1.0, or 3.0 mg/kg/day) is plotted on the x-axis. Figure (A) represents global withdrawal scores of all pups (previous published and used here with permission from the Journal of Pharmacology and Experiment Therapeutics), while Figures (B) and (E) include only pups that were also included in the Ethovision® XT analyses by Raters 1 and 2, respectively. Bars and error bars represent treatment group means and standard errors, respectively, following challenge with saline (open bars), 1 mg/kg naltrexone (gray bars), or 10 mg/kg naltrexone (filled bars). Asterisks indicate statistically significant difference versus control group (vehicle-treated, saline-challenged) at the following levels: *p < .05, **p < .01, ***p < 0.001, ****p < 0.0001.

Table 1.

Treatment group membership of pups included in present study

Prenatal treatment	Postnatal challenge	Total pups	Males	Females	Litters represented
Vehicle	Saline	22	11	11	7
	Ntx1	17	9	8	6
	Ntx10	14	6	8	5
Morphine	Saline	11	5	6	3
	Ntx1	8	4	4	3
	Ntx10	8	4	4	3
NorBUP, 0.3 mg/kg/day	Saline	14	8	6	4
	Ntx1	14	8	6	4
	Ntx10	7	4	3	2
NorBUP, 1.0 mg/kg/day	Saline	21	9	12	6
	Ntx1	19	8	11	7
	Ntx10	18	9	9	6
NorBUP, 3.0 mg/kg/day	Saline	13	6	7	5
	Ntx1	11	6	5	4
	Ntx10	11	5	6	5
TOTALS		208	102	106	26[*]

S, saline; Ntx1, naltrexone (1 mg/kg); Ntx10, naltrexone (10 mg/kg); NorBUP, norbuprenorphine.

[*] not the sum of the "Litters represented" column because pups within each litter were randomized to multiple postnatal challenges. Within prenatal treatments, unequal numbers of litters across postnatal challenges were due to not having enough pups per litter per sex to test all three postnatal challenges following division of some litters for a separate study. For other litters, multiple pups of each sex were tested for each postnatal challenge. For group analyses, scores for littermates receiving the same postnatal challenge were averaged such that no litter or no sex was overrepresented (see "Materials and Methods").

Table 2.

Number of litters with sex-balanced single scores by treatments, endpoint, and rater

Prenatal treatment	Postnatal challenge	Rater 1			Rater 2	
		GWS, n	Ethovision® XT, n	Exclusions, n	Ethovision® XT, n	Exclusions, n
Vehicle	Saline	7	4	3	4	3
	Ntx1	5	3	2	3	2
	Ntx10	4	4	0	4	0
Morphine	Saline	3	3	0	3	0
	Ntx1	3	2	1	2	1
	Ntx10	3	2	1	3	0
NorBUP, 0.3 mg/kg/day	Saline	4	4	0	4	0
	Ntx1	4	4	0	4	0
	Ntx10	2	2	0	2	0
NorBUP, 1.0 mg/kg/day	Saline	6	5	1	6	0
	Ntx1	5	4	1	5	0
	Ntx10	6	5	1	6	0
NorBUP, 3.0 mg/kg/day	Saline	4	3	1	3	1
	Ntx1	4	1	3	2	2
	Ntx10	5	1	4	1	4

Within each litter, scores from all same-sex littermates were averaged to provide a score for each sex. Male and female scores from each litter (where available) were then averaged to provide the litter's sex-balanced single score for global withdrawal (GWS) and for the Ethovision® XT endpoints. For purposes of comparing results of the Ethovision® XT analysis with the previous results (Figure 6), the table displays the number of successful analyses per treatment group. Ntx1, naltrexone 1 mg/kg; Ntx10, naltrexone 10 mg/kg

Table 3.

Behavioral definitions of opioid withdrawal signs in the neonatal rat

Behavior	Definition
Body curl	Ventral or lateral flexion of trunk
Body stretch	Extension or dorsal flexion of trunk causing apparent lengthening of body
Body tremor	Mild lateral movements of head that progress to a full-body lateral tremor
Face wiping	Wiping of both forelimbs across the face
Foreleg movements	Flexion, extension, or rotation of one or both forelegs
Hindleg movements	Flexion or extension of one or both hindlegs
Head movements	Ventral, dorsal or lateral rotary motion of the head
Wall climbing	Placing at least two forepaws on the wall of the observation cage and moving them up and down
Locomotion	Walking across the cage
Quiet	Sedated appearance with no movement for at least 50% of the sample

Used with permission from the Journal of Pharmacology and Experimental Therapeutics[6]

Table 4.

Results of two-way ANOVA of prenatal treatment and postnatal challenge on each endpoint

Sign	Predictor	*F*-Value	*P* Value	% of Total Variation
GWS (original)	Interaction	F (8, 50) = 3.051	.0071[a]	8.721
	Prenatal	F (4, 50) = 34.81	< 0001[a]	49.74
	Postnatal	F (2, 50) = 26.78	<.0001[a]	19.14
				77.60
RATER 1				
GWS	Interaction	F (8, 32) = 3.189	.0089[a]	12.60
	Prenatal	F (4, 32) = 27.72	<.0001[a]	54.74
	Postnatal	F (2, 32) = 24.32	<.0001[a]	24.02
				91.36
Distance Moved (cm)	Interaction	F (8, 32) = 2.263	.0483[a]	10.95
	Prenatal	F (4, 32) = 22.62	<.0001[a]	54.73
	Postnatal	F (2, 32) = 18.71	<.0001[a]	22.63
				88.31
Movement Duration (s)	Interaction	F (8, 32) = 2.242	.0503	10.37
	Prenatal	F (4, 32) = 24.84	<.0001[a]	57.46
	Postnatal	F (2, 32) = 17.06	<.0001[a]	19.74
				87.57
RATER 2				
GWS	Interaction	F (8, 37) = 2.700	.0190[a]	11.10
	Prenatal	F (4, 37) = 24.15	<.0001[a]	49.65
	Postnatal	F (2, 37) = 21.30	<.0001[a]	21.90
				82.65
Distance Moved (cm)	Interaction	F (8, 37) = 2.006	.0729	9.446
	Prenatal	F (4, 37) = 21.39	<.0001[a]	50.36
	Postnatal	F (2, 37) = 19.00	<.0001[a]	22.37
				82.18
Movement Duration (s)	Interaction	F (8, 37) = 2.257	.0448[a]	10.18
	Prenatal	F (4, 37) = 23.24	<.0001[a]	52.42
	Postnatal	F (2, 37) = 19.28	<.0001[a]	21.75

Sign	Predictor	*F*-Value	*P* Value	% of Total Variation
				84.35

a indicates statistical significance (p < .05).

Prenatal, prenatal treatment factor; Postnatal, postnatal challenge factor; Interaction, interaction effect between prenatal treatment and postnatal challenge.