CORE

Metadata, citation and similar papers at core.ac.uk

Provided by University of Birmingham Research Portal

UNIVERSITY OF BIRMINGHAM

University of Birmingham Research at Birmingham

Reducing attrition in drug development: smart loading preclinical safety assessment

Roberts, Ruth; Kavanagh, Stefan; R. Mellor, Howard; E. Pollard, Christopher ; Robinson, Sally; Platz, Stefan J.

DOI:
10.1016/j.drudis.2013.11.014

License:
Creative Commons: Attribution-NonCommercial-NoDerivs (CC BY-NC-ND)

Document Version
Peer reviewed version

Citation for published version (Harvard):
Roberts, R, Kavanagh, S, R. Mellor, H, E. Pollard, C, Robinson, S & Platz, SJ 2014, 'Reducing attrition in drug development: smart loading preclinical safety assessment', *Drug Discovery Today*, vol. 19, no. 3, pp. 341-347. https://doi.org/10.1016/j.drudis.2013.11.014

Link to publication on Research at Birmingham portal

General rights
Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of ‘fair dealing’ under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy
While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Download date: 01. Mar. 2020

Reducing Attrition in Drug Development: smart loading pre-clinical safety assessment

Ruth A Roberts, Stefan L Kavanagh, Howard R Mellor*, Christopher E Pollard, Sally Robinson and Stefan J Platz

Drug Safety and Metabolism, AstraZeneca, Alderley Park,
Macclesfield, SK10 4TG, UK
*Current address – Vertex Pharmaceuticals (Europe), Milton Park, Abingdon, OX14 4RW, UK

Corresponding author: Roberts, R. A.
(ruth.roberts@astrazeneca.com)
Phone: +44 1625 512849
Fax: +44 1625 513779

Key words: Attrition, Drug Safety, Toxicity, Drug Development, Drug Discovery, Pre-clinical Safety Strategies, Telemetry

Abstract

Entry into the critical pre-clinical ‘Good Laboratory Practice (GLP)’ stage of toxicology testing triggers significant R&D investment yet >20% of AstraZeneca’s potential new medicines have stopped for safety reasons in this GLP phase alone. How could we avoid at least some of these costly failures? An analysis of historical ‘stopping toxicities’ showed that > 50% were attributable to target organ toxicities emerging within two weeks of repeat dosing or to acute cardiovascular risks. By frontloading 2 week repeat dose toxicity studies and a comprehensive assessment of cardiovascular safety, we anticipate a potential 50% reduction in attrition in the GLP phase. This will reduce animal use overall, save significant R&D costs and improve drug pipeline quality.

Teaser: More than 20% of AstraZeneca’s potential new medicines were halted for safety reasons in the GLP testing phase alone – it is envisaged that ***a new approach could avoid at least some of these failures***

The overall decline in R&D productivity in pharmaceuticals over the past two decades is well documented [1, 2, 3, 4] In addition to the clear cost to patients in need of new therapies, the financial and reputational impact of late stage failures to the industry has also been significant. There are numerous factors underlying the productivity decline including a widespread industry focus on pipeline quantity rather than quality [5], heightened regulatory scrutiny[6] and an increasing focus of pharmaceutical company investment in areas of unmet medical need and unexploited biological mechanisms where there is a high risk of failure [2]. Yet, for the pharmaceutical industry to be sustained in its present form, a dramatic increase in innovation and R&D productivity is required to replace the continued loss of revenues from patent expirations on successful products [3].

Although a lack of clinical efficacy is the major cause of drug attrition, an unacceptable safety profile is also a major cause of development candidate failure [1, 7, 8], not only in the clinic, but at stages from drug candidate nomination through clinical development to post-marketing surveillance. Building on this, some failures attributed to poor efficacy may actually arise because the toxicity profile prevents a sufficiently robust exploration of the efficacy dose response curve (dose-limiting toxicity). In facing this significant safety challenge, the pharmaceutical industry must improve its ability to predict the probability of subsequent failure due to toxicity earlier in the drug discovery process and become increasingly selective regarding the targets and compounds nominated for progression.

