Validation of a New Scoring Scale for Behavioral Assessment of L-Dopa-Induced Dyskinesia in the Rat: A New Tool for Early Decision-Making in Drug Development

Simon Loiodice, Anne-Sophie Denibaud, Wendy Deffains, Magali Alix, Pierre Montagne, Marine Seffals, Christophe Drieu La Rochelle

To cite this version:
Simon Loiodice, Anne-Sophie Denibaud, Wendy Deffains, Magali Alix, Pierre Montagne, et al.. Validation of a New Scoring Scale for Behavioral Assessment of L-Dopa-Induced Dyskinesia in the Rat: A New Tool for Early Decision-Making in Drug Development. ACS Chemical Neuroscience, American Chemical Society (ACS), In press, <10.1021/acschemneuro.7b00426>. <hal-01693938>
Validation of a new scoring scale for behavioral assessment of L-dopa-induced dyskinesia in the rat: a new tool for early decision-making in drug development

Simon Loiodice¹*, Anne-Sophie Denibaud¹, Wendy Deffains¹, Magali Alix¹, Pierre Montagne¹, Marine Seffals², Christophe Drieu La Rochelle²

¹Non-Clinical Department, Biotrial Pharmacology, 7-9 rue Jean-Louis Bertrand 35042 Rennes, France
²Plate-Forme H2P2, Université de Rennes 1, Biosit, 2 Av. du Prof. Léon Bernard, 35043 Rennes, France.

*Corresponding Author:
Simon Loiodice
Non-Clinical Department
Biotrial Pharmacology
7-9 rue Jean-Louis Bertrand
35042 Rennes, France
Email: simon.loiodice@biotrial.com
Abstract

The 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP)-treated non-human primate (NHP) has been described as the most translatable model for experimental reproduction of L-dopa-induced dyskinesia (LID). However, from a drug discovery perspective, the risk associated with investment in this type of model is high due to the time and cost. The 6-hydroxydopamine (6-OHDA) rat dyskinesia model is recommended for testing compounds but relies on onerous, and non-standard behavioral rating scales. We sought to develop a simplified and sensitive method aiming at assessing LID in the rat. The purpose was to validate a reliable tool providing earlier insight into the antidyskinetic potential of compounds in a time/cost-effective manner before further investigation in NHP models.

Unilaterally 6-OHDA-lesioned rats were administered L-dopa (20 mg/kg) and benserazide (5 mg/kg) daily for 3 weeks starting 4 weeks post-lesion, then co-administered with amantadine (20-30-40 mg/kg). An adapted rating scale was used to score LID frequency and a severity coefficient was applied depending on the features of the observed behavior.

A gradual increase (about 3-fold) in LID score was observed over the 3 weeks of L-dopa treatment. The rating scale was sensitive enough to highlight a dose-dependent amantadine-mediated decrease (about 2.2-fold) in LID score.

We validated a simplified method, able to reflect different levels of severity in the assessment of LID and, thus, provide a reliable tool for drug discovery.

Keywords: Parkinson’s disease; L-dopa; Dyskinesia; 6-OHDA; Drug discovery
Introduction

To date, L-dopa is the most effective drug for treating the signs and symptoms of Parkinson’s disease (PD) and is still considered the standard therapeutic agent \(^1,2\). However, after several years of treatment, L-dopa-induced dyskinesia (LID) occurs in the majority of patients \(^1,3–5\). These highly debilitating motor complications manifest as a variety of abnormal involuntary movements (AIMs), including severe and painful dystonic cramps (sustained abnormal muscle contractions), hyperkinetic and purposeless (choreiform) movements \(^5–7\).

The pathophysiological mechanisms underlying LID remain poorly understood despite substantial efforts in research \(^8,9\). LID has been proposed to be an aberrant form of neuroplasticity triggered by the combined effects of dopaminergic (DA) denervation and pulsatile stimulation of the DA receptors through repeated L-dopa intake \(^1,6\). With increasing duration of treatment, there is an increase in both the frequency and the severity of dyskinesia \(^6\). Given that L-dopa is still required for the treatment of advanced PD, the management of LID is a crucial challenge \(^8,9\). The limited success of anti-dyskinetic drugs in the clinic emphasizes the need to use appropriate animal models with good predictive value. The MPTP-treated non-human primate (NHP) has been described as the most translatable model and the gold standard for experimental reproduction of PD features, as well as dyskinetic syndrome \(^7,10,11\). However, the use of NHPs is costly, time-consuming and raises potential ethical issues. Hence, a drug candidate failing due to lack of efficacy in NHP dyskinesia models may result in losing a significant investment in the research program (potentially resulting in significant consequences for a small biotech company, and in no new drug being available for patients). In agreement with this idea, it has been proposed that early attrition of non-efficacious compounds is crucial for successful central nervous system (CNS) drug development \(^12,13\).

To counteract the limitations associated with the use of NHPs in antidyskinetic drug discovery, other simpler models were developed. The unilateral 6-hydroxydopamine (6-OHDA) rat dyskinesia model remains one of the most popular for modeling peak-dose dyskinesia \(^7\). This model was first described
by Cenci and colleagues in 1998 \cite{14} and is still widely used and recommended for testing compounds, because it produces a stable and reproducible behavioral outcome \cite{7,15}. This model relies on a behavioral assessment of LID using a rating scale that includes measurement of locomotor behavior (contralateral turning) as well as abnormal involuntary movements (AIMs), such as forelimb dyskinesia, dystonic posturing of the contralateral side of the body, axial dystonia, manifest as contralateral twisted posture of the neck and upper body, and orolingual dyskinesia, characterized by stereotyped jaw movements and contralateral tongue protrusion \cite{14}. Lundblad and colleagues elegantly demonstrated progressive worsening of AIMs over a 3-week L-dopa treatment using this scoring method \cite{16}. However, the severity of the AIMs was determined based on the frequency of occurrence of different subtypes of behaviors, without taking into account their intensity level, which is known to increase over time along with LID frequency with continued L-dopa treatment in PD patients, as reviewed by Bezard and colleagues \cite{6}. This limitation was later counterbalanced by adding an amplitude score (based on the extent of the movements) in order to differentiate between small but clear forelimb movements and dystonic-like movements involving the whole shoulder \cite{17,18}.

