Mini-review


G.T. Whiteside a,*, A. Adedoyin b, L. Leventhal a

a Neuroscience Discovery Research, Wyeth Research, CN 8000, Princeton, NJ 08543, USA
b Drug Safety and Metabolism, 500 Arcola Road, Collegeville, PA 19426, USA

Received 8 August 2007; received in revised form 3 January 2008; accepted 7 January 2008

Abstract

A number of previous reviews have very eloquently summarized pain models and endpoints in animals. Many of these reviews also discuss how animal models have enhanced our understanding of pain mechanisms and make forward-looking statements as to our proximity to the development of effective mechanism-based treatments. While a number of reports cite failures of animal pain models to predict efficacy in humans, few have actually analyzed where these models have been successful. This review gives a brief overview of those successes, both backward, providing validation of the models, and forward, predicting clinical efficacy. While the largest dataset is presented on treatments for neuropathic pain, this review also discusses acute and inflammatory pain models. Key to prediction of clinical efficacy is a lack of side effects, which may incorrectly suggest efficacy in animals and an understanding of how pharmacokinetic parameters translate from animals to man. As such, this review focuses on a description of the pharmacokinetic–pharmacodynamic relationship for a number of pain treatments that are effective in both animals and humans. Finally we discuss where and why animal pain models have failed and summarize improvements to pain models that should expand and improve their predictive power.

Keywords: Inflammation; Neuropathic; Rat; Exposure; Efficacy; Analgesic; Translation

1. Introduction

A number of reviews of pain models have been published in recent years (Beggs and Salter, 2006; Blackburn-Munro, 2004; Eaton, 2003; Honore, 2006; Negus et al., 2006; Walker et al., 1999; Zimmermann, 2001). Often focusing on neuropathic pain, these provide in-depth explanations of the disease models and measurements (endpoints) that are used to quantify the extent of pain that is present. In addition, a number of reviews have summarized studies investigating mechanisms involved in pain transduction (Blackburn-Munro and Erichsen, 2005; Campbell and Meyer, 2006; Dickenson et al., 2002; Littlejohn and Guymer, 2006; Moalem and Tracey, 2006; Urban et al., 2001). More difficult to summarize is how measurement of pain in animals correlates with pain in man and, as such, only a minority of reviews specifically address this issue (the reader is directed towards Blackburn-Munro (2004) for further details). These multiple reviews highlight the tremendous increase in our understanding of pain as a disease that have been made in recent decades. In addition, they demonstrate that a number of approaches, based on an understanding of pain mechanisms, are being employed to develop new therapies for the treatment of acute and chronic pain.

Articles discussing the mechanistic basis of pain states often focus on novel therapies that are at the preclinical or early clinical stage (Dickenson et al., 2002; Rice and Hill, 2006; Urban et al., 2001); therefore, the validity of the models that were utilized in their discovery cannot yet be determined.
A limited number of articles have discussed the pharmacology of known analgesics in animals and tried to correlate this with pharmacology in the clinic (Blackburn-Munro and Erichsen, 2005; Fishbain et al., 2000; Pullar and Palmer, 2003); however, each is limited to a single molecular target, and none have considered the impact of pharmacokinetics on this correlation. One study that stands out was conducted by Kontinen and Meert (2003) who performed a semi-quantitative evaluation of the predictive validity of four established peripheral nerve injury models across more than 3000 studies. Here the authors concluded that the pharmacological sensitivity of these models ranged from 61% to 88%. In such an undertaking, it is impossible to consider significant experimental variables (such as drug exposure, route, dose-range, vehicle, testing methodology, etc.); however, this study clearly indicates the value of these nerve injury models.

