The evolution of high complexity companion testing for targeted and immuno-oncology

“...we are beginning to see a convergence in the companion testing required for targeted and immuno-oncology.”

First draft submitted: 7 July 2016; **Accepted for publication:** 15 July 2016; **Published online:** 9 August 2016

Keywords: companion • immuno-oncology • oncology • precision • sequencing

Companion testing in precision oncology: where are we today?

As this article goes to press in late 2016, the precision medicine community has amassed over 20 years of companion product development experience in the field of oncology. This article reflects on what we have achieved and learned along the way, how the companion paradigm is playing out in immuno-oncology and why we are likely to achieve much more in the next 10 years than we have in the previous two decades. Also considered here are some of the disruptive operational and business model changes we will see in the coming years as these new opportunities are realized.

The current companion scenario has played out in targeted oncology since 1998, with the rate of innovation increasing in recent years. Indeed, over 45 new oncology drugs launched between 2010 and 2014 in more than 53 indications now extend life for many patients, while avoiding harm (toxicity) for others [1]. Nevertheless, despite the long-recognized promise of high precision medicine, highly specific targeting is still lacking, with only seven unique ‘required’ *in vitro* companion diagnostic oncology tests cleared to date by the US FDA (*HER2*, *CKIT*, *EGFR*, *KRAS*, *BRAF*, *ALK*, *BRCA1/2*). These markers have underpinned important therapeutic advances and enabled drug development in subpopulations while providing for the exclusion of patients lacking specific target variants/expression, but lack the posi-

tive predictive value to identify responders with high precision. This is a direct result of the fact that, while these markers are generally implicit in the underlying target biology, capturing either differential target expression or molecular variants, they represent only a single molecular parameter in a complex molecular ecosystem.

With the singular exception of *BRCA1/2* testing, these companion tests have been developed as distributed *in vitro* diagnostic (IVD) product (‘kit’) offerings which then compete commercially with various independently developed laboratory-developed test (LDT) offerings (also known as in-house tests or IHTs in the EU). Development has generally been under a collaborative agreement between a pharmaceutical and diagnostic product (IVD) company under a mandated FDA pre-market approval (PMA) path, the most stringent regulatory option. This model, while ensuring quality, has tended to limit companion business model innovation and created a scenario whereby static companion IVD products are launched into a marketplace dominated by dynamic LDTs.

Next-generation sequencing (NGS), as presently deployed, generally represents an important evolution, but not yet a step change, in this targeted therapy model. For example, a typical small NGS panel deployed in the well-characterized non-small-cell lung cancer (NSCLC) context brings important clinical workflow efficiencies via the highly parallel analysis of several ‘actionable’ genetic driv-

Iain D Miller
Healthcare Strategies Group, London,
UK
iain@healthcarestrategiesgroup.com

ers simultaneously (e.g., *EGFR*, *KRAS*, *ALK*, *BRAF*), thereby conserving scarce tissue and potentially speeding time to ultimate actionable result if several tests are required. Some of these targeted panels leverage circulating tumor DNA (ctDNA), which offers both a tissue-sparing opportunity and the potential for real-time monitoring of evolution of tumor status and clonality. Important as these contributions are, NGS as a whole has yet to change the single-marker therapy guidance scenario in the clinic. By contrast, larger panels, such as the 343-gene FoundationOne® panel developed by Foundation Medicine (Cambridge, MA, USA) [2], extend the concept of actionability into the research domain and have already begun to yield important insights. However, such larger panels are not yet part of routine care. This commentary will argue that the clinical contribution of such complex molecular analysis is set to change markedly in the coming years. Furthermore, as will be discussed later in this article, the FDA is also evolving its oversight model toward a more flexible approach which should enable more dynamic and innovative IVD and LDT companion test offerings from multiple providers going forward.

“Next-generation sequencing, as presently deployed, generally represents an important evolution, but not yet a step change, in this targeted therapy model.”

