Review

Validation of experimental medicine methods in psychiatry: The P1vital approach and experience

Gerard R. Dawson*, Kevin J. Craig, Colin T. Dourish

P1vital, Department of Psychiatry, University of Oxford, Warneford Hospital, Oxford OX3 7JX, UK

1. Introduction

Munos recently observed that the rate of production of new drugs by pharmaceutical companies has been constant since the 1950s and nothing that has been done in the past 60 years has affected this rate [1]. Companies have produced new molecular entities (NMEs) at steady rates whether they are large or small organisations, focused on small molecules or biologics or operating in the twenty-first century or in the 1950s. By contrast, in the last 25–30 years both academia and industry have invested heavily in biomedical research, much of it directly aimed at elucidating the mechanisms underlying brain function. Consequently, our knowledge of brain systems and functions has vastly increased. New experimental techniques have been developed and a wealth of new drug targets aimed at improving the treatment of psychiatric disorders has been produced. Unfortunately, the clinical translation of these basic research findings to the benefit of psychiatric patients has been disappointing to date. This poor rate of return on

* Corresponding author. Tel.: +44 1865 247934.
E-mail address: gdawson@p1vital.com (G.R. Dawson).

0006-2952/$ – see front matter © 2011 Published by Elsevier Inc.
doi:10.1016/j.bcp.2011.03.013
investment has prompted a number of pharmaceutical companies to decrease their focus on psychiatry as a therapeutic area and two major players, GlaxoSmithKline and AstraZeneca [2], have recently announced that they will stop basic research aimed at developing new treatments for psychiatric disorders. This has occurred despite the fact that there is still a significant unmet need for new and more effective treatments for psychiatric disorders with an enormous market potential of tens of billions of dollars (see Table 1). The rationale for pharmaceutical companies to reprioritise or withdraw from drug development in psychiatry is largely based on two factors, the low probability of success (POS) of developing a new drug for CNS disorders and the escalating costs. The latest assessments of the costs of drug development (including lost opportunity costs) show that they continue to increase rapidly with the average cost of development for a new drug now estimated to be greater than $1 billion [1]. In addition, the POS of a novel compound at the Phase 1 clinical trial stage for the treatment of a CNS disorder is lower than the average across all therapeutic areas at only 8% [4]. Therefore it is important to consider why it appears to be more difficult to discover and develop drugs for psychiatric disorders in addition to why the success rate for drug development is stagnating.

In the clinical development of new drugs the principal cause of failure is insufficient efficacy and safety, accounting for approximately 30% and 20% respectively of all failures [4]. This is a particular problem in CNS drug development which has a lower than average POS due to the poor translation from preclinical models to clinical efficacy. As noted above, in the preclinical arena both academia and industry have contributed to a significant increase in fundamental knowledge and new classes of candidate compounds for the treatment of psychiatric disorders. However, clinical translation of these findings to benefit patients has to date been very limited. Almost all of the medicines we think of as new, atypical antipsychotics for example, were conceived as refinements of previous treatments. The failure, to date, to translate the innovations of molecular biology to the clinical arena has been disappointing. For example, preclinical and early clinical studies with novel putative antidepressants such as neurokinin NK1 antagonists have often identified promising trends in preclinical and early clinical studies that have not been confirmed by the results of subsequent Phase 3 clinical trials [7]. Unfortunately, placebo controlled trials are difficult to conduct in patients with the type and degree of depression that most requires pharmacologic intervention. What remains are populations for study in which placebo response rates are high thereby confounding detection of positive treatment effects. Attempts to compensate for poor signal detection by increasing sample size has often led only to very expensive failed trials. Thus, there is a growing gap for many Pharma companies between their pre-clinical portfolio and willingness to invest in such large scale clinical studies. In the
absence of any clear successes, it is very difficult to calculate the costs of development for truly novel CNS drugs, a fact which can have a chilling effect on those in pharmaceutical companies who must decide on whether to risk the unknown. Consequently, both the lack of translation from promising preclinical data to clinical efficacy and the rising cost of drug development have highlighted an evident gap between pre-clinical promise and failure in large scale clinical trials. In addition, this has occurred at a time when the number of compounds entering the clinic has quadrupled due to the increasing introduction of robotic screening technology during the past 15 years.

