Early Decision-Making in Drug Development: The Potential Role of Pharmaco-EEG and Pharmaco-Sleep

Frederick J. Wilson a Philippe Danjou b

a Medical Imaging and Physiological Measurements Consultant, Canterbury, UK; b Biotrial, Rennes, France

Key Words
Electroencephalography · Quantitative electroencephalography · Pharmaco-EEG · Pharmaco-sleep · Polysomnography · Central nervous system · Drug development · Pharmacokinetics · Pharmacodynamics · Biomarkers · Translatable electroencephalography biomarkers

Abstract
The pharmaceutical industry has been suffering from low clinical success rates of new drugs for some time with particularly high attrition in early clinical development, especially for drugs aimed at central targets. Both pharmaco-electroencephalography (EEG) and pharmaco-sleep, along with other biomarker techniques, have significant potential to assist with this problem by enabling early decisions to be made about the likelihood of a compound proving successful in the clinic. This paper discusses the role and points of application of biomarker techniques in early drug development. It proposes a framework for the use of pharmaco-EEG and pharmaco-sleep in drug development that (i) relies on the combination of preclinical data and an understanding of translatability to generate robust hypotheses for testing in early clinical studies and (ii) is backed up by a clear decision-making process. The areas that need further development before this framework can be put fully into practice are discussed, along with some possible routes by which this could be achieved through precompetitive co-operation within the industry.

Introduction

The pharmaceutical industry has been facing considerable challenges over the last 20 years or so with falling approval rates for new drugs adding to economic pressures as governments around the world try to contain costs within the health care budget. As demonstrated by recent research and development site closures and layoffs across the industry, the situation has become unsustainable, with revenues from the number of new drugs reaching the market being too low to sustain previous levels of research and development spending that have simply not been delivering enough new products.

The greatest problem is the very low success rate in clinical development, from first-in-human to registration, exacerbated by the fact that the majority of this attrition occurs late in the pipeline, when considerable amounts of money have already been spent on large clinical trials. Recent analyses have shown that the overall clinical success rate for new drugs is around 11–13%, but there is considerable variation between therapeutic areas.
with central nervous system (CNS) indications faring particularly badly at approximately 8% [1, 2]. Lack of efficacy is the most common cause of failure in both phase 2 and phase 3 [3, 4]. Also, the failure rate due to lack of efficacy has not improved over time, unlike failure due to pharmacokinetic (PK) or bioavailability problems, for example, which decreased 4-fold from 1991 to 2000 [1], representing significant advances in both biological understanding and chemistry.

One possible approach to improving the situation is to determine much earlier in the development pipeline which drugs are not destined to succeed and to terminate these projects before they enter large clinical studies, thus significantly reducing the overall development cost. Such an approach may also enable resources to be better focused on drugs in which there is a higher confidence of success, potentially accelerating their development and bringing them to market earlier, with clear advantages both for patients and pharmaceutical company revenues.

The purpose of this paper is to review the role of biomarkers in general, highlighting the roles of pharmacoelectroencephalography (EEG) and pharmaco-sleep in particular, in making such early decisions about the likelihood of success of a new drug.

**The Role of Biomarkers in Early Drug Development**

The three key points during the early part of the drug development process at which biomarkers can be employed to address different questions are illustrated in figure 1. Firstly, preclinical studies can be carried out with the aim of providing predictions as to clinical efficacy/success and/or to enable the development of a hypothesis that can be tested in an early clinical study. Secondly, the clinical plan can include biomarkers, with the aim of providing confirmatory clinical data at the earliest possible stage. These measures may either be included in one of the phase 1 safety studies or in a separate biomarker study when there are practical limitations or the required statistical power cannot be achieved within a typical first-in-human design. The aim of this is to enable the likelihood of success to be determined so the project can be fast-tracked, proceed normally or be terminated. Ideally, such an approach will enable identification of compounds likely to be unsuccessful so that their development can be terminated prior to the first study in patients (phase 2a). As success in phase 2a would typically result in proof of concept (PoC) being declared, this can be broadly termed as aiming to ‘de-risk PoC’. Finally, biomarkers can be used to enhance patient selection in the phase 2 study, with the aim of selecting only those most likely to respond favourably to the drug, hence enhancing the likelihood of success or reducing the size and cost of the PoC study. In this case there is the possibility to include a wider target population in later larger studies, once efficacy has been demonstrated in a specific group.