Despite the obvious contributions that animal models have made to our understanding of pain pathobiology and the fact that new drugs have been developed based on efficacy in animal models (Campbell and Meyer, 2006), substantial criticism has been levied for their lack of perceived predictive power (Blackburn-Munro, 2004; Hill, 2004; Rice and Hill, 2006; Vierck, 2006). By way of investigating the predictive validity of animal pain models, this review focuses on compounds that are currently used clinically to treat pain, both approved and off-label, and describes pharmacokinetic—pharmacodynamic relationships in both animals and humans. As such the value of these models in predicting dose and exposure for specific analgesic mechanisms is investigated. In contrast to the systematic review of Kontinen and Meert (2003), here we focus on a smaller number of clinically relevant compounds and studies; in doing such, we can hold constant the variables described above. In addition, we felt it was critical to, first, ensure that efficacious doses (and exposures) in animal models were not confounded by the presence of side effects, and second, take into account consideration of drug exposure in both animals and humans. The majority of the animal efficacy, side effect and pharmacokinetic data cited were generated in-house (and supplemented from the literature), while the human efficacious doses and exposures were found via literature search. It is noteworthy that the acute, inflammatory and nerve injury models selected for analysis were chosen based upon their common use in the industrial setting for drug discovery (Iyengar et al., 2004; Jarvis et al., 2002; Sullivan et al., 2007; Valenzano et al., 2005). In addition, the area under the curve (AUC) is used as the measure of drug exposure and in order to make the species comparisons more comparable, single dose pharmacokinetic data was used for both rat and human data. Furthermore, identical AUC measures (e.g., AUC (0–infinity) or AUC (0–12 h)) in animals and humans were used for each compound analyzed to ensure that the comparisons are based on similar data sets. The maximum plasma concentration (Cmax) at the efficacious dose in rat was also compared to the plasma concentration at the maintenance dose in humans. Towards the end of the review, we explore the limitations of the commonly used models and close with a summary of current research aimed at bettering the predictivity of preclinical pain models and endpoints.

2. Commonly used animal pain models in drug discovery

Methodologies employed for assessing pain in animals can be broken down into endpoints and models. Endpoints are the tests conducted to ascertain the extent of pain. Endpoints are most commonly either spontaneous pain-related behaviors or thresholds to a ramping stimulus. Pain-related behavior, such as biting, licking, guarding and flinching are absent, or minimal, in normal animals and are only elicited on establishment of a model. An evoked stimulus—response measurement consists of an application of a stimulus of increasing intensity, which is commonly thermal or mechanical in nature, followed by measurement of a threshold or latency at which the animal displays nocifensive behavior. When any of these stimuli—response measurements are applied to normal animals, they constitute a measure of normal nociception and can be used to assess the effect of frank analgesics (defined as those that inhibit non-pathological, nociceptive pain) such as opioids and local anesthetics. Hargreave’s apparatus, von Frey fibers, hot plate, tail-flick, tail-dip and Randall–Selitto apparatus are all tools for applying a ramping stimulus to evoke a response (Campbell and Meyer, 2006; Honore, 2006; Sullivan et al., 2007; Valenzano et al., 2005).

Models describe manipulations of animals that are performed in order to generate a pain state, which is commonly manifest as behavioral hypersensitivity such as hyperalgesia, allodynia or both and/or spontaneous pain behavior. Commonly used models can be broken down into three main groups: the first involves local injection of a pain causing substance, such as capsaicin, bradykinin or dilute acid. The second involves injection of substances, either locally or systemically, that cause an inflammatory response and pain subsequent to the inflammation. Examples of such substances include carrageenan, zymosan and Freund’s complete adjuvant. The final group involves injury to the nervous system by direct mechanical, metabolic or chemical means. Examples of each include spinal nerve ligation (mechanical), streptozocin treatment (metabolic) and taxol treatment (chemical). Each of these models is generally paired with one, or a number of, endpoints. In this way, the extent of the hypersensitivity can be measured, and reversal of pain, back to “normal” levels, by pharmacological intervention can be assessed. Effective treatments are known as anti-hyperalgesics or anti-allodynic depending on the stimulus modality they reverse; however, it is important to note that the frank analgesics mentioned above will also reverse pain in these hypersensitivity assays.

3. Predictive value of animal models of acute pain

Clinical treatment of moderate to severe acute pain, such as that caused by a surgical incision, continues to be dominated by opioids (Leykin et al., 2007). As such, we have summarized preclinical and clinical data for morphine and oxycodone as
prototypic opioid agonists. Efficacy data for morphine in the hot plate assay were generated in-house using methods described in detail elsewhere (Whiteside et al., 2005). Efficacy data for oxycodone were identified from literature reports that utilize equivalent methodology (Lemberg et al., 2006) to those used in-house. Stated minimal effective doses (MEDs), shown in Tables 1–3, are doses that do not produce statistically significant motor deficits in our in-house rotarod assay of ataxia, using methodology as previously described (Valenzano et al., 2005). All in-house pharmacokinetic studies were conducted according to previously described methods (Sullivan et al., 2007).

