The *Discussion Forum* provides a medium for airing your views on any issues related to the pharmaceutical industry and obtaining feedback and discussion on these views from others in the field. You can discuss issues that get you hot under the collar, practical problems at the bench, recently published literature, or just something bizarre or humorous that you wish to share. Publication of letters in this section is subject to editorial discretion and company-promotional letters will be rejected immediately. Furthermore, the views provided are those of the authors and are not intended to represent the views of the companies they work for. Moreover, these views do not reflect those of Elsevier, *Drug Discovery Today* or its editorial team. Please submit all letters to Steve Carney, Features Editor, *Drug Discovery Today*, e-mail: S.Carney@elsevier.com

Quantitative image analysis: software systems in drug development trials

In a recent issue of *Drug Discovery Today* [1], Pathak *et al.* published an instructive and timely article on the use of quantitative image analysis software in drug development trials. This article is a rewarding read because it comprehensively describes important issues concerning the design of (and demands on) professional software systems for large-scale clinical trials. These demands are clearly different from those in small-scale scientific studies, and go beyond those that relate to image analysis and image processing alone. The safety, integrity and traceability of electronic data, together with intuitive and efficient graphical user systems and secure database structures, are of great relevance here. This is particularly important because the digital nature of the images facilitates the processing of data from large multi-center studies in single centers that specialize in image analysis, and which might receive hundreds of datasets over relatively short periods. This letter aims to add a few general points on this topic, as well as some specific ones from a perspective of cartilage and osteoarthritis (OA) research.

Medical images have provided tremendous new insights into *in vivo* anatomy and pathology, and could provide powerful quantitative endpoints in conjunction with digital post-processing. It is important (and often difficult), however, to establish the relationship between 'imaging endpoints' and 'clinical endpoints' – features that are directly related to patient prognosis. For example, in osteoporosis, a drug must not only be shown to improve bone mass, density and structure, but it must also reduce the rate of fracture to gain regulatory approval. Furthermore, imaging endpoints must lie directly on the disease pathway and must register the relevant beneficial and adverse effects of the drug [2].

Pathak and colleagues correctly state that semi-automated or fully automated post-processing techniques have the potential to reduce the variability of the data and to increase measurement precision. However, as mentioned in their introduction, there are some potential drawbacks with uncritical automated processing. In cases where images display high noise levels, complex features and relatively low contrast (such as in magnetic resonance images of articular cartilage in OA), automated processing can be enforced by adding model-based approaches that 'help out' where image-based information is scarce. However, these algorithms might 'overlook' the subtle changes that one aims to measure. Defining the right amount of user interaction is thus a delicate process, and one must be aware that automated techniques are not always superior to careful manual or semi-automated approaches conducted by a carefully trained and experienced user.

When testing the intra- and inter-reader variability of the measurements (i.e. reproducibility and/or precision), and when comparing reports on precision among studies, one needs to account for the different levels of confidence and how the measures have been obtained statistically [3]. Furthermore, the analysis of measurement precision should not be confined to short-term conditions, but should also test for the effects of spreading image acquisition and or image processing over those time intervals that are typically encountered in longitudinal trials [3,4].

One crucial (and probably the most important) step in evaluating the performance of algorithms is 'validation'. This is achieved by testing to what degree the measurement (and the inter-subject variation of measurements) corresponds to the 'true value', as derived by the best available and established technique (gold standard). Tests on geometric test objects (phantoms) are certainly worthwhile, but these cannot replace direct validation on the biological object of interest. Where no other *in vivo* technique is available, validation has to involve cadaver studies. In the case of quantitative cartilage imaging, total knee arthroplasty, however, has provided a unique opportunity to directly validate quantitative *in vivo* measurements in relation to post-operative *in vitro* analysis [5,6]. Automated algorithms must withstand rigorous testing to ensure that savings in time and financial resources (and potential reductions in measurement variability) do not compromise the accuracy (technical validity) of the analysis.

 1359-6446/03/$ – see front matter PII: S1359-6446(03)02758-2

Quantitative imaging certainly holds tremendous potential for the evaluation of the 'structure-' and 'disease-' modifying capability of drugs versus purely symptomatic relief. However, like Odysseus, we should resist the sound of the sirens. Rather than being seduced by shiny images and/or technically appealing algorithms, we ought to maintain a critical view, and to neither trust images or algorithms, until proven otherwise.