The use of surrogate EEG-based biomarkers in preclinical experimentation in order to provide predictions of likely clinical efficacy is covered in a review by Leiser et al. [5]. Whilst scientifically interesting and worthy of further investigation, the validity of this approach in drug development remains in question, largely due to past failures which are only recently becoming understood [6].

Similarly, the use of EEG-based techniques for diagnosis and patient selection is covered in detail for a wide range of disorders. There is clear value to such techniques in drug development, for example in separating distinct pheno-
types as part of patient selection for a phase 2 study (for a famous example, see the reviews by Barry and colleagues [7, 8] of the use of EEG and event-related potential techniques in attention deficit hyperactivity disorder), although development and validation are needed on a case-by-case basis.

A more immediately applicable approach is to use appropriate pharmacodynamic (PD) biomarkers to increase confidence at the point of preclinical to clinical transition, with the aim of reducing the risk of later PoC failure due to poor PK-PD characteristics or failure to modulate the pharmacological target. Whilst this may seem simplistic, a recent retrospective review of 44 drug development projects at Pfizer [9] found that such an approach was highly predictive of the likelihood of success in phase 2. In particular, each project was assessed to determine the availability of evidence for (i) exposure at the target site of action, (ii) binding to the pharmacological target and (iii) expression of pharmacology. Based on these principles, it was concluded that in only 24 out of 44 cases had the mechanism been tested robustly, meaning that confidence was high that the target had been pharmacologically modulated sufficiently to test for efficacy and hence that a negative result would confirm that the target was inappropriate for the indication and should not be pursued further. More strikingly, of the 14 compounds for which there was clear positive evidence against all 3 criteria, 12 achieved a positive PoC result, whereas for the 12 compounds for which there was no, or only incomplete, evidence against any of the criteria, all were negative at PoC.

A PK-PD-based approach thus has a potentially important part to play in early decision-making and hence reducing attrition. From the perspective of using pharmacoeffect and pharmaco-EEG in the early development of drugs aimed at targets in the CNS, it also has the advantage that an empirical approach can be used (as discussed in the section on a framework for decision-making below), thus avoiding difficult questions about whether an EEG change is a true surrogate marker for efficacy. Another way to think about this is to consider the simplified scheme of neuroanatomical structures and neurotransmitter systems comprising the homeostatic system that regulates electrical brain rhythms, as proposed by John and Prichep [10]. The complexity of even this simplified scheme amply illustrates the problem with a detailed interpretation of EEG changes as predictors of efficacy. However, an alternative interpretation is to conclude that EEG is likely to be sensitive to pharmacological modulation of all the major neurotransmitter systems and hence to provide a useful PD signal, but that the observed effects will often be downstream due to the complexity of the neurological networks, and hence the exact mechanism is hardly ever known. The complexity of the homeostatic processes also means that the effect size may vary for different mechanisms (for example, EEG has a high sensitivity to the effects of benzodiazepines [11] and scopolamine [12] but a lower sensitivity to the selective serotonin re-uptake inhibitors [13]) and also in different vigilance states (for example, the 5-HT3 receptor antagonists have effects on the EEG during sleep [14] but not during the daytime [15]). Also, in some cases, such as for 5-HT2a antagonists [16], sleep architecture changes may be more sensitive than the spectral EEG end points. Therefore an approach in which EEG and sleep changes are considered as intermediate biomarkers that reflect pharmacological changes in the brain, rather than as surrogate efficacy biomarkers, is likely to be more successful at enabling robust decision-making.

Finally, in addition to being scientifically robust, an approach that involves additional work must be economically viable. Wise and Preston [17] used an economic model-based approach to address this question, concentrating in particular on the case of adding a functional magnetic resonance imaging (fMRI) study to the project between phase 1 and phase 2, exactly as shown in figure 1, although a similar argument can be applied to other biomarker technologies. Their analysis assumed that the biomarker study acted as a ‘filter’ to remove compounds with little chance of success prior to the phase 2 study, whilst the overall success rate to market remained unchanged. Overall, including the fMRI study reduced the overall cost of bringing one drug successfully to market.