Safety attrition is relatively common in the discovery (design) phase of drug development that takes place before candidate selection, where the identification of decision-driving chemical- or target-related toxicity is the primary aim of discovery toxicology strategies. At this early stage, there is an opportunity to ‘successfully’ halt programs prior to major investment or to influence chemical design away from the toxicity issue. However, these approaches, which often include *in silico* and *in vitro* methodologies, have not yet been consistently successful in predicting toxicities, particularly chemical- rather than target-driven ones, emerging upon repeat-dose *in vivo* exposure in the later regulatory Good Laboratory Practice (GLP) studies.

Many potential new medicines are halted in the GLP Phase, before entry into humans.

Before a new medicine can enter human clinical trials for the first time, safety and tolerability must be assessed in pre-clinical rodent and non-rodent studies both as a regulatory requirement and, more importantly, to assess, limit and manage risk to human volunteers or patients. The scientific rationale surrounding the use of animals in safety assessment studies to support human dosing has been reviewed elsewhere [9, 10]. Pre-clinical studies conducted prior to the initial phase 1 clinical investigations are required to characterise effects on function, target organ toxicity, to determine dose dependency and the relationship to drug exposure. Information from these studies is used to estimate initial safe starting doses and dose ranges, as well as to provide clinicians with key information to help develop suitable monitoring and patient exclusion strategies for these trials.

The preclinical safety packages performed by pharmaceutical companies before commencing human Phase 1 clinical trials for small molecules are specified in regulatory guidance [11, 12] and, as such, the designs often have a high degree of similarity. In general, toxicology studies of up to one month duration are suitable to support single or multiple dosing for a similar duration in Phase I clinical trials in human volunteers or patients [11]. Consequently, it is common practice within pharmaceutical companies to support initial Phase I studies in humans with toxicology studies of up to one month of duration in two species, conducted in accordance with GLP. For most companies the decision to enter into this ‘GLP study phase’ is a critical milestone since it carries with it significant financial and resource investment across R&D. However, a recent internal data analysis has shown that between 2001 and 2010 >20% of AstraZeneca’s potential new medicines were halted for safety reasons in the phase between the commencement of GLP toxicology studies and entry into humans. How could these toxicities be predicted, allowing us to avoid at least some of the failures in the future?

Root causes of the attrition

An analysis of 48 candidate drugs (CDs) that were halted for adverse safety findings in the first GLP dose to FTIM window between 2001 and 2010 was conducted (in-house data – unpublished). These 48 compounds represented 24% of the 198 AstraZeneca drug candidates nominated for development during this period. The primary cause of failure or ‘stopping toxicity’ for each candidate was identified and the duration of dosing after which the toxicity was reported to emerge was then determined.

Interestingly, at least 40% (Table 1) of the historical failures were attributable to target organ toxicities (detected by histopathology), or to unexplained deaths, that emerged within two weeks of repeat dosing. For 8% of the failed candidates, it was clear that the toxicity was not or would not be detected by the end of two weeks dosing. For a further 11%, it was unclear if two weeks dosing would be sufficient for toxicity detection. This was most frequently because the toxicity had not been detected in a 1 week dose range finding(DRF) study but had been detected after a later 1 month study, yet no interim timepoint data were available. In addition there were a number of candidate drugs (10%) that were halted due to acute cardiovascular risks that emerged in GLP dog telemetry studies, designed to detect alterations in cardiovascular parameters such as heart rate and blood pressure. For the remaining 19% of projects, the decision to stop development in the GLP phase was based on other issues such as genetic toxicology, reactive metabolites or due to clinical safety signals in more advanced but closely related projects (Figure 1). Regarding the candidate drugs halted due to genetic toxicology and cardiovascular risk during the GLP phase, current best practice in *in silico* and *in vitro* screening both for cardiovascular liability and for genetic toxicology is routinely employed at AstraZeneca and many potential candidate drugs carrying these liabilities are screened out along the route [13, 14, 15]. These assays typically increase in their sensitivity during the progression in time, complexity and cost from *in silico* to *in vitro,* to small scale non-GLP and finally to larger scale GLP screening. On top of this increase in sensitivity, the hazard may have been detected in vitro but then falling margins once data is available in vivo could make progression less viable. Thus, failures in the GLP phase are reduced but not removed by the earlier screening cascades. Regarding the repeat dose toxicities that were detected at less than 2 weeks of dosing, these were distributed among a variety of target organs such as liver, lung, muscle, pancreas, cardiovascular system and thyroid gland (Table1).