In addition to the predictive limitations of current companion test offerings in the clinic, many commentators have described the challenge inherent in implementing homogenous testing in a laboratory landscape characterized by multiple LDT test clones with variable levels of evidentiary underpinning [3]. In the case of the relatively ‘simple’ example of HER2 testing, real-world discordance rates as high as 26% persisted for several years after the 1998 companion product launch [4], although it should be noted that some of this discordance resulted from lack of compliance with labeling of FDA-regulated tests. Indeed, optimization of HER2 testing protocols has been the subject of continued debate until the publication of the revised American Society of Clinical Oncology/College of American Pathologists guidelines in 2013 [5], which have contributed to current discordance rates as low as 4% in the clinical setting today [6]. Notwithstanding the advances in HER2 testing, significant heterogeneity in companion test quality is still a feature of the precision medicine landscape today.

The evolving companion testing model in immuno-oncology: first-generation tests

Immuno-oncology, as with targeted therapy (as the term is currently used), has the potential to become a

highly individualized offering in which precise diagnostic guidance informs initial selection and ongoing optimization of combination therapy regimens. The two fields are increasingly seen as synergistic, as witnessed by the proliferation of combination clinical trials. In common with the history of targeted therapy guidance, the first generation of immuno-oncology testing is based on a single target-related biomarker, in this case PD-L1, the programmed death ligand which communicates with the PD-1 receptor.

The last 2-year period has witnessed multiple approvals within the PD-1/PD-L1 class for Opdivo® (BMS, NY, USA), Keytruda® (Merck, Kenilworth, NJ, USA) and Tecentriq® (Roche-Genentech, CA, USA) in indications including melanoma, NSCLC and urothelial cancer. Across these indications, companion biomarker impact has been rather mixed, with only Keytruda in NSCLC being approved with a companion (PD-L1) test, with Opdivo in NSCLC and Tecentriq in urothelial cancer being approved instead with newly designated ‘complementary’ diagnostics. The FDA felt the need to introduce this new diagnostic class (defined as ‘tests which enrich for response, but which are not deemed necessary for selection’ [7]) in order to highlight both the higher response rates in PD-L1 positives (vs all-comers) and the fact that a subpopulation of PD-L1 negatives still responded to the drug. Notably, the various PD-L1 assays used in these studies differ in their specification of positivity cut-off, and leverage different antibodies and staining protocols. As a consequence, use of the PD-1/PD-L1 therapy class in the clinic today is associated with a rather complex biomarker picture, albeit one focused on a single biological marker.

With the foregoing in mind, the FDA launched its immuno-oncology-focused Blueprint initiative in March of 2015 [8]. Blueprint was also motivated by the FDA’s awareness of the long validation journey of HER2, another immunohistochemistry (IHC)-based test, and of general challenges associated with interchangeable use of multiple tests of increasing complexity to select amongst therapies in the same drug class.

The objective of the Blueprint initiative is to achieve interoperability between available tests and drugs in the PD-1/PD-L1 class, initially by assessing analytical concordance in a small (n = 38) representative NSCLC cohort between four available assays used at their respective cut-off points. Recently reported Blueprint Phase I results were rather discouraging, however [9]. Specifically, of the four tests studied across a total of 156 ICH slides, only 50% of cases yielded concordant positive results across the four assays, a result driven in part by the different cut-off rate of each test. Investi-

gators concluded that, for now, “clinicians should still follow the recommendations of specific assay-drug pairs,” a rather unsatisfying outcome when measured against the project’s interoperability objective. An ongoing Phase II Blueprint trial with a larger cohort may add further clarification. At the same American Association for Cancer Research meeting, a team from Astra Zeneca (London, UK) reported rather more encouraging results in a larger cohort of 500 NSCLC patients using three of the available PD-L1 assays [10]. Specifically, this team found that over 90% concordance could be realized if interassay cut-offs were optimized (varied), suggesting the possibility of future algorithms to extrapolate results from one assay to the other.

The aforementioned studies, together with their ongoing successor studies, offer the possibility of ultimate harmonization of the PD-L1 assay landscape. This will be important for health technology assessment bodies such as NICE, which noted in its recent appraisal consultation documents on use of Opdivo in both the nonsquamous and squamous NSCLC settings that one reason for the preliminary noncoverage recommendations was that the sponsor (BMS) had not sufficiently powered its trials to demonstrate the PD-L1 positive subgroup effect [11]. Indeed, the FDA-defined category of ‘complementary’ tests may ultimately come to have a similar meaning to payers as companion tests, providing the necessary degree of enrichment to meet cost–effectiveness criteria and (potentially) pay-for-performance mandates, especially as these drugs migrate toward the first-line NSCLC setting.