The Biomarker Technologies’ ‘Toolbox’

When choosing a suitable biomarker technique for use in a particular drug development project, it is important to see the range of available techniques as complementary and forming a ‘toolbox’ from which the appropriate modality, or modalities, can be selected for a given situation. Table 1 summarises the characteristics of the main imaging- and physiological measurement-based biomarkers available for use when developing a drug that putatively acts on the CNS. Broadly, these can be divided into the three groups below.

**Target Binding**

Receptor occupancy positron emission tomography (RO-PET), using either a labelled drug or an appropriate ligand for the receptors being targeted, is unique in that it can provide a direct measure of molecular target engage-
ment but it provides no information on the activation of that target. An obvious limitation is the availability of an adequate ligand for human use and the costs associated with such development. Correct interpretation of the results also requires knowledge of the level of target occupancy that is adequate for a therapeutic effect. For innovative mechanisms of action, this must be extrapolated from animal data and is not always clear-cut. For example, before clozapine was demonstrated to be efficacious, a D₂ receptor occupancy of less than 60% would not have been considered therapeutically useful, and similarly, before rimonabant, a quasi-complete block of the CB₁ receptors was considered necessary for efficacy.

**Metabolic Evidence of Pharmacological Expression**

Fluorodeoxyglucose positron emission tomography (FDG-PET) measures glucose metabolism in different areas of the brain and hence can provide indirect evidence of a drug effect on neuronal activity. Similarly, fMRI can detect neuronal activity, and pharmacologically induced changes thereof, by measuring changes in blood flow and/or oxygenation in the brain.

**Neuronal Evidence of Pharmacological Expression**

Magneto-encephalography and EEG offer the most direct markers of neuronal activity available non-invasively by directly measuring, respectively, the magnetic and electric fields resulting from a coherent simultaneous activity of groups of neurons. Polysomnography (PSG), which assesses sleep, is based on such direct measures (including EEG, electromyography and other biosignals), but the final output consists of sleep staging information, derived by applying scoring rules, such as those published by the American Academy of Sleep Medicine [18].

Each of the techniques described has specific advantages and disadvantages. The imaging techniques (RO-PET, FDG-PET and fMRI) have high spatial but low temporal resolution, whereas magneto-encephalography and EEG have low spatial but high temporal resolution, and PSG produces easily interpretable derived end points that can be highly clinically relevant. Techniques involving task-based activation of specific neuronal networks produce results that are highly task specific. Finally, it is important to note that EEG has the unique advantage that the same methodology can be applied both preclinically and clinically (by contrast, rodent imaging generally requires the animal to be anaesthetised, thus confounding interpretation even when the imaging technique is similar), thus giving the promise of excellent translatability for both EEG and sleep end points.

### An Empirical Framework for Decision-Making Using Pharmaco-EEG and Pharmaco-Sleep

In order to use an empirical PK-PD-based biomarker approach in drug development, a robust framework is needed to define criteria for choosing the appropriate biomarker and making decisions based upon the results. In the case of EEG and sleep, this can be developed based on the following considerations:

<table>
<thead>
<tr>
<th>Table 1. Comparison of available functional CNS biomarker techniques</th>
</tr>
</thead>
<tbody>
<tr>
<td>****</td>
</tr>
<tr>
<td>Direct measure of molecular target engagement?</td>
</tr>
<tr>
<td>Measure of pharmacological action and engagement of target neural circuit?</td>
</tr>
<tr>
<td>Direct measure of neuronal function?</td>
</tr>
<tr>
<td>Temporal resolution</td>
</tr>
<tr>
<td>Spatial resolution</td>
</tr>
<tr>
<td>Can it be integrated with SD/MD studies?</td>
</tr>
<tr>
<td>Portability/equipment can be standardised between sites</td>
</tr>
<tr>
<td>Availability</td>
</tr>
<tr>
<td>Cost</td>
</tr>
</tbody>
</table>

MEG = Magneto-encephalography; qEEG = quantitative EEG; n.a. = not assessed; SD = single dose; MD = multiple dose.
upon knowledge of translation. Continuous rodent EEG studies are straightforward, yield both EEG and sleep data, and are usually performed as a matter of course as part of the preclinical development of putatively CNS-active compounds. Examination of these results should determine whether or not a clinical EEG or PSG biomarker study in healthy volunteers is appropriate. In essence, if a robust preclinical signal has been measured and if it is expected that this effect will translate to humans (based on knowledge of the mechanism, the nature of the effect, etc.), then it is appropriate to include a clinical biomarker study to confirm that similar pharmacology is observed in man before proceeding with patient studies.