Another rating scale aimed at further assessing the intensity of LID was meanwhile developed by Steece-Collier and colleagues \cite{19}. Interestingly, in this latter method, the score was assigned not only based on the duration of a behavior, but also on its features, in order to better reflect LID severity \cite{19}. All of these rating scales have recently been compared in a validation study, in order to assess their reliability and translatability \cite{20}. The authors obtained slight differences in the responses to reference compounds depending on the rating scale, and suggest (as do other authors) a need to standardize the method \cite{7,20}. Notably, they stated that methods enabling the detection of variations in both duration and intensity of LID resulted in improved sensitivity to the effect of the reference compound amantadine, a non-competitive NMDA receptor antagonist \cite{7}.

In clinical research, it has been highlighted that the development of a single sensitive and robust rating scale was challenging because of the different types of dyskinesias and their different temporal patterns, anatomical distributions and associated disabilities \cite{8}. Similarly, although well
validated, the different methods described above reflect a lack of harmonization in the behavioral assessment of LID in animal models. Most of the existing scales are relevant for profound investigations into the neurobiological mechanisms of LID (if not designed specifically for this purpose) and rely on direct visual observation of a large number of fine behavioral parameters. As a consequence, the impact of subjectivity is high (the experimenter may introduce a subjective bias for each behavioral subtype observed, which influences the global LID score), and the use of a particular method can be challenging for an experimenter who was not trained by the laboratory in which the method was developed. This suggests a need to develop an ‘easy-to-use’ LID scoring scale that could be used in the 6-OHDA rat as a de-risking strategy, to have earlier insight into drug efficacy before carrying out further expensive studies using NHP models of dyskinesia.

In the present study, we sought to develop a new scoring method, adapted from previous rating scales for the assessment of LID in the rat 6-OHDA dyskinesia model. The purpose was to validate a simplified method focusing on a reduced number of behavioral subtypes easily identified in order to lower the impact of subjectivity and provide reliable data concerning the antidyskinetic potential of a compound in a time/cost-effective manner.

**RESULTS & DISCUSSION**

As shown in Fig. 2a and Fig. 2b, the animals that received a stereotaxic injection of 6-OHDA showed a significant decrease in the number of TH+ cells within the SNc (81.1% cell loss compared to the unlesioned side, \( t_{34} = 11.09; p<0.0001 \)). However, 2 rats showed a more partial lesion, of 31% and 44%, respectively (black dots in Fig. 2b). TH immunostaining performed in 2 sham animals revealed a number of 22,267 and 15,800 TH+ cells in the SNc, respectively, which was similar to the numbers of TH+ cells observed in the unlesioned side of lesioned rats (18,696 on average). Forelimb use asymmetry and akinesia were assessed using the cylinder test and the stepping test, respectively. As
shown in Fig. 2c, the percentage of contralateral forelimb paw use was significantly decreased in 6-OHDA-lesioned rats (29.4%) compared to sham rats (48.4%) (p=0.0005). This motor impairment was reversed after treatment with L-dopa (8 mg/kg) co-administered with benserazide (6 mg/kg). Consistently, the stepping test procedure reflected a similar motor impairment that was alleviated with the same L-dopa treatment (Fig. 2d). Indeed, while sham rats made an average of 23.8 adjusting steps with their contralateral paw over 3 trials, vehicle-treated lesioned rats made only 4.4 steps (p<0.0001) and L-dopa-treated lesioned rats made 16.8 steps (p<0.0001 compared to vehicle, p=0.0057 compared to sham). Fig. 2e, reveals that apomorphine challenge significantly induced contralateral rotations, especially in 8 rats, for which the challenge induced over 90 turns/45 min. Based on immunohistochemistry analysis, these rats underwent a dopaminergic lesion causing a TH+ cell loss of over 80% compared to the unlesioned side and were selected for the assessment of L-dopa-induced dyskinesia (apomorphine did not produce contraversive turn in 10 rats including 2 partially lesioned rats represented by black dots in Fig. 2e and Fig. 2b.

Drug-induced dyskinesia is a frequent debilitating complication in PD, associated with physical and social disabilities. For decades, substantial research efforts have been invested into novel therapeutics that could improve patients’ quality of life, with limited success. Consistent with the need to develop standardized and cost/time-effective preclinical tools for anti-dyskinetic drug discovery, we provide evidence highlighting the value of a novel ‘easy-to-use’ reliable method for behavioral assessment of LID in the rat.