For most programs the only repeat dose *in vivo* data available for decision making at the time of compound nomination into development was from the maximum tolerated dose/dose range finding (MTD/DRF) studies conducted to set doses for the subsequent definitive GLP studies using only small numbers of animals (Figure 2a). Previously in AstraZeneca these non-GLP MTD/DRF studies traditionally followed a standard design where single rising doses are given to establish an MTD; once this is established a dose is selected for a short repeat dose phase routinely with 7 days of repeat dosing in the dose-range finding phase (Figure 2a). For scientific and welfare reasons, common practice in AstraZeneca is to explore adverse effects in rodent species prior to non-rodent species. This increases the amount of information available for the design of the non-rodent studies; for example, data from the initial rodent study can be used to help set the starting dose, or to allow for specific monitoring of adverse effects in non-rodents. The non-rodent maximum tolerated dose/dose-range finding study design was based on an industry consensus paper [16]. The primary objective of the study was to establish the limit of tolerance to set the high dose for regulatory studies using one group of animals. This design was used in AstraZeneca and was an optimal design alongside the pipeline quantity approach using the minimum number of non-rodents to rapidly move to the GLP studies. It is therefore not surprising that the ‘stopping toxicities’ observed at 2 weeks were previously undetected at the decision to enter the GLP study phase.

Regarding the acute cardiovascular risks, these would not normally emerge until the non-rodent telemetry study usually conducted in the GLP phase to support the risk assessment before entering into humans [17]. An exception to this would be if early, small animal cardiovascular studies have been performed perhaps driven by knowledge of a hightened, specific risk in this area.

What is the significance for human risk assessment?

A detailed analysis conducted by Olson *et al* examining the concordance of toxicities between animals and humans for 150 drugs demonstrated that while many adverse effects in humans were predicted by single-dose data, a significant proportion of human toxicities were only detected upon repeat dosing. The number of toxicities detected increased in line with dosing period, between one day and one month, with very few additional toxicities identified in longer experiments [18]

Further to this Rozman and Doull emphasised the quantitative relationship between dose, time and toxicological outcome, particularly at the low end of the dose-response curve [19]. Increasing the dosing duration can reveal toxicity at lower exposure levels relative to short-term studies, so it is important to explore the full toxicological dose-response curve to allow better contextualisation of any toxicity findings with respect to the predicted therapeutic exposure.

What have we put in place?

Based on the analysis, we have implemented a new pre-GLP approach with two key components (Figure 2b): 1) an extended repeat dose study (with a minimum of 14 days repeat dosing) to allow a more rigorous analysis of the probability of subsequent ‘stopping’ toxicities and their dose dependency and 2) a comprehensive assessment of cardiovascular safety in non-rodent models (previously conducted in the GLP phase). Regarding the extended repeat dose study, we have introduced 3 dose groups plus control in both the rodent and non-rodent study so that a dose response is established to give an early indication of margins to intended clinical exposure. A full clinical pathology and pathology examination (see Table 2a and b) will be conducted on these studies to maximize the information available for decision making. Echo cardiography and ophthalmoscopy are not routinely carried out on these studies unless a specific risk to be mitigated has been identified. Regarding animal numbers, the rodent study will use 3 animals per sex per group.

The non-rodent study will be based on just one animal per sex per group rising to 2 if scientific justification indicates the need. Clearly, these studies are designed to detect obvious safety signals early while minimizing animal usage and the numbers/group used have proven adequate for the purpose and are in line with industry practice [20]. Another key aspect of this approach is to ensure that the repeat dose study will also meet the requirements of setting doses for the GLP toxicology studies that should follow on rapidly for successful projects.