It is increasingly recognised that the introduction of experimental and translational medicine models at the interface between Phase 1 and Phase 2 clinical trials will be crucial for future success in psychiatric drug development. Experimental medicine studies are essential to bridge the gap between animal and human studies and to help mitigate and manage the risk in this particularly difficult area of drug development. While this strategy commands a broad consensus, the tactics to enable its development are in their infancy. If, using their own limited resources, any one pharmaceutical company were to attempt to develop and validate experimental medicine models as effective decision making tools for the development of new drug treatments, progress would be slow and very expensive. The fragmentation implied by this analysis is counter-productive at a fragile stage in the development of this new approach. Moreover, expertise in experimental medicine resides largely in academic centres that have access to volunteers and patients to develop such models [8,9].
P1vital’s solution was to initiate a coordinated effort across academia and industry by establishing a pre-competitive consortium to develop technology in this key area of clinical drug development. The objective was to develop and validate human models that would help to predict whether a novel drug is likely to be effective in the intended patient population. These validated experimental medicine models would then be made available by P1vital to the pharmaceutical industry to incorporate into their early drug development programs.

The pharmaceutical industry has a history of joining forces to develop technology in key areas of drug discovery [10,11]. These pre-competitive consortia are established with clear aims and objectives that are determined and funded by the Pharma consortium members. The advantages to the members are significant: at a fraction of the cost new experimental methods can be developed than can be achieved using in-house resources alone. There are other obvious advantages such as economies and efficiencies of scale, coordination and funding that a pre-competitive consortium brings. In addition, in the P1vital consortium, additional, less obvious, advantages have become apparent such as the greater understanding that is achieved by industry and academic groups through interacting with each other in the pursuit of a common goal and the increased cooperation between different academic groups that normally compete with each other for scarce resources. To achieve success the consortium approach requires a clear management framework to ensure that the projects deliver results that can inform decision-making in drug development. This approach also requires vision and a long term commitment, as a single experimental medicine study can take up to 18 months to complete and a series of experiments (including positive and negative controls) may require 3–4 years.

2. The establishment, aims and objectives of the P1vital pre-competitive consortium in CNS experimental medicine

Bringing together diverse academic groups and representatives from Pharma for initial consortium meetings was a challenging process. During the early months of 2005 P1vital visited a number of academic centres and Pharma companies to present their vision, aims and objectives for a pre-competitive consortium in CNS experimental medicine. As a result representatives from seven Pharma companies opted to join the initial meeting held in Merton College, Oxford in September 2005. At that meeting academic groups from the University of Bristol, Cardiff University, the Institute of Psychiatry, the University of Manchester, the University of Oxford and Newcastle University, presented preliminary work on experimental medicine models in anxiety, schizophrenia, cognition and depression. During these presentations one of the early and enduring benefits of the consortium discussions that emerged was the wealth of experience provided by Pharma attendees who had extensive experience with experimental medicine models in psychiatry. Based on literature reports of successful results with licensed drugs in experimental medicine models, a number of companies had initiated trials with proprietary compounds that included a licensed drug as a comparator. In many instances there was a failure to replicate the published results with the comparator drug. Such failures were reported at the meeting to be a common occurrence and in a number of instances, different companies, unknowingly, had collaborated with the same research teams and had each failed to replicate the published results. Unfortunately, it emerged that these negative results are rarely published leading to unnecessary duplicated work without advancing the field. Consequently, these discussions further underlined the need for a more systematic approach to validating experimental methods in psychiatry where the positive effects of NMEs observed in preclinical models are often not predictive of their effects in patients. Moreover, the various lessons learned from these collective experiences served the consortium well as they provided a strong foundation on which to base the principles for the future validation of the experimental medicine models proposed by the university groups.

Once the proposed studies had been reviewed, the meeting attendees discussed and agreed principles for the validation of experimental medicine models. It was proposed that both positive and negative controls should be included in the validation studies and that the effects of these drugs on biomarkers in healthy volunteers or surrogate populations should be, where possible, replicated in patients. The questions of when a model would be considered to be validated and which models should be validated were discussed at length during this initial meeting and subsequent consortium formation meetings held in March, September and November 2006 (which were attended by representatives of a number of additional Pharma companies). Although it proved difficult to agree a definitive answer, it was generally accepted that an experimental medicine model would be considered to be validated for use with proprietary compounds when the following criteria are met:

(a) sensitive to currently available drugs;
(b) sensitive to drugs with different mechanisms of action;
(c) insensitive to one or two negative controls; and
(d) results are replicated in patients.