In acute pain, the efficacious exposure in rats for morphine is 3 times greater than that observed in humans (Table 1). In contrast, the efficacious exposure for oxycodone is almost 40-fold greater than that observed in humans. Considering \( C_{\text{max}} \) the ratios are reversed with a 51-fold higher exposure in rats as compared to humans for morphine, while the efficacious concentration for oxycodone in rats is 0.8 times that in humans. It is worth noting that the human exposures were determined from immediate release and sustained release formulations for morphine and oxycodone, respectively, which may make the human to rat correlation less accurate than using similar formulations for both. This likely explains why comparing exposures for morphine and plasma concentration for oxycodone yield very close ratios (2.9 and 0.8 respectively) whereas the reverse yields ratios that do not approximate. Beyond this caveat, the observed difference may be due to species differences in metabolism, brain and tissue penetration, plasma protein binding or other factors altering availability of compound at the target tissue. It is noteworthy that rat exposures are often described at the MED, for a single administration, while clinical exposures are described at the maintenance dose based on repeated administration. The MED is the lowest dose that elicits a statistically significant effect; it is therefore, by definition, an effect of limited magnitude. This is likely to be in contrast to a maintenance dose in patients, which is expected to have an effect of larger magnitude such that the patient realizes a substantial benefit. This discrepancy may result in the rat efficacious exposure underestimating the exposure necessary to maintain efficacy in humans. In addition, efficacy in animals is based upon single acute dosing, while that in man is typically based upon chronic administration thus chronic dosing in preclinical studies may improve the predictivity of the models and exposure. However, chronically administered drugs can produce tolerance that requires increasing doses to maintain efficacy; alternatively, they can cause metabolic induction, leading to decreased exposures. Such effects would alter the interpretation of efficacious exposures compared across species. This is one potential limitation of the analysis presented here however, these limitations are commonly encountered in drug discovery.

The data in Table 1 focuses on the relationship between rat and human efficacious plasma concentrations and drug exposures. We can conclude from the table that overall, efficacious drug exposure in the rat approximates to efficacious exposure in humans. Although the routes of administration differ, it is assumed that efficacious exposure and plasma concentration is independent of route of administration. Comparisons based on dose, however, cannot be made, since the routes of administration between rat and human are not consistent (subcutaneous versus oral, respectively). While the hot plate assay is commonly used as a measure of acute pain it actually is more an assay for normal nociceptive pain and as such may only be predictive for a subgroup of treatments such as opioids and local anesthetics. Caution is therefore warranted in using this model to predict clinical efficacy in conditions such as post-operative pain (also referred to as acute pain); in this case, and as discussed in Section 7 more appropriate models are now available. In addition it should also be realized that data generated in rodent studies have successfully predicted the occurrence of at least some of the side effects (e.g.,

---

**Table 1**

Comparison of the pharmacokinetic—pharmacodynamic relationship of acute pain drugs in rats and humans

**A. Exposure**

<table>
<thead>
<tr>
<th>Compound</th>
<th>Human daily dose (mg)</th>
<th>Human maintenance dose (mg/kg)</th>
<th>Rat MED hot plate (mg/kg)</th>
<th>Rat exposure (AUC; ng h/ml) at MED</th>
<th>Human exposure (AUC; ng h/ml)</th>
<th>Exposure ratio (rat/human)</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morphine</td>
<td>60</td>
<td>0.9</td>
<td>3</td>
<td>799</td>
<td>279</td>
<td>2.9</td>
<td>Rat, Wyeth in-house; human, Anonymous (2007)</td>
</tr>
<tr>
<td>Oxycodone</td>
<td>160</td>
<td>2.3</td>
<td>0.6</td>
<td>71100</td>
<td>1856</td>
<td>38</td>
<td>Rat, Wyeth in-house, Huang et al. (2005); human, Anonymous (2007)</td>
</tr>
</tbody>
</table>

**B. Concentration**

<table>
<thead>
<tr>
<th>Compound</th>
<th>Human daily dose (mg)</th>
<th>Human maintenance dose (mg/kg)</th>
<th>Rat MED hot plate (mg/kg)</th>
<th>Rat ( C_{\text{max}} ) (ng/ml) at MED</th>
<th>Human ( C_{\text{max}} ) (ng/ml)</th>
<th>Concentration ratio (rat/human)</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morphine</td>
<td>60</td>
<td>0.9</td>
<td>3</td>
<td>976</td>
<td>19</td>
<td>51</td>
<td>Rat, Wyeth in-house; human, Anonymous (2007)</td>
</tr>
<tr>
<td>Oxycodone</td>
<td>160</td>
<td>2.3</td>
<td>0.6</td>
<td>123</td>
<td>156</td>
<td>0.8</td>
<td>Rat, Wyeth in-house; human, Anonymous (2007)</td>
</tr>
</tbody>
</table>