In addition to this extended in vivo rodent and non-rodent MTD/DRF study, we have brought forward the non-rodent telemetry to give a comprehensive assessment of cardiovascular safety before commitment to GLP toxicology and all the investment that this step triggers.

What do we hope to achieve?

Overall, the new strategy should detect compounds that induce target organ toxicity within two weeks and also identify compounds with acute cardiovascular risks. This retrospective analysis suggests that the strategy should bring drug candidate safety attrition forward, delivering the opportunity for up to a 50% reduction in attrition in the GLP phase (Figure 1). This makes some assumptions, primarily that our future candidate drugs will have a similar breadth of associated toxicological issues to those observed historically; this is not an unreasonable assumption since future drug candidates are likely to be derived from similar chemical space. When modeled hypothetically with 100 drug projects entering discovery, a 50% reduction in attrition in the GLP Phase (from 24% to 12%) (table 3) results in an improvement from 4 to 5 successful drug registrations, all other parameters on success at each stage of the pipeline being equal (Figure 3).

Undoubtedly, the significant effort that has been invested in early *in silico* and *in vitro* methodologies has yielded results in enhancing compound selection to reduce attrition but has not yet been consistently successful in predicting toxicities, particularly chemical- rather than target-driven ones, emerging upon repeat-dose *in vivo* exposure in the later regulatory Good Laboratory Practice (GLP) studies. The new approach means a greater investment earlier in the development programme since more animals are now used in the MTD/DRF studies and the non-rodent telemetry study is conducted earlier. This in turn requires more compound to be synthesized at an earlier stage to enable more robust decision making. Additionally, the new approach will take slightly longer during this MTD/DRF phase but this can be offset in efficiencies of time during the GLP phase and should have minimal impact on time taken to get new candidate drugs to the clinic. However, if successful, this new approach will save significant R&D costs, increase efficiency by avoiding progressing unsuitable compounds thus reducing animal use overall for those individual programmes and deliver an overall improvement in the quality of the drug discovery and development pipeline.

References

1. Kola, I. and Landis, J. (2004) Can the pharmaceutical industry reduce attrition rates? Nature Reviews Drug Discovery. 3, 711-715.
2. Pammolli, F. et al. (2011) The productivity crisis in pharmaceutical R&D. 2011. Nature Reviews Drug Discovery. 10, 428-438.
3. Paul, S.M. *et al*. (2010) How to improve R&D productivity: the pharmaceutical industry's grand challenge. Nature Reviews Drug Discovery. Vol 9, 203-214.
4. Herper, M. The Truly Staggering Cost of Inventing New Drugs. Forbes. March 2012.
5. Linder, M. (2007) Clinical attrition due to biased preclinical assessments of potential efficacy. Pharmacology and Therapeutics, 115, 148-175.

6. Munos, B. (2009) Lessons from 60 years of pharmaceutical innovation. Nature Reviews Drug Discovery. 8, 959-968.

7. Arrowsmith, J. and Miller, P. (2013) Trial Watch: Phase II and Phase III attrition rates 2011-2012. **P.** Nature Reviews Drug Discovery, 569.

8. Redfern, W.S, et al (2010) Impact and frequency of different toxicities throughout the pharmaceutical life cycle. The Toxicologist. 2010;114:1081.
9. Greaves, P et al. (2004) First dose of potential new medicines to humans: how animals help. Nature Reviews Drug Discovery. Vol. 3, 226-236.
10. ABPI (2011) Why are animals needed in medicines research? http://www.abpi.org.uk/industry-info/newmedicines/Documents/CaseStudy_WhyAnimalsNeeded.pdf

11. ICH Guideline M3(R2), (2009). Guidance on Nonclinical Safety Studies for the Conduct of Human Clinical Trials and Marketing Authorization for Pharmaceuticals. http://www.ich.org/products/guidelines/multidisciplinary/article/multidisciplinary-guidelines.html 12. ICH Guideline S9, (2009). Nonclinical Evaluation for Anticancer Pharmaceuticals. http://www.ich.org/products/guidelines/safety/article/safety-guidelines.html

13. Harmer AR, Abi-Gerges N, Morton MJ, Pullen GF, Valentin JP, Pollard CE. Toxicol Appl Pharmacol. 2012 Apr 15;260(2):162-72. doi: 10.1016/j.taap.2012.02.007. Epub 2012 Feb 21. Validation of an in vitro contractility assay using canine ventricular myocytes.