The consortium formation meeting discussions also focused on the process for validation of experimental medicine models. P1vital and the Pharma attendees agreed that the long-term strategy for the development of efficacy biomarkers should follow an established pattern similar to that of drug research and development. Early exploratory work would identify promising candidate efficacy biomarkers, surrogate populations or combinations of both, based on available evidence or published research. Once candidate efficacy models show promise in early validation
of the cognitive abnormalities associated with the disease. If the underlying hypothesis is correct, then individuals with high schizotypy may provide a face-valid, genetically and mechanistically related and rapidly accessible surrogate group to determine the effects of novel treatments for schizophrenia.

The primary objective of the first schizophrenia study (Schiz 1) was to validate a series of biomarkers in individuals with high schizotypy and determine the effects of two antipsychotics (amisulpride and risperidone) and a putative cognition enhancer (nicotine) on their cognitive performance. In addition, we also determined the feasibility of recruiting high schizotypes in sufficient numbers to support proprietary compound studies and whether these studies could be successfully executed as a single study across three academic centres.

Participants were recruited from three UK sites: the University of Manchester, the Institute of Psychiatry (King’s College) in London and Cardiff University. Recruitment was via an online questionnaire measuring schizotypy the Schizotypal Personality Questionnaire (SPQ) in its short (SPQ-B [14]) and full version (SPQ [15]). In a nine month period, 13,275 people completed the full version of the online SPQ and a further 9098 filled out the short version, SPQ-B. From these responses approximately 800 participants were telephone screened from which 244 participants, 97 in Manchester, 83 in London and 64 in Cardiff were selected to take part in the study. The high and average schizotypy groups consisted of 122 participants each. Of these, 59 participants were randomized to placebo, 62 to amisulpride (400 mg), 62 to nicotine (7 mg contained in a patch) and 61 to risperidone (2 mg) treatment groups.

The initial pattern of results from these studies provide support for utilising the high schizotypy group in further validation studies and, if successful, in clinical trials with proprietary compounds. Much of the data from this large study is still being analysed. However, preliminary results suggest that the biomarkers used in this assessment are sensitive both to the effects of high schizotypy and its modulation by antipsychotics, particularly amisulpride which, for example, tended to disrupt saccade inhibition in controls but improve performance in high schizotypes [16]. The second schizophrenia study (Schiz 2) was initiated in early 2010 with the primary objective of validating the same biomarker set employed in Schiz 1 to assess the cognitive deficits of healthy volunteers infused with ketamine and of patients with schizophrenia. The results obtained from Schiz 1 and Schiz 2 will be compared to determine whether schizotypy or ketamine infusion

3. P1vital consortium studies in schizophrenia

Increasing evidence suggests schizophrenia spectrum disorders lie at the extreme of a normally distributed trait of psychosis proneness manifested as a personality trait called ‘schizotypy’ in the general population [12]. There is a growing literature on the high schizotypy group that suggests high schizotypy may provide a good ‘model’ for schizophrenia. For example, this group shares many of the features found in patients with schizophrenia, albeit in an attenuated form, including cognitive deficits, positive symptoms, negative symptoms and anhedonia [13]. Schizotypy is also related to risk factors for psychosis. Research at the University of Manchester and Cardiff University confirmed that healthy participants with high schizotypy experience isolated psychotic symptoms and show some
best models the cognitive and negative symptom deficits in schizophrenia and whether particular biomarkers are more sensitive when used in one model or the other. While we await the results of these initial validation studies it is clear from recent independent market research conducted by P1vital that the development of drugs for the treatment of cognitive deficits and negative symptoms in schizophrenia remains a priority for many Pharma companies and consequently is likely to be a significant focus of the next P1vital Consortium which is currently in the planning stage (see Figs. 2 and 3).