Harmonization of PD-L1 testing, as noted, offers the potential for better targeted immunotherapy in some clinical oncology settings, together with more favorable cost–effectiveness for healthcare systems. However, the potential benefit of PD-L1-based stratification should not be overestimated and is best seen as being comparable to the current generation of approved FDA target-specific companion tests, with their significant attendant limitations. Indeed, PD-1/PD-L1 is but one checkpoint mechanism and checkpoint blockade itself is but one part of a complex multifactorial immunity cycle in which anticancer immunity evolves. The real step change for precision medicine in immuno-oncology will likely come from the definition and operational deployment of a multifactorial test which more fully describes the broader interactions between cancer and the immune system, the ultimate arbiter of long-term outcome. Such multifactorial testing – and the combination therapy options it will inform – has the potential to better serve the patients who do not currently respond and those who experience significant toxicity.

Next-generation immuno-oncology & targeted therapy companion testing

Successful deployment of predictive medicine in immuno-oncology will require that a broader assessment of the cancer immune system interaction be undertaken (in conjunction with the assessment of molecular actors necessary to inform use of any combination targeted therapy). Several conceptual frameworks are starting to provide insight on what this might mean for companion testing. One such framework is the ‘cancer immunity cycle’, recently elucidated by various authors [12]. Other groups [13] have attempted to capture how this diverse dataset might be deployed, by introducing the concept of an actionable and individualized ‘Cancer Immunogram’.

“Such multifactorial testing – and the combination therapy options it will inform – has the potential to better serve the patients...”

Taken together, these new frameworks, evidenced by a number of distinct separate studies by other groups of individual “Immunogram” subcomponents, suggest that consideration be given to a number of factors during a patient evaluation. These factors include tumor ‘foreignness’, general immune status, immune infiltration, checkpoint status, presence of soluble inhibitors, tumor sensitivity to immune effectors and other local tumor metabolic factors [13]. Collectively, these factors address immune recognition, activation, infiltration (‘inflamed’ phenotype) and impact, with T-cell activity seen as the ultimate effector mechanism. The visual cancer immunogram concept captures these parameters as part of a 7-dimensional visual ‘radar plot’ of immunological status with putative consequences for selection of specific therapy combinations. As noted by the authors, the information required for Immunogram assessment may be obtained from a combination of various diagnostic approaches including tumor genomics, IHC and standard assays on the peripheral blood compartment. Some of these diagnostic tools are already deployed in the targeted therapy context, suggesting that integrated deployment may not require a step change in current workflows.

The ‘Cancer Immunogram’, while conceptual in nature, is likely to be increasingly evidenced by studies of the impact of its individual tumor microenvironment variables on patient outcome. Indeed, several existing studies already provide such proof of concept, leveraging technologies with already established impact on selection of targeted therapy. Collectively, this body of work offers a tantalizing glimpse of what might become a universal companion diagnostic scenario, albeit not necessarily delivered by a single technological platform. For example, multigene NGS

exome panels can be used to both direct and monitor use of targeted therapy, while also providing important information on immunological metrics such as mutational load and mismatch repair defects, both of which are proxies for tumor foreignness [14,15], together with detail on T- and B-cell clonality [16], which dictates immunological status/diversity. Similarly, IHC, long a mainstay of targeted therapy selection, underlies the Immunoscore T-cell infiltration test which was recently shown to successfully prognose stage I/II/III colon cancer [17] and which has been under study for many years. Indeed, immune responsivity and associated T-cell infiltration may prove to be gating diagnostic parameters for early deployment of immunotherapy. As noted in my previous article on this topic [18], varying degrees of evidence exist to support use of a range of other technologies within the Immunogram concept. These technologies include assessment of functional T-cell status using multimer-based fluorescence-activated flow sorting (FACS) or Elispot techniques, analysis of molecular *IF- γ* signatures, and use of various multiplexed protein/molecular platforms. Associated treatment algorithms could, for example, first gauge immunological responsivity ('inflamed phenotype'), then address checkpoint status in inflamed phenotypes, together with other factors [20].