14. Harmer AR, Abi-Gerges N, Easter A, Woods A, Lawrence CL, Small BG, Valentin JP, Pollard CE. J Pharmacol Toxicol Methods. 2008 Jan-Feb;57(1):30-41. Epub 2007 Sep 26. Optimisation and validation of a medium-throughput electrophysiology-based hNav1.5 assay using IonWorks.

15. Pollard CE, Abi Gerges N, Bridgland-Taylor MH, Easter A, Hammond TG, Valentin JP. Br J Pharmacol. 2010 Jan;159(1):12-21. doi: 10.1111/j.1476-5381.2009.00207.x. An introduction to QT interval prolongation and non-clinical approaches to assessing and reducing risk.

16. Smith, D et al. (2005) Optimising the design of preliminary toxicity studies for pharmaceutical safety testing in the dog. Reg. Toxicol. Pharmacol. 41, 95-101

17. Leishman, J. et al (2012) Best practice in the conduct of key nonclinical cardiovascular assessments in drug development: current recommendations from the Safety Pharmacology Society. Pharmacol Toxicol Methods. 65, 93-101.

18. Olson, H. et al. (2000) Concordance of the toxicity of pharmaceuticals in humans and animals. Regul. Toxicol. Pharmacol. 32, 56-67.

19. Rozman, K. K. and Doull, J. (2001) The role of time as a quantifiable variable of toxicity and the experiemental conditions when Haber's c x t product can be observed: implications for therapeutics. J. Pharamcol. Exp. Ther. 296, 663-668.

20 NC3Rs/LASA (2009) Guidance on dose level selection for regulatory general toxicology studies for pharmaceuticals. http://www.nc3rs.org.uk/document.asp?id=1317

Figure Legends

Figure 1. 50% (green) of the compounds that failed in the phase between entry into GLP toxicology testing and entry into humans would have evoked 'stopping toxicity' in either extended 14-day DRF studies or the non-rodent telemetry studies normally conducted later in the development of a drug. For an additional 23%(yellow) of projects it is unclear if these measures would allow detection after 2 wks repeat dosing as the toxicity was originally only assessed after 1 month. The new strategy would certainly not have detected the toxicities for 8% (red) of projects as the toxicity was confirmed not to be present after 2 weeks repeat dosing. The remaining 19% (grey) were stopped for reasons other than non-clinical *in vivo* toxicity such as clinical toxicity of CD1 informing a stop for CD2 in the pre-clinical phase.

Figure 2. Panel A depicts the model used for most drug programs, where dose range finding studies in rodents and non-rodents are conducted prior to the candidate selection investment decision (CSID) to set doses for the subsequent definitive GLP studies using only small numbers of animals. Some programs may have early, small animal cardiovascular studies as an assessment of acute cardiovascular risks. In the new model (**Panel B)**, multi-purpose MTD/14 day repeat dose toxicity studies will be conducted in rodents and non-rodents at 3 doses plus control in support of the CSID decision. The rodent study will use 3 per group both males and females whereas the non-rodent will be based on just one animal per group rising to 2 if compound amounts allow. Non-rodent telemetry will also be moved prior to CSID to provide an assessment of acute cardiovascular risk.

Figure 3. Modeling the impact of a 50% reduction in attrition in the GLP Phase on subsequent drug project success. Darker bars show 100 drug projects entering the lead generation phase of discovery transitioning through 7 key stages each with its own probability of success (POS). Lighter bars show the impact on portfolio size of a reduction in attrition in the GLP phase from 24% to 12%, all other parameters being equal. Note logarithmic axis starts at 1.