4. P1vital consortium studies in anxiety

Generalised Anxiety Disorder (GAD) is a relatively under-researched condition, and much still remains to be learned about its causes and treatment. The 7.5% carbon dioxide (CO₂) challenge assay has been proposed as a human model of GAD that reliably reproduces GAD symptoms in healthy participants [17]. Breathing an air gas mixture containing 7.5% CO₂ for 20 min significantly increases self-rating scores of anxiety and other associated symptoms, such as tension and worry. A number of benzodiazepines, such as lorazepam (2.0 mg) reliably reduce these symptoms [18]. The primary objective of the anxiety study (Anx 1) was to determine whether the CO₂ model is sensitive to two non-benzodiazepine drugs with a label for the treatment of GAD; venlafaxine (150 mg) and pregabalin (200 mg). The results from this study suggested that chronic treatment (three weeks) with venlafaxine and pregabalin at therapeutic doses did not reliably reduce the anxiety induced by breathing 7.5% CO₂ or the panic-like symptoms induced by a single breath of 35% CO₂. However, a recent Phase 1B study conducted by P1vital with a proprietary CRF receptor antagonist showed that this compound was effective in the same 7.5% CO₂ assay [19], suggesting that the model is sensitive to non-benzodiazepine compounds that reduce acute anxiety. This model is now considered to be sufficiently validated and further consortium studies are not envisaged. Interestingly, independent market research conducted by P1vital indicates that the development of novel treatments for acute anxiety is not currently a priority for the Pharma industry (see Figs. 2 and 3).

5. P1vital consortium studies in cognitive disorders

In collaboration with researchers at the Institute of Psychiatry, P1vital has employed virtual reality techniques to develop human analogues of two well characterised animal models of spatial learning and hippocampal function, the Arena (Morris water maze) and Platform (Olton radial maze) tasks. By combining virtual reality and fMRI it has been shown that the Arena task reliably activates the hippocampus in a predictable way [20]. In this task both age and the muscarinic receptor antagonist, scopolamine, reduce hippocampal activation and impair spatial learning. Interestingly, age (>65 years) induces a markedly greater behavioural deficit than scopolamine [21]. To further characterise and validate the Arena and Platform tasks, the objectives for the first cognition study (Cog 1) were to determine the behavioural performance and brain activations of patients with mild cognitive impairment (MCI) performing the tasks. An important secondary objective was to determine the feasibility of recruiting participants with MCI in sufficient numbers to support proprietary compound studies. The MCI group was of interest because of the relatively intact nature of neuropsychological functioning that enables sensitive neuropsychological tests to be used and because many people with MCI go on to develop dementia. In common with a number of similar studies we have found the recruitment of MCI patients to be very challenging. Few individuals present for assessment and treatment early enough to be classified as having MCI and many who attend for the first time meet criteria for mild Alzheimer’s disease on formal testing, despite sustaining a moderate level of functioning. There has also been a great deal of research interest in the MCI group and therefore there is significant competition between researchers for a relatively small pool of potential participants. Therefore, it seems unlikely that this group will provide a rapidly accessible surrogate population for testing the effects of new cognition enhancing compounds.

In the second cognition study (Cog 2) we determined the behavioural performance and brain activations of high and average schizotypes (see above) in the Arena and Olton mazes. We found that although the groups performed equally well on both tasks there were significant differences in brain activations between high and average schizotypes. Consequently it is possible that when the cognitive processes of high schizotypes are challenged they may not be as efficient as those of controls and deficits may be detectable [22].

6. P1vital consortium studies in depression

The validation of biomarkers for emotional bias in depression would ideally be carried out in depressed patients as they may be more sensitive to drug-induced effects. However, their recruitment through primary care centres is extremely slow and therefore, although this type of patient recruitment is possible [23], it is not feasible to execute an experimental medicine study in a timely manner using such an approach. The rationale for the first depression study (Dep 1) was to determine whether it is possible to recruit participants with dysphoria, that is participants with low mood, but not necessarily clinically depressed, at a rate that would be feasible for experimental medicine studies. A further aim was to determine whether dysphoric individuals are more sensitive to emotional stimuli than healthy participants and thereby provide a better signal to noise ratio for studies with proprietary compounds.

The results from the Dep 1 study are encouraging. Firstly using novel internet recruitment techniques we were able to recruit dysphoric participants, many of whom met the criteria for major depressive disorder (see Table 2) at a much faster rate than was observed for the recruitment of depressed patients through primary care centres. Secondly the pattern of behavioural and fMRI results was similar to that reported in studies that included more severely depressed participants. Thirdly the negative biases in dysphoric participants were particularly apparent in fMRI models of emotional processing and involve hyperactivity of the
amygdala and decreased prefrontal cortex responses to negative vs positive stimuli [24]. These findings validate the use of dysphoric participants for future investigation of novel antidepressant drug effects and the second, multi-centre consortium depression study (Dep 2) to compare the effects of citalopram in non-dysphoric and dysphoric participants is ongoing. This is, to our knowledge, the first multicentre randomized controlled trial with fMRI signals as the primary endpoint.