MED, minimum efficacious dose; AUC, area under the curve. All rat data were generated following subcutaneous administration. All human data were generated following oral administration and based on 70 kg body weight. Rat efficacy data for morphine were generated in-house at Wyeth and data for oxycodone were literature derived (Lemberg et al., 2006). AUC for all studies is AUC (0–infinity). Extrapolations of pharmacokinetic data assume linearity. Morphine rat exposure is extrapolated from data after a 10 mg/kg dose and \( C_{\text{max}} \) is extrapolated from a 1 mg/kg dose. Oxycodone rat exposure is extrapolated from data after a 5 mg/kg dose and \( C_{\text{max}} \) is extrapolated from a 2 mg/kg dose.
sedation, constipation and respiratory depression) observed clinically for this compound class (Anonymous, 2007).

4. Predictive value of animal models of inflammatory pain

Pain relief for patients with inflammatory diseases, such as rheumatoid arthritis, is largely based upon the use of non-steroidal anti-inflammatory drugs (NSAIDs). Included in this group are the troubled COX-2 inhibitors; while celecoxib is still marketed for the treatment of pain (Leykin et al., 2007), patient use has radically declined (Bresalier et al., 2005), rofecoxib was voluntarily withdrawn in 2004 (Bresalier et al., 2005) and the FDA rejected etoricoxib in 2007 (Fitzgerald, 2007). As such, we have summarized preclinical and clinical data for celecoxib. In addition, we show data for indomethacin as a prototypic NSAID. Efficacy data for both compounds in the Freund’s complete adjuvant (FCA) model of chronic inflammatory pain with the Randall–Selitto endpoint were generated in-house using methods that are described in detail elsewhere (Valenzano et al., 2005). All in-house pharmacokinetic studies were conducted according to previously described methods (Sullivan et al., 2007).

In inflammatory pain, the efficacious dose, plasma concentration and exposure is under 5-fold higher in human as compared to rats for both celecoxib and indomethacin (Table 2). As discussed for acute pain, the observed differences may be due to issues affecting availability of compounds as well as the inherent difficulties in comparing rat MEDs to clinical maintenance doses in humans. As before, data generated in rodent studies have successfully predicted the occurrence of at least some of the side effects (e.g., gastric lesions) that are observed clinically for these classes of compounds (Anonymous, 2007).

5. Predictive value of animal models of neuropathic pain

Only 5 drugs are FDA approved for the treatment of neuropathic pain. These are the anticonvulsant gabapentin (for post-herpetic neuralgia), the anticonvulsant pregabalin (for post-herpetic neuralgia and diabetic neuropathy), the anticonvulsant carbamazepine (for trigeminal neuralgia), the local anesthetic lidocaine (topically for post-herpetic neuralgia), and the antidepressant duloxetine (for diabetic neuropathy). While these treatments have proven efficacious in controlled clinical trials, substantial improvements are needed due to the limited extent of pain relief, in terms of both the individual and the percent of the population satisfactorily treated (see review, Rice and Hill, 2006). These treatments also have been associated with dose-limiting side effects as discussed elsewhere (see review, Rice and Hill, 2006). In addition to these approved therapies, a number of other drugs are used off-label. These include, among others, opioids (although controversy exists as to their effectiveness) additional anticonvulsants such as lamotrigine, additional antidepressants such as amitriptyline and milnacipran and the calcium channel blocker ziconitide (given intrathecally). This review focuses on a subset of these approved and off-label treatments for neuropathic pain to establish a pharmacokinetic–pharmacodynamic relationship across species. Efficacy data for all compounds in the spinal nerve ligation (SNL) model of neuropathic pain with Randall–Selitto endpoint were generated in-house using methods that are described in detail elsewhere (Leventhal et al., 2007; Valenzano et al., 2005). We focused our analysis on the SNL model of neuropathic pain as this is commonly used for preclinical...
drug screening of compounds for neuropathic pain indications (Iyengar et al., 2004; Jarvis et al., 2002; Sullivan et al., 2007; Valenzano et al., 2005). Alternative neuropathic pain models are available such as partial sciatic nerve injury (PSNL) and chronic-constriction injury (CCI); preclinically, compounds presented in the current review also are efficacious in these models. The PSNL and SNL models correlate very closely (Iyengar et al., 2004; Jarvis et al., 2002; Sullivan et al., 2007; Valenzano et al., 2005). All in-house pharmacokinetic studies were conducted according to previously described methods (Sullivan et al., 2007).