The late stage (Phase II, Phase III) attrition rate of CNS drug-candidates is known to be particularly high, mainly due to lack of efficacy (or safety issues)\textsuperscript{12,13,23,24}. The successful development of CNS drugs relies on the development of appropriate animal models for efficacy testing in preclinical studies and a new paradigm for drug development that will give early readouts for proof of concept in order to allow attrition to occur much earlier in the process \textsuperscript{12,13}. This challenge applies to PD drug
discovery, especially for the development of antidyskinetic drugs. To date, the MPTP-treated NHP is considered the most predictive animal model of LID. Indeed, after chronic treatment with L-dopa, MPTP-treated monkeys develop choreic and dystonic movements that can be scored using clinical rating scales after only modest adaptations. But the cost and time required to implement this type of studies is associated with a risk of losing a significant investment. The 6-OHDA dyskinesia rat model appears to be a good option to provide a first insight into drug efficacy in a time/cost-effective manner. This model has been extensively described as valuable for antidyskinetic drug discovery and relies on behavioral scoring of the LID. A recent study sought to validate 3 contemporary AIM scales available in the literature and highlighted the interest of adding a severity parameter to the scoring of AIM frequency. The scoring scale developed in our study attempts to take this crucial parameter into account. Both severity and adjustment coefficients were applied to a set of specific behavioral subtypes that were selected to reflect different levels of severity. The goal was to provide an overview of the dyskinetic state of each animal rather than a very deep behavioral assessment. However, we designed our scale to reflect as much contrast as possible within the different levels of AIM severity to increase its sensitivity. It is worth noting that such an approach, relying on assigning a severity coefficient depending on the features of the observed behavior has been used in previous validated scales.

The evaluation of the evolution of the LID over the 3-week period of L-dopa treatment revealed a significant time-dependent increase in the LID score from week 0 to week 3 (p=0.0362). It worth noting that the LID scores correspond to ranked data. Hence, a square transformation was performed in order to ensure normal distribution and apply parametric tests. The evaluation of the evolution of the LID over the 3-week period of L-dopa treatment revealed a significant time-dependent increase in the LID score from week 0 to week 3 (p=0.0362) (Fig. 3a). It worth noting that the LID scores correspond to ranked data. Hence, a square transformation was performed in order to ensure normal distribution and apply parametric tests. **Fig 3b** shows that “increased locomotion with contralateral bias” and “dystonia” were the main AIM components of the global LID score at week 0 and week 3 respectively (score significantly higher compared to other components). Furthermore, **Fig 3b** illustrates a significant increase of the score for “increased locomotion with grabbing” on weeks 1 and 2 (p=0.0008, p=0.0038) compared to week 0. On week 3, a significant increase of the score for “axial posture” (p=0.0009) and “dystonia” (p=0.0151) was
observed compared to week 0. **Fig. 3c** shows a dose-dependent decrease in the LID score when the animals received amantadine, with a maximum effect at the 40 mg/kg dose (p=0.0034). **Fig. 3d** illustrates that at this dose, most of the effect occurred between 30 min and 70 min post L-dopa/benserazide treatment (p=0.007 at 30 min, p=0.0019 at 40 min; p=0.002 at 50 min and p=0.0155 at 70 min after L-dopa/benserazide administration). It is worth noting that the 30 mg/kg dose produced a significant decrease in the LID score at 50 min (p=0.0443). **Fig. 3e** further illustrates the effect of the active dose of amantadine (40 mg/kg) on individual LID scores. **Fig. 3f** shows the contribution of the different behavioral subtypes of the scale to the global LID score among the different treatment groups after the 3-week period of L-dopa treatment. It reveals that “dystonia” and “axial posture” constituted the main components of the global LID score and the main parameters that were alleviated by amantadine since a significant decrease was observed for “dystonia” score at 30 mg/kg (p=0.0379) and 40 mg/kg (p=0.0350). It is worth noting that the “axial posture” score was not alleviated at 30 mg/kg and slightly, but non-significantly, improved at 40 mg/kg.

Our data demonstrate that our rating scale was able to distinguish slight behavioral changes, as shown by the measurement of the time-dependent worsening of LID. Interestingly, we reported a significant increase of the LID score after one week of L-dopa dosing, consistent with a previous report indicating that AIM and LID can occur after the first-ever dose of L-dopa. Furthermore, the analysis of the different behavioral subtypes separately suggests a moderate increase of the most severe AIM components (i.e. dystonic circling, axial posture and dystonia) on weeks 1 and 2 accompanied by a significant augmentation of contralateral turning with grabbing. The two AIM subtypes “axial posture” and “dystonia” were significantly increased on the week 3 post-L-dopa treatment. This further illustrates the capability of our rating scale to highlight a progressive worsening of the dyskinetic state suggesting the sensitivity of the method. Moreover, the sensitivity of our rating scale allowed to show the amantadine-mediated alleviation of AIMS in a dose-dependent manner (using a tight dose range). It is worth noting that amantadine-mediated alleviation
of the global LID score was mainly associated with an improvement of “dystonia” and, in a lesser extent, “axial posture” AIM components with very limited effect on other behavioral subtypes observed in our rating scale. This is in line with previous data reporting amantadine-mediated alleviation of AIM score with no effect of contralateral turning. Consistent with previous reports, our study illustrates the importance of performing assessments at multiple time-points, as we were able to pinpoint the time window corresponding to the optimal effect of amantadine. This is particularly relevant in experiments investigating compounds for which poor pharmacokinetics data are available, as some agents may reduce the severity of the AIM score at a peak time but prolong the temporal course of motor dysfunction, or have a delayed effect.