References

1 Pathak, S.D. *et al* (2003) Quantitative image analysis: software systems in drug development trials. *Drug Discov. Today* 8, 451–464
2 Peterfy, C.G. (2000) Scratching the surface: articular cartilage disorders in the knee. *Magn. Reson. Imaging Clin. N. Am.* 8, 409–430
3 Glüer C.C. *et al* (1995) Accurate assessment of precision errors: how to measure the reproducibility of bone densitometry techniques. *Osteoporos. Int.* 5, 262–270
4 Eckstein F. *et al* (2002) Long-term and resegmentation precision of quantitative cartilage MR imaging (qMRI). *Osteoarthritis Cartilage* 10, 922–928
5 Burgkart R. *et al* (2001) Magnetic resonance imaging-based assessment of cartilage loss in severe osteoarthritis: accuracy, precision, and diagnostic value. *Arthritis Rheum.* 44, 2072–2077
6 Graichen, H. *et al*. Quantitative assessment of cartilage status in osteoarthritis with qMRI – technical validation in total knee arthroplasty beyond cartilage volume. *Arthritis Rheum*. (In Press)

Felix Eckstein
Musculoskeletal Research Group
Institute of Anatomy
Ludwig-Maximilians-Universität München
Pettenkoferstr. 11
D-80336 Munchen, Germany

K.H. Englmeier
Institute for Medical Informatics and Systems Research (MEDIS)
GSF-National Research Center for Environment and Health
Ingolstädter Landstrasse 1
D-85764 Neuherberg, Germany

Ultra high quality uHTS

The application of fluorescence detection to HTS and its recent successor, ultra-HTS (uHTS), is well established. This is clear from the wealth of techniques and commercial systems that are discussed in the excellent overview by Stefan Jäger and co-workers that was published recently in *Drug Discovery Today* [1].

In the early days of HTS – the mid-1990s – clients were often presented with a single commercial solution to meet a particular need. The reliability of the chosen instrument was questionable and years of development were needed before satisfactory results could be guaranteed. Together, suppliers and consumers in the field have progressed a long way; diverse and skilled teams of engineers and scientists have worked synergistically to bring much improved instruments to the market.

Eggeling *et al.* also describe the many established and novel fluorescence methods that have been adapted to the demanding conditions of uHTS [1]. Among these, microscopic fluorescence shows great promise in the delivery of not only ultra high throughput but also ultra high quality. Despite similar advances in liquid handling in the low microliter and nanoliter range [2], experimental error in reagent dispensing is the main cause of false positives in screening campaigns. Miniaturization enables replication of experiments to minimize the impact of errors without increasing spending on reagents. Furthermore, microscopic fluorescence provides an insight into the binding events and reagent concentrations in every well, enabling the correction of errors in liquid handling or sample processing (e.g. mixing or evaporation).

Although the new fluorescent techniques are promising, uHTS will never become the universal solution for drug discovery, as recently pointed out by Schmid [3]. However, following years of capital investment, automated compound testing can be seen as an efficient engine to facilitate drug discovery if, and only if, it is fueled with the right targets and compounds.

Discovery scientists are still trying to answer the fundamental question: where are the drugs that genomics and HTS promised to deliver? The average time frame for the development of a drug is estimated to be 15 years. HTS started to be applied in the early 1990s, only being seriously refined and implemented in the late 1990s (for example, Lipinski's landmark paper [4] is from 1997), and thus, it might be too early to expect the answer now. Development pipelines show that the number of projects emerging from this new paradigm is steadily increasing and will continue to do so. Attrition rates are alarming, but this is likely to be a consequence of using novel targets and chemotypes, rather than a consequence of employing a chemistry programme from an engineered, serendipitous event (diversity screening) instead of a structure-based idea (rational design).

Furthermore, to improve the otherwise slim chances of success, drug hunters should apply all available options, exploiting HTS and rational design in a parallel and complementary way – there is no reason to rule out either one or the other.

References

1 Eggeling, C. *et al*. (2003) Highly sensitive fluorescence detection technology currently available for high throughput screening. *Drug Discov. Today* 8, 632–641
2 Taylor, P. *et al*. (2002) A standard operating procedure for assessing liquid handler performance in high-throughput screening. *J. Biomol. Screen*ing 7, 554–569
3 Schmidt, E. (2003) Trust me – I'm a doctor. *Drug Discov. Today* 8, 439–438
4 Lipinski, C.A. *et al.* (1997) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* 23, 3-25

Ricardo Macarrón
Director of Molecular Screening-Philadelphia
GlaxoSmithKline
709 Swedeland Rd
King of Prussia, PA 19406, USA
E-mail: Ricardo_Macarron@gsk.com