In neuropathic pain, the minimal efficacious exposure in rats for all compounds is between 1- and 15-fold greater than that observed in man (Table 3). Interestingly, the efficacious exposure for gabapentin is almost identical in rat and humans. The efficacious exposure for amitriptyline could not be determined, since no efficacy was observed in the SNL model (highest dose tested was 100 mg/kg, p.o.). Ratios based on plasma concentration confirmed the ratios based on exposure; the concentration ratio for all compounds, except for duloxetine, was 1—12. Duloxetine was the extreme in both cases with an exposure ratio of 15 and a concentration ratio of 37. In contrast to exposure, the efficacious doses for all compounds, except gabapentin and lamotrigine, were more than 15-fold greater in rat as compared to human. Overall, the difference between rat and human is again within 15-fold for exposure. All compounds investigated required higher plasma concentrations and exposures in the rat to achieve efficacy as compared to humans. In addition, the efficacious dose for all but two of the compounds was considerably higher in rat as compared to human. As previously discussed, the observed differences may be due to issues affecting availability of compounds as well as the difficulties comparing MED in rats to a maintenance dose in man. In addition, as the mechanisms of action of these compounds are disparate, comparisons between compounds may be less appropriate.

The data in Table 3 focuses on the relationship between rat and human efficacious plasma concentrations and drug

Table 3
Comparison of the pharmacokinetic—pharmacodynamic relationship of neuropathic pain drugs in rats and humans

A. Exposure

<table>
<thead>
<tr>
<th>Compound</th>
<th>Human daily dose range (mg) [maintenance]</th>
<th>Human maintenance dose (mg/kg)</th>
<th>Rat MED SNL (mg/kg)</th>
<th>Rat exposure (AUC; ng h/ml) at MED</th>
<th>Human exposure (AUC; ng h/ml)</th>
<th>Exposure ratio (rat/human)</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duloxetine</td>
<td>40—120 [60]</td>
<td>0.9</td>
<td>30</td>
<td>8673</td>
<td>584</td>
<td>15</td>
<td>Rat, Wyeth in-house; human, Chan et al. (2007)</td>
</tr>
<tr>
<td>Gabapentin</td>
<td>300—3600 [1800]</td>
<td>26.0</td>
<td>100</td>
<td>146000</td>
<td>125370</td>
<td>1.2</td>
<td>Rat, Radulovic et al. (1995); human, Gidal et al. (1998)</td>
</tr>
<tr>
<td>Lamotrigine</td>
<td>100—500 [200]</td>
<td>2.9</td>
<td>10</td>
<td>208200</td>
<td>69754</td>
<td>3</td>
<td>Rat, Castel-Branco et al. (2004); human, Theis et al. (2005)</td>
</tr>
<tr>
<td>Carbamazepine</td>
<td>100—1200 [1200]</td>
<td>17.0</td>
<td>100</td>
<td>55780</td>
<td>14120</td>
<td>4</td>
<td>Rat, Chen et al. (2002); human, Theis et al. (2005)</td>
</tr>
<tr>
<td>Milnacipran</td>
<td>50—150 [50]</td>
<td>0.7</td>
<td>30</td>
<td>6732</td>
<td>939</td>
<td>7</td>
<td>Rat, Wyeth in-house; human, Puzzo et al. (2006)</td>
</tr>
<tr>
<td>Amitriptyline</td>
<td>10—150 [150]</td>
<td>2.1</td>
<td>100</td>
<td>&gt;2526</td>
<td>3540</td>
<td>Not able to determine</td>
<td>Rat, Wyeth in-house; human, Park et al. (2003)</td>
</tr>
</tbody>
</table>