Having determined the expected effect, the clinical study should be designed with the primary objective of testing for this effect in humans, with additional standard EEG and/or PSG measures included as secondary endpoints to provide additional information to aid interpretation later. In this case, the basic decision rules for the clinical study can be relatively simple:

- If the expected effect (primary end point) is observed: translation of pharmacology is confirmed and the project continues
- If the expected effect (primary end point) is not observed but a plausible alternative pharmacological effect is seen: translation of pharmacology is not confirmed but central pharmacological action is likely and the project pauses for re-evaluation of the data
- If no pharmacologically plausible effects are observed: central pharmacological action is unlikely and therefore the project is put on hold or terminated

The most challenging interpretation arises in the second case, where the expected effect is not observed but another plausible pharmacological effect is seen, as this can occur either due to a failure of translation of the effect between species or a difference in pharmacology. Such a situation becomes less likely to arise with more experience as knowledge of translation improves. However, when it does occur, the best approach is to re-evaluate all the data to try to assess the level of confidence that the drug is indeed modulating the correct target in humans and hence what the next steps for the project should be (i.e. whether a further confirmatory study, perhaps with a different biomarker modality, is needed, or whether the project should continue with acknowledgement of the residual risk). A good place to start may be to re-evaluate the rodent data to assess whether the effect seen clinically was indeed present, although at a lower level than the primary dominant effect on which the hypothesis was based.

When interpreting the results, it must be remembered that EEG and PSG endpoints are potentially sensitive to non-central drug effects (e.g. due to systemic changes in blood pressure, peripheral discomfort disturbing sleep, etc.). However, this is also the case for other biomarker techniques in the 'toolbox' for assessing CNS activity. For example, fMRI is sensitive to any changes in the haemodynamic response function, which can be invoked by changes in blood CO₂ levels resulting from respiratory or cardiac changes. In the first case described in the decision rules above, where central activity has been confirmed in animals through prior detailed experiments and the expected translational EEG/PSG changes are observed, the risk that these changes are caused by a peripheral mechanism is low. However, in the second case, it is important to examine all available data to reduce the risk. Ideally, target engagement data such as RO-PET would also be available, which will significantly assist in this process.

Finally, in addition to the process described above, which lends itself to binary decisions, the study design should, where possible, also allow for quantitative PK-PD analysis, using the biomarker end point as the PD measure. This is most likely to be feasible with EEG because the influence of diurnal variations and other factors on sleep processing and the derived nature of the end points mean that PSG measures are generally less suitable for use in PK-PD modelling. Comparing PK-PD profiles between species will provide useful information about the relative potency of the compound between species and can be helpful in selecting a suitable dose to take forward into the first patient study.

Several case studies assessing the potential utility of this framework using EEG based on evidence available in the literature are described in detail in Wilson et al. [19]. In addition, an example based on the use of sleep end points in a real-world scenario has recently been reported [20]. In this case, EEG studies in rodents of the drug under development, an inhibitor of fatty acid amide hydrolase, showed both a sleep effect in the form of a modest suppression of rapid eye movement (REM) or stage R sleep and an EEG effect in the form of a small increase in beta band spectral power. In this case, the sleep end point was chosen as the primary end point for the clinical study based on statistical considerations – the beta band power change was too small to be robustly detected in a study of a reasonable and cost-effective size. A broadly similar suppression of REM sleep was observed in humans, thus confirming central activity, which was extremely important in decision-making following failure of a study in knee osteo-arthritis patients to demonstrate analgesic efficacy [21].
Obstacles

This framework is clearly only viable with a clear and robust knowledge of translatability, and this is often the greatest obstacle to its implementation. Whilst sleep end points are found to be highly translatable [22] and can be used for robust decision-making, EEG-based end points are more problematic. A large part of the reason for this is the lack of standardisation of the methods for recording and analysis of EEG studies, both preclinically and clinically, which makes meta-analyses and data-pooling problematic, hence frustrating attempts to assess translatability. As a result, there are still many questions about the optimum EEG methodology for translation, including the optimum electrode positions in different species, the equivalence of spectral bands for quantitative EEG analysis, the equivalence of preclinical and clinical paradigms for controlling vigilance and eye state and differences between human scalp EEG recordings and preclinical cortical recordings (including differences between gyrencephalic and lissencephalic brains). Another issue is the statistical handling of multiplicity due to the large number of potential variables [23].