“From a developmental perspective, the landscape of pharma-diagnostic partnerships has already shifted.”

These examples illustrate that we are already implicitly building some of the pieces of the Cancer Immunogram, albeit in distinct development activities across the precision oncology community. A shift to such a high-complexity test paradigm raises a number of developmental and operational questions, which will be explored in the remainder of this article.

Development & deployment considerations for high-complexity tests

As noted, the precision oncology field is presently undergoing a transition from target and drug-specific single gene tests to high complexity offerings reflecting a more holistic view of relevant biological pathways. Current evidence suggests that this new generation of tests has the potential to bring a step change in the precision with which we are able to target patients today. Implementation of these new companion and baseline tests will require new development and validation models, a flexible regulatory approach and new provider models.

From a developmental perspective, the landscape of pharma-diagnostic partnerships has already shifted. Relative to the first generation of targeted therapy

partnerships, several deals being done today increasingly involve more complex panels, often have hybrid LDT-IVD structures and are generally of higher value. Illustrative examples are the partnerships which Astra Zeneca has struck with Myriad Genetics (UT, USA) and Foundation Medicine for its DNA damage response portfolio. Both partnerships leverage clinical laboratory improvement amendment (CLIA)-based sequencing of several homologous recombination repair genes, one of which (Myriad's BRACAnalysis CDx[®]) was the first LDT companion diagnostic approved under the PMA process, the most stringent regulatory option available to the FDA. Other examples of high complexity companion programs include partnerships negotiated between Nanostring (WA, USA) and Celgene (NJ, USA) and (separately) with Merck for several complex gene expression signatures, and between Foundation Medicine and Clovis Oncology (CO, USA) leveraging the 343-gene FoundationOne assay described earlier to interrogate DNA damage response genes. Pharmaceutical companies have traditionally been wary of centralized testing, fearing market access challenges, but are now recognizing the benefits of such complex testing and adopting strategies to regionalize testing.

In parallel, the FDA is keen to adopt a flexible and pragmatic approach to enable development of innovative complex tests. Indeed, the FDA seems likely to regulate some NGS panels as class II devices with general and/or special controls, as it did with Illumina's cystic fibrosis panel in 2013, leveraging well-curated databases such as NCBI's ClinVar to provide evidence on the clinical significance of variants [18]. Notably, a class III (PMA) submission was not required for the cystic fibrosis panel, providing a window into the pragmatic path the FDA is charting for both product and service (laboratory) developers of future test panels, which would not individually be required to submit duplicative dossiers on clinical validation. Indeed, the FDA has characterized its proposed approach as being more quality-systems oriented, focused on a combination of recognition of the interpretive relevance of dynamically evolving variants within recognized databases together with a design concept approach leveraging established metrics for diagnostic accuracy, precision, sensitivity and specificity against said variants. This evolving FDA view was recently captured in draft guidance documents directed toward database recognition and germline test development [19]. These documents, while not specifically addressing oncology panels or the ongoing LDT oversight controversy, are starting to provide important operational clarity for the emerging high-complexity oncology companion testing paradigm from both central laboratories and distributed IVD product suppliers and will be followed

by additional oncology NGS- and LDT-specific guidance in the future. Indeed, the recent (August, 2016) acceptance by FDA and the US Centers for Medicare and Medicaid Services (CMS) for review of Foundation Medicine's FoundationOne NGS test under the FDA's expedited access pathway (EAP) and the joint (FDA/CMS) parallel review process is a significant development. Specifically, from a regulatory perspective, not only will this shine a light on FDA's evolving regulatory approach for NGS-guided oncology precision medicine, but it may also herald the beginning of a post-PMA companion testing scenario which better informs multiple clinical settings and decision points. Furthermore, funding of US\$10 million in 2016 was allocated to the FDA under President Obama's US\$215 million Precision Medicine Initiative to acquire additional expertise and advance the development of high quality, curated databases to support the required regulatory structures. In parallel, the new EU IVD regulation, recently approved (though not yet published in full) by the EU Council, seems set to build on the analytical framework established in its 18-year-old IVD Directive predecessor, while engaging EMA for companion product reviews and clarifying in-house testing exemption limitations. Such pragmatic regulatory approaches on both sides of the Atlantic may increasingly blur the lines between IVD and LDT companion product offerings and should serve to level the playing field, while assuring quality.