B. Concentration

<table>
<thead>
<tr>
<th>Compound</th>
<th>Human daily dose range (mg) [maintenance]</th>
<th>Human maintenance dose (mg/kg)</th>
<th>Rat MED SNL (mg/kg)</th>
<th>Rat C\textsubscript{max} (ng/ml) at MED</th>
<th>Human C\textsubscript{max} (ng/ml)</th>
<th>Exposure ratio (rat/human)</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duloxetine</td>
<td>40—120 [60]</td>
<td>0.9</td>
<td>30</td>
<td>1439</td>
<td>39</td>
<td>37.2</td>
<td>Rat, Wyeth in-house; human, Chan et al. (2007)</td>
</tr>
<tr>
<td>Gabapentin</td>
<td>300—3600 [1800]</td>
<td>26</td>
<td>100</td>
<td>32800</td>
<td>11940</td>
<td>2.7</td>
<td>Rat, Radulovic et al. (1995); human, Gidal et al. (1998)</td>
</tr>
<tr>
<td>Lamotrigine</td>
<td>100—500 [200]</td>
<td>3</td>
<td>10</td>
<td>5420</td>
<td>4479</td>
<td>1.2</td>
<td>Rat, Castel-Branco et al. (2004); human, Theis et al. (2005)</td>
</tr>
<tr>
<td>Carbamazepine</td>
<td>100—1200 [1200]</td>
<td>6</td>
<td>100</td>
<td>8550</td>
<td>3320</td>
<td>2.6</td>
<td>Rat, Chen et al. (2002); human, Theis et al. (2005)</td>
</tr>
<tr>
<td>Milnacipran</td>
<td>50—150 [50]</td>
<td>1</td>
<td>30</td>
<td>1678</td>
<td>144</td>
<td>11.7</td>
<td>Rat, Wyeth in-house; human, Puzzo et al. (2006)</td>
</tr>
<tr>
<td>Amitriptyline</td>
<td>10—150 [150]</td>
<td>2</td>
<td>100</td>
<td>&gt;232</td>
<td>109</td>
<td>Not able to determine</td>
<td>Rat, Wyeth in-house; human, Park et al. (2003)</td>
</tr>
</tbody>
</table>

SNL, spinal nerve ligation; MED, minimum efficacious exposure; AUC, area under the curve. All compounds were administered orally except for lamotrigine, which was administered intraperitoneally. Human data are based on a 70 kg body weight. Rat efficacy data were generated in-house at Wyeth. AUC for all compounds is AUC (0— infinity) except for human lamotrigine, carbamazepine and amitriptyline data which are AUC (0—12 h), AUC (0—24 h) and AUC (0—96 h), respectively. Extrapolations of pharmacokinetic data assume linearity. Gabapentin rat exposure is extrapolated from data after a 50 mg/kg dose and human exposure from data after a 600 mg dose. Lamotrigine rat exposure is extrapolated from data after a 20 mg/kg dose. Amitriptyline human exposure is extrapolated from data after a 50 mg dose.
exposures. We can conclude from the table that, first, the SNL rat model of neuropathic pain predicts efficacious exposure in humans and, second, efficacious plasma concentration and exposure in the rat approximates to efficacious plasma concentration and exposure in humans. Data generated in rodent studies, similar to acute and inflammatory pain, have successfully predicted the occurrence of at least some of the side effects (e.g., sedation and ataxia) that are observed clinically for these classes of compounds (Anonymous, 2007). This is particularly noteworthy in the case of amitriptyline, in which efficacy was not observed with acute dosing in rats. This is in line with the clinical situation, in which amitriptyline is dosed chronically with the dose being titrated to reduce severity of side effects (Jose et al., 2007).

6. Failures of animal pain models

The goal of a targeted drug discovery effort is to identify and ultimately develop a small molecule or biologic modulator of a particular disease target. The process involves a number of steps, each of which has a substantial risk of failure. Animal models play key roles at a number of points along this pathway. Commonly, the process begins with target identification; often this can be accomplished by making use of tissue from an animal model combined with molecular techniques to gain an indication that a particular target is implicated in a disease process (Wang et al., 2002). Alternatively, inhibition of the disease process using tool compounds or molecular manipulations, such as knock-out animals or antisense mediated knock-down, can be used to implicate a target or disease mechanism. Following this, in vitro assays are developed; these typically employ construction of recombinant cell lines that express the target receptor of interest or cell-free assays using purified or recombinant proteins. These assays facilitate identification of compounds that interact with or modulate the target/system in the desired fashion. They can then be used as read outs of medicinal chemistry efforts to improve upon desired characteristics such as potency, efficacy and selectivity. In addition, other characteristics such as metabolic stability and desirable pharmacokinetic profile are optimized as part of the iterative synthetic chemistry effort. Following demonstration of efficacy in an animal disease model, a compound can then move into animal models of safety and toxicology before ultimately moving to the clinic.