7. Initial findings and lessons learned

Pre-competitive consortia have many advantages and a number of findings and observations have emerged during the first three years of the P1vital consortium which may benefit any future consortia. The validation of experimental medicine models in psychiatry is by necessity a time consuming process. Biomarkers and clinical efficacy assessment methods need to be validated in placebo controlled trials with both positive and negative controls in healthy volunteers and in studies with patients. The positive controls used are generally marketed drugs for the therapeutic indication under investigation. In contrast, negative controls may be compounds that showed promising evidence of efficacy in early clinical trials but which failed in large scale clinical trials or drugs that are marketed for a different indication. For example, the component tasks of the P1vital [18] Oxford Emotional Test Battery (ETB) for detecting potential antidepressant efficacy have been validated over a period of more than 10 years in studies with both positive controls [25–35], and negative controls [36–38] in healthy volunteers and in patients with depression [23]. During that time the tasks have remained relatively unchanged allowing comparison between results over an extended period. However, the time horizons of industry are somewhat shorter as areas of scientific focus can change relatively rapidly. For example, since the P1vital pre-competitive consortium was initiated in May 2007 the relative importance of psychiatry research has reduced for two of the consortium members (AstraZeneca and GSK) [2] and two other consortium members (Wyeth and Organon) have been acquired. Therefore, maintaining continuity and completing the model validation process can be challenging and requires careful planning. The first P1vital pre-competitive consortium was planned to be completed over a period of three years and to deliver its first results after 18 months. Thus while recognising that a validation can be lengthy P1vital adopted a balanced approach of completing the validation of two extensively investigated experimental medicine models (the ETB and the CO2 model of anxiety), while initiating the standardisation and validation of the schizotypy and dysphoria models. A potential constraint of this approach is that two cycles of consortium funding may be required for sufficient validation of a particular efficacy biomarker, method or phenotype to enable the evaluation of proprietary compounds with confidence in the predictive nature of the results.

8. Conclusions and plans for further consortia

Although much of the data from the initial consortium studies is currently being analysed, the initial results from the P1vital consortium were sufficiently encouraging to prompt the Pharma members to ask P1vital to investigate the feasibility of establishing a second consortium. To determine the key CNS therapeutic areas of focus for a second consortium, P1vital commissioned an independent market research company to seek the views of Pharma company representatives. The respondents in this survey included individuals from consortium members and non-member companies who had previously expressed an interest in the P1vital consortium. The survey results suggested that the next consortium should focus on three main disease areas, schizophrenia and associated negative and cognitive deficits, depression and bipolar disorder, and cognitive improvement (see Figs. 2 and 3). It is also clear from this market research that a consortium approach to validating experimental medicine models in pain, including both neuropathic and non-neuropathic pain would be widely supported.

Finally, the efforts of the consortium in validating experimental medicine efficacy models would be of little value if the validated models were not made available for both the testing of proprietary compounds and further academic research. Although the consortium is only in its third year a number of proprietary compound studies have already been conducted by P1vital at university and hospital sites to the standard required by Pharma for Phase 1 trials. At UK universities where Clinical Research Facilities (CRFs), essentially research hotels providing all of the clinical facilities necessary to conduct early phase trials are available, experimental medicine studies have progressed smoothly both from a participant and researchers perspective. However, such CRF facilities currently exist in only a few of the large clinical research centres in UK. Consequently, the next challenge in the development of experimental medicine models will be the creation of sufficient clinical research facilities with expertise in neuroimaging and cognitive testing to cope with the increasing demand for proprietary compound studies. When more of these CRF facilities come on stream experimental medicine studies can become routine and an integral part of drug development programs in psychiatry, thereby accelerating new and more effective treatments for the patients who so desperately need them.

References


Table 2

<table>
<thead>
<tr>
<th></th>
<th>Non-dysphoric</th>
<th>Dysphoric</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>24</td>
<td>24</td>
</tr>
<tr>
<td>BDI score</td>
<td>0.33 (1.0)</td>
<td>19 (10)</td>
</tr>
<tr>
<td>HAM-D score</td>
<td>0.13 (0.4)</td>
<td>12.5 (6.4)</td>
</tr>
<tr>
<td>Met criteria for MDD</td>
<td>0</td>
<td>15 (62.5%)</td>
</tr>
</tbody>
</table>

Data in the table are shown as mean with standard deviation in parentheses.