Some of the aforementioned considerations suggest a market shift toward regionalized high-complexity companion test offerings. Indeed, the very term 'companion', which has long stood for a 1:1 test-drug relationship, may rapidly come to seem an outmoded term. As the evolving FDA framework and Blueprint initiatives anticipate, tests will increasingly have a 'one to many' relationship with the drugs they guide and may perhaps instead come to be seen as part of a future baseline tumor workup. Under this likely scenario, pharmaceutical companies may look to diagnostic providers as molecular information aggregators rather than providers of a single test. From a provider perspective, investment in sophisticated baseline testing is likely to yield a high return in terms of optimization of patient management and cost-effectiveness.

The increasing complexity of companion/baseline testing may, at least in the short term, result in a 'Centre of Excellence' (COE) type deployment model in some countries. For example, in the UK context, 13 Genomic Medicine Centers have been established to transform diagnosis and treatment for patients with cancer and rare diseases, while supporting the GB£300 million plus 100k Genomes project. France in turn has recently committed to investing €670 mil-

lion in the establishment of 12 regional NGS 'platforms' to sequence 175,000 genomes/year from 50,000 patients with relapsed/refractory disease by 2020 [19]. Within Germany, KBV (the National Association of Statutory Health Insurance Physicians) introduced on 1 July 2016, specific EBM (Uniform Value Scale) fee schedule coverage for multiparametric panels (including NGS panels) for tumor genetic studies. However, the type of centrally driven COE provision now emerging in the UK and France may not prevail in Germany, or indeed in the US, with their robust public-private laboratory mix. Instead, it is likely that a mix of highly specialist private laboratories and leading academic medical centers will continue to compete to develop the required competencies in these and other countries. It is in turn likely that such COE/expert laboratory providers will look to IVD and IT providers to provide key subcomponents and/or oversight of this high complexity testing infrastructure, including local deployment of validated (IVD) test subcomponents and information technology, while traditional sole-source (CLIA) providers, often based only in the US, may have to partner and regionalize their offerings. Such expert laboratories may then become the sole recipients of scarce patient samples, and will have to develop algorithms for maximum information yield from sequential or highly parallel testing. Indeed, some authors have begun to propose such hypothetical algorithms for combination immunological and targeted therapy [20].

“...pharmaceutical companies may look to diagnostic providers as molecular information aggregators...”

As the foregoing illustrates, we are beginning to see a convergence in the companion testing required for targeted and immuno-oncology. Indeed, this author believes that immuno-oncology itself, notwithstanding some of the 'all-comers' drug labels today, will increasingly be seen as a subset of targeted therapy. As noted, some conceptual frameworks for a universal baseline testing paradigm are starting to emerge, with the potential to deliver a step change in the precision with which patients are targeted today. In parallel, regulators are aligning around pragmatic assessment models with the potential to harmonize quality, reflect the dynamic nature of variant annotation and stimulate innovation across both the product and service (LDT/IHT) sectors. Single parameter 'companion' tests are likely to give way to high complexity baseline patient assessment and liquid-biopsy mediated continuous monitoring frameworks. All of this is very good news for the cancer patient and the payer, both of which will be well served by what promises to be an exciting period of creative disruption.