When discussing “failure” or discontinuation of a drug discovery effort, it is crucial to determine and specify exactly why a directed effort to generate a small molecule or biologic modulator of a particular drug target “failed” to culminate in a marketed drug. There are a number of reasons why drug development efforts may be discontinued; however these failures often are incorrectly attributed to a failure of the target to yield a clinically meaningful effect, implying a lack of predictivity of the animal models. While this scenario does occur in the pain field, most drug discovery discontinuations are due to reasons outside of animal model predictivity particularly true when pursuing unprecedented targets. In contrast, the failure of NK1 antagonists in the clinic is certainly an example of animal models not predicting clinical efficacy. In this case, the compounds were efficacious in animal pain models (Hill, 2000), they demonstrated suitable systemic exposure in humans (Bergstrom et al., 2004), there was sufficient penetration into the central nervous system (Bergstrom et al., 2004) and receptor occupancy studies showed that the compounds gained access to the target (Bergstrom et al., 2004). The conclusion is that the target, although relevant to pain in rodents, is not relevant to pain in man. This is the only published case where such a definitive conclusion can be made.

Let us consider alternative reasons for “failure” at each stage of drug discovery and development. First, during target identification, a target may be incorrectly associated with pain; examples of this are plentiful as exemplified in the numerous reports of fold change in RNA expression level that are correlated with importance to disease mechanism (Wang et al., 2002). Though intuitive one must remember that in animal models and in patients, many disease/model related changes occur that are not necessarily linked to pain. Target validation efforts may suffer from a lack of selectivity of tool compounds or non-specific effects of molecular techniques, again leading to an erroneous association between a particular target and pain. In cases where a target has been “correctly” identified and validated in animals, drug discovery efforts may still be discontinued. Such targets may be valid in humans, and in some cases, there is very good evidence supporting clinical use (Hamilton et al., 2000). Why then are they dropped? The reasons are many and include: an inability to find modulators of the target; inability to develop selectivity versus other targets; the involvement of the target in other systems that leads to unwanted side effects (e.g. the rewarding effects of opioids); species differences in the molecular biology of the target; an inability to design-in adequate pharmacokinetic parameters; an inability to design-in adequate tissue penetration to gain access to the target; on-target or off-target toxicity issues. Finally, conducting clinical trials in an inappropriate patient population or utilizing an unsuitable clinical trial design may result in trial failure and cessation of further investigation, when if a different population or different trial design were used, efficacy may have been revealed.

In the systematic review of neuropathic pain models by Kontinen and Meert (2003) the authors concluded that “The models should not be held accountable for unrelated failures in the drug development process”. Unfortunately, the reasons for discontinuation of a drug discovery effort are neither well investigated nor published, leading to speculative and unsubstantiated claims. We hope to have reiterated the message of Kontinen and Meert and expanded on it by highlighting potential reasons for those failures. Ultimately, animal pain models have proven to be useful both as instruments to teach us about the basic biology of pain in addition to having proven predictive value for drug discovery. Care should be taken to not let a single widely publicized failure cast doubt upon the utility of all animal models of pain. These models are tools that can help prioritize the relative importance of pain mechanisms in different pain states, such as acute, inflammatory, osteoarthritic or neuropathic conditions. Beyond this, they can help
prioritize compounds resulting from drug discovery efforts in order to reduce inherent risk, but an expectation of 100% clinical predictivity is unrealistic. In fact, due to the reasons already discussed, a lower expectation of animal model predictivity should be tolerated as the norm. Even if animal models are only 10% predictive they are still essential tools in the drug discovery repertoire. This same fact is well understood, and the limitations are fully accepted in other fields, such as schizophrenia, anxiety and depression (Castagne et al., 2006; McArthur and Borsini, 2006; Sams-Dodd, 1998). The gap between models and trials may well be filled with translational studies mentioned below. With regard to animal models of pain, the community is avidly working to develop and industrialize improved animal models and endpoints; it is expected that these will be more disease relevant and should further improve clinical predictivity.

7. Improvements in animal pain models

Advances in animal pain model development fall into two distinct categories. The first involves improvements to the models themselves, while the second involves development of additional endpoints. The driving principles of these efforts are that the new models will be more disease relevant and that the novel endpoints will be more reflective of pain in patients. Novel disease relevant animal models include those for osteoarthritis pain, such as intraarticular monoiodoacetate (Bove et al., 2006; Pomonis et al., 2005), post-surgical pain, such as plantar incision (Brennan et al., 1996), painful cystitis, such as intraperitoneal cyclophosphamide (Wantuch et al., 2007) and pain due to bone metastasis, such as intratibial or intrafemoral implantation of tumours (Medhurst et al., 2002; Schwei et al., 1999). In addition to these models having obvious face validity, published pharmacology reports are beginning to appear utilizing these models in the evaluation of both reference compounds and novel treatment strategies (El Mouedden and Meert, 2007; Fernihough et al., 2004; Ghilardi et al., 2005; Hamamoto et al., 2007; Whiteside et al., 2004).