One of the reasons proposed to explain the limited success of anti-dyskinetic drugs in the clinic is the lack of validated clinical outcome measures that are responsive to treatment despite the validity of multiple dyskinesia rating scales, and it has been pointed out that the development of a single sensitive and robust rating scale was challenging because of the different types of dyskinesias and their different temporal patterns, anatomical distributions and associated disabilities. Furthermore, the clinical assessment of LID is highly subject to the placebo effect. Similarly, in rat models of dyskinesia, various rating scales were developed and validated over the last decades. These methods rely on direct visual observation of a limited number of behavioral subtypes including orolingual/jaw movements (easily impacted by subjectivity) but only taking into account the severity level of AIM through frequency scores or amplitude scores depending on the observation of parameters easily subjectively appreciated (e.g. angle of torsion in moving animal). Other scales using a simpler method to determine a severity score focus on a large number of various fine subtypes of behavior or on observation of different parts of the animal’s body separately (neck, truck, limbs). While these methods are appropriate for investigating the neurobiological mechanisms of LID their subjectivity remains high: a large number of subjective parameters may increase the chance of experimenter-dependent errors in assessment influencing an animal’s global dyskinetic score. Similarly, assessment of very fine behavioral parameters such as orolingual
movements or torsion angle in moving rats may introduce a subjective bias, impact the reproducibility and harm to the need of standardization. The reproducibility could only be ensured by significant training provided by the laboratory in which the method was first validated, as well as internal validation in the laboratory where the new research will be carried out. In a recent study comparing different rating scales, the authors suggested a need for standardization of the methods for assessing LID, as they obtained slight differences in the responses to reference compounds depending on the method used. We believe that focusing on a limited number of behavioral subtypes could lower the impact of subjectivity and improve reproducibility while providing a global overview of the dyskinetic state of each animal. In our study, we observed a restricted number of behavioral subtypes to (i) minimize the impact of subjectivity (assuming that, as subjective bias would be applied to a smaller number of behavioral subtypes, there would be a lower impact on the final LID score due to fewer chances to produce “error”) and to (ii) obtain an overall view of the dyskinetic state of the animal at a particular time-point. Orolingual AIM were considered as too easily impacted by subjectivity (potentially hardly identified in dyskinetic moving rats) and were not assessed to minimize the risk to introduce a subjective bias. Our data demonstrate that this parameter may not be required to have a first insight into the antidyskinetic potential of a drug. We believe that this type of method could be more easily harmonized between different laboratories and meet expectations for the preclinical assessment of developing compound. Indeed, the purpose of our scale is not to support mechanistic studies aiming at understanding the neurobiological mechanisms of LID, but rather to provide, from a drug discovery perspective, a valuable, time/cost-effective outcome. While different subtypes of AIM may be mediated by different neurobiological mechanisms, assessing the effect of a test compound on the “global dyskinetic state” may help as part of a de-risking strategy to support decision-making during drug development (e.g., a go/no-go decision for initiating NHP studies). Our rating scale focuses on a limited number of behavioral subtypes, including non-dyskinetic behavior (contralateral circling) as well as AIMs with dystonic and hyperkinetic features. The behavioral subtypes assessed were associated with severity coefficients in
order to reflect the effect of a test compound on the global dyskinetic state of an animal. Our data demonstrate a time-dependent worsening of LID and a dose-dependent alleviation by amantadine, illustrating the sensitivity of this method.

In the present study, we provided data supporting the reliability of our new LID rating scale. However, as for other scales, the behavioral assessment relies on subjective observation of the animal, which is the main limitation of this type of a tool. The development of methods relying on more objective readouts, such as fluidic biomarkers or EEG monitoring, would be of major interest. For instance, substantial efforts have been made to identify gene/protein expression changes associated with LID in 6-OHDA animal models of PD \textsuperscript{29–31}, while other studies have attempted to identify fluidic biomarkers in PD patients \textsuperscript{32}. Other EEG-based approaches have pointed out different specific patterns that are potentially associated with LID in humans \textsuperscript{33–35} and in the 6-OHDA rat model \textsuperscript{36,37}. It could be of great interest to combine similar approaches with our behavioral rating scale and assess whether results obtained using these methods correlate with each other and provide valuable information during the assessment of drug efficacy.

Many companies are trying to reduce costs by outsourcing drug discovery to academic labs or contract research organizations due to unfavorable risk/reward balance with CNS drug development \textsuperscript{23,38–40,41,41}. These new partners need to develop standardized and accessible tools for CNS drug discovery. Increasing confidence in preclinical data during the development of anti-dyskinetic agents is of crucial importance. In the present study, we provide data validating a simplified method that is able to reflect different levels of severity when assessing LID. The sensitivity of the rating scale allowed us to highlight a time-dependent worsening of LID during 3 weeks of L-dopa treatment and a dose-dependent reversion of LID by amantadine. We propose this method as a valuable, reliable and sensitive tool for the initial testing of novel antidyskinetic compounds. Although the method proposed here may not be appropriate for investigations on the neurobiological mechanisms of LID (no detailed assessment of various AIM subtypes), it may provide early efficacy data to support
decision-making during the drug development process. Indeed, early attrition of inefficacious compounds is thought to be crucial for successful drug development (especially for CNS drugs), hence the need to develop reliable methods that can rapidly provide efficacy data in a cost-effective manner, before moving the compound forward in expensive and time-consuming studies. We believe that the use of our rating scale in the 6-OHDA dyskinesia model can be valuable if integrated in a de-risking strategy during the preclinical development of an antidyskinetic agent.

**Materials & Methods**

**Animals**

Adult male Sprague-Dawley rats (175-200 g) from Janvier Labs (Saint Berthevin, France) were maintained in a controlled environment (lights on 07:00-19:00, ±22°C) with food and water freely available. They were housed 3-4 per cage. This study was carried out in AAALAC-accredited facilities in strict accordance with the European Communities Council Directive (2010/63/EU) guidelines for the care of laboratory animals. The protocol was approved by the Biotrial Pharmacology Committee on the Ethics of Animal Experiments “Comité de Réflexion Ethique en Expérimentation Animale” (CR2EA), and in accordance with French Research Ministry regulations. All possible efforts were made to minimize suffering.