These issues are not, however, technically insurmountable and can all potentially be resolved with a co-ordinated and well-structured effort.

Opportunities

Contrary to the obstacles described above, technological and computational advances over the last few years give cause for significant optimism with regard to the future of pharmaco-sleep and pharmaco-EEG as biomarker techniques. In particular, improved and smaller recording equipment enables easier use in clinical studies and much higher data quality (higher sampling and bit rates, improved artefact and noise isolation). Increased data storage capabilities enable all raw data to be stored easily for analysis/reanalysis offline and, combined with improved computational power, this enables faster and more automated analysis. Novel measures such as coherence and cordance can be studied, EEG source localisation algorithms such as LORETA [24] are available, and advanced techniques are being developed for the extraction of true neuronal gamma band signals from scalp recordings [25]. Finally, exciting enhanced analysis techniques are emerging, such as generalised semi-linear canonical correlation analysis (GSLCCA), which is a hypothesis-free technique that produces both a PD output and interpretable mechanistic information from EEG recordings without the need to define the frequency bands and electrodes of interest a priori and whilst avoiding many of the multiplicity concerns [26, 27].

Development Needs for Pharmaco-Sleep and Pharmaco-EEG

The first and most important step in unleashing the potential of pharmaco-EEG in drug development is a significant improvement in standardisation across industry and academia. To this end, the International Pharmaco-EEG Society has recently produced updated guidelines covering pharmaco-EEG in man [28]. Standardisation of pharmaco-sleep studies in man has generally been less of a problem due to the widespread use of the Rechtschaffen and Kales scoring rules [29] and the recent adoption of the American Academy of Sleep Medicine guidelines [18]. However, as these are aimed at clinical practice and hence do not cover all aspects of pharmaco-sleep, the International Pharmaco-EEG Society has also recently produced guidelines on pharmaco-sleep studies in man [30] to augment them.

In order for the empirical framework described above to be useful in as wide a range of situations as possible, work is needed to better understand the translation of pharmaco-sleep and pharmaco-EEG, to improve hypothesis setting and interpretation of results. The first step to this is to generate a wide range of standardised pharmaco-sleep and pharmaco-EEG data sets covering a wide range of different drug classes, produced in accordance with the new guidelines, to enable interspecies comparison. The most efficient way to achieve this is likely to be through precompetitive sharing of reference data across the pharmaceutical industry, possibly starting with a precompetitive consortium project to collect a large body of such data, particularly for clinical recordings which are much more expensive to obtain.

Finally, academic work is needed to develop novel and advanced techniques for signal processing and analysis, such as those needed for the gamma band and GSLCCA mentioned above. A significant part of the support for such work could come from the pharmaceutical industry making appropriate data more widely available to underpin the algorithm development work and enable meaningful comparative testing.
Conclusion

When applied in an appropriate way, and with some well-structured development work to further enhance their utility, pharmaco-sleep and pharmaco-EEG have significant potential to be amongst the key tools that enable early decision-making in the development of CNS drugs, hence helping to reduce the serious attrition problem faced by the pharmaceutical industry.

To achieve this goal, precompetitive work within the industry and co-operation with academia are needed to address the issues of standardisation and to improve knowledge of translatability, as well as to improve signal processing and data analysis techniques.

References

5. Leiser SC, Dunlop J, Bowly MR, Devilbiss Street SD: Can the flow of medicines be improved? Fundamental pharmacokinetic and pharmacodynamic modeling of the electroencephalogram help improve survival of central nervous system drugs, hence helping to reduce the serious attrition problem faced by the pharmaceutical industry.

To achieve this goal, precompetitive work within the industry and co-operation with academia are needed to address the issues of standardisation and to improve knowledge of translatability, as well as to improve signal processing and data analysis techniques.