Considering endpoint development, researchers are starting to investigate pain-related behaviors that may be more indicative of clinical pain. These often involve interrogating supraspinal mechanisms and can be viewed as the preclinical researcher asking the rat how much pain they feel. Specific examples of these behaviors include, feeding (Negus et al., 2006), sleep (Andersen and Tufik, 2003), rearing (Matson et al., 2007), locomotion (Negus et al., 2006), analgesic self administration (Colpaert et al., 2001), weight bearing (Pomonis et al., 2005; Whiteside et al., 2006), alteration in gait (Coulthard et al., 2003), grip strength (Kehl et al., 2000) and behavioral methods for quantifying the affective component of pain (Johansen et al., 2001; LaBuda and Fuchs, 2005; Pedersen et al., 2007). As changes in many of these behaviors are documented in pain-affected patients (McWilliams et al., 2003; Williams et al., 2006) and contribute to decreased quality of life and social functioning, it is reasonable to assume that these models will have utility and improve the overall predictive validity. Vierck (2006) makes a strong argument for operant models while strongly criticizing stimulus—response assays, which he claims to be intrinsically flawed and neither sensitive nor specific predictors of efficacy in humans. This view should be balanced with that of Campbell and Meyer (2006) who remind us that while stimulus response assays have a motor component that cannot be dissociated from a pain response, operant assays have a motivational component that cannot be dissociated from an analgesic effect. It is likely that once fully characterized, these kinds of novel endpoints will augment well-established stimulus—response assays rather than replace them. Recognizing the limitations of both animal models and endpoints should lead to their use in a rational, integrated manner, a strategy that is most likely to yield the richest information and best inform clinical decisions.

In addition to advances in both animal models and endpoints, improvements in the clinical arena should also advance the pain field toward discovering and developing effective therapies. First, the use of human surrogate models may allow the more efficient rejection of compounds with pharmacokinetic or tissue penetration problems and, in some cases, such as with capsaicin, may allow us to conclude an on-target effect. In the same way, development of imaging technologies in combination with compound, mechanism and ultimately pain-relevant, biomarkers will further reduce the risk of progressing compounds while increasing fundamental knowledge of pain mechanisms. Initiatives aimed at standardizing sensory testing in the clinic (Rolke et al., 2006), analogous to what largely was achieved in the preclinical setting, should allow more direct comparisons between efficacy measures in rodents and humans. In addition, it would be advantageous to have agreed-upon descriptors of efficacy that are consistent between preclinical and clinical studies. Finally, expanded and more in-depth collaboration between preclinical, translational, and clinical scientists, should improve our use of existing pain models, facilitate development of new models and may also help the interpretation and understanding of clinical findings.

8. Conclusions

From the analysis presented here we conclude that, overall, the rat predicts efficacious drug exposure for specific analgesic mechanisms, across models of acute, inflammatory and neuropathic pain for the clinically relevant compounds investigated here. The authors consider a ~10-fold difference between species (rat to human) indicative of predictive utility, however, a shift of this magnitude within a species would be considered problematic. Although there was considerable variation between individual compounds, it is worth noting that for some compounds very close correlations were observed. It is also worth noticing that for some compounds the correlation was not close (50-fold being the worst ratio) and that in some cases comparing efficacious exposures gave close correlations while in other cases comparing efficacious plasma concentrations gave a better correlation. Taking into account that comparisons were made between rat MEDs and human maintenance doses, likely differences that exist in both drug concentration at target and the efficacy measures employed, we
consider the overall correlation across three very different animal models to be very encouraging. While this analysis is inherently biased in that we chose compounds based on efficacy in both rat models, and humans, we feel this analysis is revealing and is the first of its kind. We also hope this review will stimulate further testing of these models using different compounds and under conditions not yet evaluated so that we can further understand their validity, and most importantly, where they work and where they are inappropriate so that they can be more appropriately applied and utilized. In addition, advances in animal models and endpoints, as well as improved clinical trial design and use of earlier stage translational studies should improve the predictive validity of animal pain models, though this is unlikely to ever reach the 100% ideal.

Acknowledgements

The authors wish to thank Rob Allen and Jeff Kennedy for critical reading and discussion of this manuscript.

References