**Unilateral 6-OHDA lesion**

Twenty-six (26) rats were anesthetized with xylazine (Rompun®) 10 mg/kg, i.p., and ketamine (Imalgene®) 80 mg/kg, i.p., before being placed in a stereotaxic frame (David Kopf Instruments, CA, USA). The animals received a stereotaxic injection of 6-OHDA (4 µg/µL) or vehicle into the medial forebrain bundle. This injection consisted of a 2.5 µL deposit at the following coordinates (from bregma): AP: -4.4 mm; ML: -1.8 mm; DV: -7.9 mm (tooth bar at -2.4 mm) according to the Paxinos rat
brain atlas. 6-OHDA or vehicle was administered at a rate of 0.4 µL/minute. To avoid reflux, the needle was maintained at the injection site for 5 minutes after the injection.

To limit damage to noradrenergic neurons, imipramine (15 mg/kg, i.p.) was administered 15 min before 6-OHDA lesioning.

The extent of the lesion was evaluated 3 weeks post-surgery, based on net apomorphine-induced contralateral rotations. Animals were injected with apomorphine (0.05 mg/kg, s.c.), immediately placed in a 45 x 45 cm plexiglas open-field and video-tracked over 45 min. The number of contralateral rotations was measured afterwards using EthoVision® XT 9.0 (Noldus, Netherlands). Animals with a number of rotations greater than 90 were considered to have greater than 95% DA cell loss.

Prior to the cylinder and stepping tests (see Behavioral procedures), lesioned animals were randomized according to their number of contralateral turns in order to have a homogeneous distribution between the L-dopa- and saline-treated groups.

Drugs

All drugs (Sigma, France) were dissolved in saline, except 6-OHDA, which was dissolved in a saline solution containing 0.02% ascorbic acid.

In the cylinder and stepping tests, L-dopa methyl ester and benserazide were administered at 8 and 6 mg/kg, ip, respectively.

To induce LID, L-dopa methyl ester and benserazide were administered daily (between 9.00 and 10.00 AM) for 3 weeks at 25 and 5 mg/kg, s.c., respectively, starting from the fourth week post-surgery. This dose of L-dopa is the same as typically used in the behavioral sensitization paradigm and this type of sensitized context was associated with LID in PD patients. Amantadine or its vehicle was administered at 20, 30 and 40 mg/kg, i.p., 40 min before L-dopa/benserazide administration.
Behavioral procedures

Cylinder test

Forelimb use asymmetry was assessed as previously described 57 3 weeks after cerebral injection of 6-OHDA. Briefly, animals were placed in a glass cylinder (approximately 20 cm in diameter and 35 cm high) and video-recorded for 10 minutes to allow retrospective analysis of the behavior. A blinded observer scored the number of contacts made by individual forepaws with the cylinder wall. The percentage of left paw touches out of a total of 20 touches was determined. Limb use asymmetry was demonstrated by expressing the use of the impaired paw as a percentage of the total number of touches, with unbiased animals having a score of 50%.

Stepping test

Forelimb akinesia was assessed as previously described 58 3 weeks after cerebral injection of 6-OHDA. Briefly, the experimenter firmly held the rat’s hindquarters while it supported its weight on its contralateral forelimb. Then, the experimenter moved the rat forward along the table (0.9 m in 5 seconds) three consecutive times per session. All sessions were video-recorded and the number of adjusting steps was counted afterwards by a blinded investigator. For each session, the total score calculated was the sum of the number of adjusting steps observed for the contralateral paw in the three tests. The sessions took place between 10.30 and 11.30 AM, 3 weeks post-surgery.

L-dopa induced dyskinesia

L-dopa mediated induction of AIMs is only possible in animals with a massive nigral DA lesion 26,59. Thus, the 8 rats displaying a full DA lesion (i.e., over 90 contralateral turns after apomorphine challenge 60,61) were included in this procedure. These rats received a daily injection of L-dopa and benserazide for 3 weeks (15 testing sessions), starting the 4th week post-surgery (after the stepping test and cylinder test procedures). Each session was performed between 9.00 and 10.00 AM.
Rats were placed in a cylinder (20 cm in diameter and 35 cm high) 20 min after the L-dopa/benserazide administration and their behavior was video-recorded for 1 hour in order to allow off-line scoring of the LID by an experimenter who was unaware of the treatment condition.

Then, 4 additional testing sessions were performed in order to assess the efficacy of amantadine in reversing the LID. Animals received vehicle on day 1, amantadine 40 mg/kg on day 2, amantadine 30 mg/kg on day 3 and amantadine 20 mg/kg on day 4 (Fig. 1). Amantadine or its vehicle was administered 40 min before L-dopa/benserazide administration and 1 hour before the animals were placed in the cylinder and video-tracked for 1 hour. Again, the behavior of the animals was video-recorded and LID scoring was performed retrospectively by a blinded experimenter.

Consistent with previous reports that assumed different levels of severity depending on the features of the AIMS\textsuperscript{17–19}, we hypothesized that different behaviors reflected different LID intensity levels. We attributed a global severity score to each rat, based on both the frequency and the intensity of the behaviors, in order to reflect gradual levels of AIM severity. The AIMS were classified into 5 subtypes according to their level of severity according to previously published rating scales\textsuperscript{14,16–22} (see Table 1). The purpose of our approach was to simplify the scoring method and provide a global AIM score reflecting the dyskinetic state of each animal. To this end, we developed a rating scale including both axial dystonia (neck/trunk torsion toward the side contralateral to the lesion) and forelimb dyskinesia (side to side, up to down tapping or circular movement of the right forelimb, “grabbing”) within the same assessment grid. For instance, the AIM subtypes “increased locomotion with grabbing” and “axial posture” include both changes in neck/trunk position and grabbing behavior (Table 1). In line with previous scales\textsuperscript{14,16}, contralateral circling (included under the item “increased locomotion with contralateral bias” in Table 1) was not considered “dyskinesia,” since contralateral turning can be induced by dopaminergic agonists with very low dykinesiogenic potential\textsuperscript{16,62}. This behavior was considered “normal behavior” following L-dopa administration in unilaterally 6-OHDA lesioned rats.
and, thus, corresponds to the lowest severity score. The purpose was to eventually weight the score if AIMs such as grabbing or others described in Table 1 were observed.