Financial & competing interests disclosure

Dr Miller is a volunteer member of the National Institute for Health and Care Excellence (NICE) Technology Appraisals Committee, which considers the cost-effectiveness of therapeutic and other medical technologies. Dr Miller does engage in consulting work in the personalized healthcare space and has consulted for Myriad Genetics. The author has no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial

conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Open access

This work is licensed under the Attribution-NonCommercial-NoDerivatives 4.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>

References

- 1 Developments in cancer treatments, market dynamics, patient access and value: Global Oncology Trend Report 2015. IMS Institute. www.imshealth.com/en
- 2 Foundation medicine. <http://foundationone.com>
- 3 Miller ID. Best practices and emerging trends for market access to personalised medicine in the US and EU: learnings for global developed and emerging markets. *Curr. Pharmacogenomics Person. Med.* 12(2), 104–113 (2014).
- 4 Roche PC, Suman VJ, Jenkins RB *et al.* Concordance between local and central laboratory *HER2* testing in the breast intergroup trial N9831. *J. Natl Cancer Inst.* 94(11), 855–857 (2002).
- 5 College of American Pathologists. www.cap.org/apps/docs/committees
- 6 Kaufman PA, Bloom KJ, Burris H *et al.* Assessing the discordance rate between local and central *HER2* testing in women with locally determined *HER2*-negative breast cancer. *Cancer* 120(17), 2657–2664 (2014).
- 7 Precision medicine at FDA. US Food and Drug Administration (2015). www.fda.gov/downloads/AdvisoryCommittees
- 8 Complexities in personalized medicine: harmonizing companion diagnostics across a class of targeted therapies. US Food and Drug Administration (2015). www.fda.gov/downloads/MedicalDevices
- 9 Hirsch FR, Philip R, Averbuch SD *et al.* The Blueprint Project: harmonizing companion diagnostics across a class of targeted therapies blueprint results. Presented at: *The American Association for Cancer Research Annual Meeting*, New Orleans, LA, USA, 16–20 April 2016.
- 10 Ratcliffe A. Comparison of three different PD-L1 NSCLC diagnostic tests shows a high degree of concordance. Presented at: *The American Association for Cancer Research Annual Meeting*, New Orleans, LA, USA, 16–20 April 2016.
- 11 National Institute for Health and Care Excellence. www.nice.org.uk
- 12 Chen DS, Mellman I. Oncology meets immunology: the cancer–immunity cycle. *Immunity* 39(1), 1–10 (2013).
- 13 Blank CU, Haanen JB, Ribas A *et al.* The Cancer Immunogram. *Science* 352(6286), 658–660 (2016).
- 14 Rosenberg J. PD-L1 expression, cancer genome atlas (TCGA) subtype and mutational load are independent predictors of response to atezolizumab (atezo) in metastatic urothelial carcinoma. Presented at: *The American Society of Clinical Oncology Annual Meeting*, Chicago, IL, USA, 3–7 June 2016.
- 15 Devarakonda SH, Masood A, Tanner MJ *et al.* Somatic mutations in mismatch repair pathway genes in non-small cell lung cancer. Presented at: *The American Society of Clinical Oncology Annual Meeting*, Chicago, IL, USA, 3–7 June 2016.
- 16 Funt S, Charen AS, Yusko E. Correlation of peripheral and intratumoral T-cell receptor (TCR) clonality with clinical outcomes in patients with metastatic urothelial cancer (mUC) treated with atezolizumab. Presented at: *The American Society of Clinical Oncology Annual Meeting*, Chicago, IL, USA, 3–7 June 2016.
- 17 Galon J, Mlecnik B, Marliot F. Validation of the immunoscore (IM) as a prognostic marker in stage I/II/III colon cancer: results of a worldwide consortium-based analysis of 1336 patients. Presented at: *The American Society of Clinical Oncology Annual Meeting*, Chicago, IL, USA, 3–7 June 2016.
- 18 Miller ID. The emergence of immuno-oncology companion products. *Per. Med.* 12(5), 435–438 (2015).
- 19 Use of public human genetic variant databases to support clinical validity for next generation sequencing 4 (NGS)-based *in vitro* diagnostics: draft guidance for stakeholders and Food and Drug Administration staff. US Food and Drug Administration (2016). www.fda.gov/ucm/groups/fdagov-public
- 20 Kim JM, Chen DS. Immune escape to PD-L1/PD-1 blockade: seven steps to success (or failure). *Ann. Oncol.* doi:10.1093/annonc/mdw217 (2016) (Epub ahead of print).