Each rat was observed for 1 min every 10 min, from 20 to 80 min post L-dopa/benserazide administration. For each subtype of behavior (Table 1), a frequency score between 0 and 3 (with 0 = absent, 1 = occasional: less than 50% of the time, 2 = frequent: more than 50% of the time, 3 = continuous, uninterrupted) was assigned.

These frequency scores were multiplied by a severity coefficient (see Table 1) that was determined prior to the behavioral analysis. These severity scores were adapted from previous studies applying similar amplitude scores \(^{17,18,63-65}\) in order to maximize the sensitivity of the scale and to reflect as much contrast as possible within the different levels of AIM severity. In our study, severity coefficients were applied as follows: 1 = horizontal body position, contralateral circling on 4 paws, 2 = horizontal body position, contralateral circling on 4 paws, grabbing movement with contralateral forelimb paw, 8 = horizontal body position with pronounced deviation toward the contralateral side (>30°) with nose close to the level of the tail, possible loss of balance due to the twisted position, 12 = vertical body position, contralateral deviation of the head, neck and upper trunk 60°-90°, grabbing and contralateral rotations on hindlimbs, 20 = vertical body position with sustained and severe torsion of neck and trunk at 90°-180° causing the rat to lose balance, purposeless "choreiform" twisting movements, animal stuck in the twisted position/possible contralateral rotations on hindlimbs. The 2 bounds coefficients (1 and 20) were determined by taking into account the maximum level of dyskinetic state of the highest dyskinetic rat (i.e. rat stuck in 180° twisted position for the entire minute of observation) after 3 weeks of L-dopa treatment and the minimum level of dyskinetic state of the lowest dyskinetic rat (i.e. classical contralateral circling with no AIM during the minute of observation) on the first day of the 3-week period of L-dopa treatment. This first step aims at calibrating the system. The intermediary coefficients (2, 8 and 12) were determined arbitrarily but in order to discriminate well between the different AIM severity levels. Thus, we applied a moderate
factor of magnitude between “increased locomotion with contralateral bias” and “increased locomotion with grabbing” (1 to 2), then a more important difference between “increased locomotion with grabbing” and “dystonic circling” (2 to 8), a mild to slight difference between behaviors in which a pronounced deviation of the body is observed (8, 12 and 20 for “dystonic circling” and “axial posture and “dystonia” respectively) and a substantial factor of magnitude between “increased locomotion with contralateral bias” and “dystonia” (1 to 20).

For animals displaying different subtypes of behaviors during the same observation period, an adjustment coefficient was also applied, as described in Table 1. This adjustment coefficient aimed to improve the accuracy of the scoring method by taking into account slight behavioral differences. For instance, we assumed that a rat displaying only “dystonia” (see Table 1) during the entire minute of observation time was more severely affected than a rat that displayed “dystonia” most of the time but that also displayed “axial posture” and/or other less severe symptoms.

A global severity score taking into account both the frequency and the intensity of the AIMs was obtained for each rat (the detailed calculations and formulae used are available in Supplementary Material 1). The maximum score that could be accumulated per testing session was 420 (maximum score per observation point: 60; number of observation points per session: 7).

**Histological analysis**

After the last day of LID scoring, the animals were decapitated under pentobarbital anesthesia (200 mg/kg) and the brain was quickly removed at 4°C on fresh ice. The brains were then post-fixed in ice-cold 4% paraformaldehyde over 24h and processed for paraffin embedding before being sectioned (7-µm slices) in the coronal plane. The slides were stained using antibodies against tyrosine hydroxylase (TH) (rabbit 1/5000: Millipore AB152) and the DAKO LSAB and HRP system (DAKO Real, DAKO France). Immunostaining was processed on a Discovery XT ® Platform (Roche Ventana, Tucson,
AZ, USA). All slides were scanned at 20x magnification for whole slide imaging using a NanoZoomer 2.0RS scanner (Hamamatsu, Japan) and analyzed using NIS-AR software (Nikon, Japan).

TH+ cells in the substantia nigra pars compacta (SNc) were counted at 3 different rostrocaudal levels (AP: -4.80 mm, -5.30 mm and -6.04 mm) according to the Paxinos rat brain atlas. At each level, the number of TH+ cells contained in the SNc was counted in three adjacent sections by a blinded experimenter. For each section, the boundaries were chosen by examining the shape of the cells and referring to the Paxinos rat brain atlas. At 100x magnification, only the cells with a pyramidal shape were counted. For each animal, the total number of TH+ cells was estimated using the Konigsmark formula: \( N_t = N_s \times (S_t/S_s) \) where \( N_t \) = total number of cells; \( N_s \) = number of cells counted; \( S_t \) = total number of sections counted; \( S_s \) = total number of sections through the SNc. The percentage of DA cell loss was estimated using the following formula: \( 100 - (\text{TH+ cell number in lesioned SNc/TH+cells number in sham SNc}) \times 100 \).

Data and statistical analysis

Data were analyzed by Student’s t-test, one-way ANOVA, repeated measures ANOVA or two-way repeated measures ANOVA depending on the experimental design. Where appropriate, post-hoc analyses were carried out with Tukey’s or Dunnett’s tests. All reported p-values are two-sided. The normality of the variables was assessed by Shapiro-Wilk test and the homoscedasticity was assessed by Levene’s test. When data were not normally distributed, the data were square-transformed. Results were expressed as means ± SEM.

Supporting Information

Excel sheet containing all formulae for LID score calculation.

Author Contributions
SL designed the study, analyzed the data and wrote the manuscript. ASD, WD, MA, PM and MS performed the experiments and CDLR reviewed the manuscript.

**Funding Sources**

This research was funded by Biotrial Pharmacology.

**Conflict of Interest**

SL, ASD, WD, MA, PM and CDLR are employees of Biotrial Pharmacology.

**Acknowledgements**

The authors thank Cliona McSweeney, Fiona McAlpine and Stéphanie Le Goaller for providing very useful comments and for reviewing the manuscript.
References


63. Lane, E. L., Soulet, D., Vercammen, L., Cenci, M. A. & Brundin, P. Neuroinflammation in the
(2008).

64. Mela, F. *et al.* Antagonism of metabotropic glutamate receptor type 5 attenuates L-DOPA-
induced dyskinesia and its molecular and neurochemical correlates in a rat model of Parkinson’s

influences on L-dopa-induced dyskinesia in the 6-OHDA model. *Brain Res. Bull.* 78, 248–253
(2009).

66. German, D. C. & Manaye, K. F. Midbrain dopaminergic neurons (nuclei A8, A9, and A10): three-
## Table Legends

**Table 1. Description and severity classification of the AIMS observed.** Severity and adjustment coefficients were applied according to the type of AIMS observed.

<table>
<thead>
<tr>
<th>AIM subtype</th>
<th>Increased locomotion with contralateral bias</th>
<th>Increased locomotion with grabbing</th>
<th>Dystonic circling</th>
<th>Axial posture</th>
<th>Dystonia</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description of the behavior</td>
<td>horizontal body position</td>
<td>horizontal body position</td>
<td>horizontal body position with pronounced deviation toward the contralateral side</td>
<td>vertical body position</td>
<td>vertical body position with sustained and severe 180° trunk torsion</td>
</tr>
<tr>
<td></td>
<td>contralateral circling on 4 paws</td>
<td>contralateral circling on 3 paws</td>
<td>nose close to the level of the tail</td>
<td>contralateral deviation of the head, neck and trunk</td>
<td>purposeless &quot;choreiform&quot; twisting movements, loss of balance, falls</td>
</tr>
<tr>
<td></td>
<td>grabbing movement with contralateral forelimb paw</td>
<td>possible loss of balance due to the twisted position</td>
<td>grabbing and contralateral rotations on hindlimbs</td>
<td>animal stuck in the twisted position/possible contralateral rotations on hindlimbs</td>
<td></td>
</tr>
<tr>
<td>Severity coefficient</td>
<td>1</td>
<td>2</td>
<td>8</td>
<td>12</td>
<td>20</td>
</tr>
<tr>
<td>Adjustment coefficient</td>
<td>-0.6</td>
<td>-0.4</td>
<td>-0.2</td>
<td>-0.1</td>
<td>none</td>
</tr>
</tbody>
</table>
Figure Legends

**Fig. 1. Study design.** General experimental procedure and timeline.

**Fig. 2. Assessment of nigral loss and forelimb use asymmetry and akinesia.** (a) Immunolabeling of tyrosine hydroxylase (TH). The bar corresponds to 1 mm. VTA = ventral tegmental area (b) Number of TH+ cells counted in the SNc (n=18). An unpaired Student’s t-test revealed severe DA cell loss in the SNc on the lesioned side compared to the unlesioned side (t_{34}=11.09; ***p<0.0001). All rats displayed a lesion with over 80% cell loss compared to the unlesioned side, except for 2 rats with partial lesions with 31% and 44% cell loss, respectively (black dots). Center lines and error bars throughout the figure represent means ± SEM. (c) Contralateral paw use in the cylinder test (n=8-10). A one-way ANOVA revealed a significant treatment effect, with F_{2,23}=15.6; p<0.0001. The percentage of contralateral paw use was significantly decreased in vehicle-treated lesioned rats compared to sham rats (***p=0.0005 with Tukey’s post-hoc test) with L-dopa/benserazide (8 and 6 mg/kg, ip respectively) reversing this impairment (no significant difference compared to sham rats and ***p<0.0001 when compared to lesioned vehicle-treated rats with Tukey’s post-hoc test). (d) Number of adjusting steps performed with the contralateral paw (n=8-10). A one-way ANOVA revealed a significant treatment effect, with F_{2,23}=42.21; p<0.0001. The number of adjusting steps was significantly decreased in vehicle-treated lesioned rats compared to sham rats (***p<0.0001 with Tukey’s post-hoc test) with L-dopa/benserazide (8 and 6 mg/kg, ip respectively) alleviating this impairment (**p=0.0057 when compared to sham rats and ***p<0.0001 when compared to vehicle-treated lesioned rats with Tukey’s post-hoc test). (e) Apomorphine-induced contralateral rotations (n=8-18). An unpaired Student’s t-test revealed that lesioned rats produced significantly more contralateral rotations compared to sham (t_{24}=2.183; p=0.0391), with 8 rats displaying over 90 turns/45 min (dotted line). Ten other lesioned rats were below this threshold (the 2 black dots represent the partially lesioned rats). Center lines and error bars throughout the figure represent means ± SEM.

**Fig. 3. L-dopa induced dyskinesia.** (a) Time course of the development of LID over 3 weeks of daily L-dopa/benserazide treatment (n=8). The ranked data were square-transformed to allow normal distribution and the use of parametric statistics. A repeated measures one-way ANOVA revealed a significant time effect, with F_{1.549,10.84}=16.6; p=0.0009 . The data reflect a time-dependent increase in the LID score over time with a significant increase in week 1, week 2 and week 3 compared to week 0.
Contribution of the different AIM components to the global LID score over the 3 weeks of L-dopa treatment (n=8). Adjustment coefficients were not applied (as their purpose is to weight the global LID score) and the ranked data were square-transformed to allow normal distribution and the use of parametric statistics (but untransformed data are displayed to improve clarity). For each time-point, a one way ANOVA followed by a Tukey’s post-hoc tests revealed a significant increase of “increased locomotion with contralateral bias” and “dystonia” at week 0 and week 3 respectively (***p<0.001 and *p<0.05 respectively) compared to other behavioral subtypes. A one-way ANOVA revealed a significant time effect for “increased locomotion with grabbing” (F_{3,28}=9.056; p=0.0002), “axial posture” (F_{3,28}=7.594; p=0.0007) and “dystonia” (F_{3,28}=4.287; p=0.0131). A Dunnett’s post-hoc analysis showed a significant increase of “increased locomotion with grabbing” on week 1 and week 2 compared to week 0 (**p=0.0008 and +p=0.0038 respectively), a significant increase of “axial posture” and “dystonia” on week 3 compared to week 0 (**p=0.0009 and +p=0.0151 respectively). Turning = increased locomotion with contralateral bias, Turning+Grab. = increased locomotion with grabbing, Dyst. Circling = dystonic circling. Amantadine-mediated alleviation of the LID score (n=8). The ranked data were square-transformed to allow normal distribution and the use of parametric statistics. A repeated measures one-way ANOVA revealed a significant treatment effect, with F_{1,167,8,168}=9.447; p=0.0129 . The data reflect a dose-dependent decrease in the LID score after treatment with amantadine, with the maximum effect at 40 mg/kg (**p=0.0034 compared to vehicle with Dunnett’s post-hoc test). Time course of the effect of amantadine (n=8). The ranked data were square-transformed to allow normal distribution and the use of parametric statistics. A repeated measures two-way ANOVA revealed no significant treatment effect (F_{3,28}=2.609; p=0.0712), but a significant time effect (F_{6,168}=8.9; p<0.0001) and a significant time x treatment interaction (F_{18,168}=1.756; p=0.0345). The Dunnett’s post-hoc analysis showed a significant effect of amantadine 40 mg/kg compared to vehicle at 30 min (**p=0.007), 40 min (**p=0.0019), 50 min (**p=0.002) and 70 min (*p=0.0155) post-L-dopa administration. The 30 mg/kg dose also significantly decreased the LID score compared to vehicle at 50 min post-L-dopa administration (*p=0.0443 with Dunnett’s post-hoc test). Ama20, Ama30 and Ama40 = amantadine 20 mg/kg, 30 mg/kg and 40 mg/kg respectively. Individual effect of amantadine 40 mg/kg on LID score (n=8). The ranked data were square-transformed to allow normal distribution and the use of parametric statistics. A paired Student’s t test revealed a significant decrease in the LID score compared to vehicle-treated rats (t_{7}=3.233; p=0.0144). Effect of amantadine on the different behavioral subtypes. Adjustment coefficients were not applied (as their purpose is to weight the global LID score) and the ranked data were square-transformed to allow normal distribution and the use of parametric statistics (but untransformed data are displayed to improve clarity). For each group of treatment, a one way
ANOVA followed by a Tukey’s post-hoc tests revealed a significant increase of “dystonia” and “axial posture” scores compared to other behavioral subtypes (*p<0.05; **p<0.01 and ***p<0.001 respectively). A one-way ANOVA revealed a significant treatment effect for the dystonia component (F_{3,28}=3.066; p=0.0442). A Dunnett’s post-hoc analysis showed a significant effect of amantadine 30 mg/kg and 40 mg/kg compared to vehicle (^p=0.0379 and ^p=0.035 respectively) for this particular behavioral subtype.
FIGURE 1

6-OHDA MFD unilateral lesion

Apomorphine challenge (selection of severely lesioned rats)

L-dopa + Benserazide

Motor assessment
- Cylinder test
- Stepping test

Dyskinesia assessment

Day 1: Vehicle
Day 2: Amantadine 40 mg/kg
Day 3: Amantadine 30 mg/kg
Day 4: Amantadine 20 mg/kg

Brain sampling
Histological analysis
FIGURE 2

(a) Lesioned side versus unlesioned side. TH = substantia nigra pars compacta (SNc), ventral tegmental area (VTA), dorsal raphe nucleus (DRN). Scale bar = 200 μm.

(b) Number of TH+ cells in the SNc. Unlesioned side versus 6-OHDA. ***p < 0.001.

(c) % of contralateral paw use. L-Dopa treatment significantly increases use over saline. **p < 0.01.

(d) Number of adjusting steps. L-Dopa treatment significantly increases adjusting steps over saline. ***p < 0.001.

(e) Contralateral turns/45 min. *p < 0.05.
FIGURE 3

(a) [Graph showing LID score distribution over time with different treatments.]

(b) [Graph showing the evolution of AIM scores over time with different treatments.]

(c) [Bar graph showing LID score by amantadine treatment levels.]

(d) [Line graph showing LID score increase over time with different treatments.]

(e) [Graph showing LID score increase with amantadine treatment, p=0.0144.]

(f) [Bar graph showing AIM score by treatment, with specific annotations for significance.]

Legend:
- Turning
- Turning + Grab.
- Dyst. Circing
- Axial Posture
- Cystonia

Vehicle, Ama20, Ama30, Ama40
Validation of a new scoring scale for behavioral assessment of L-dopa-induced dyskinesia in the rat: a new tool for early decision-making in drug development

Simon Loiodice, Anne-Sophie Denibaud, Wendy Deffains, Magali Alix, Pierre Montagne, Marine Seffals, Christophe Drieu La Rochelle

Validation of a simplified rating scale providing a global view of the possible anti-dyskinetic effect of a developing compound to facilitate early decisions-making during drug development