Point of View

Harnessing Complexity: Companion Diagnostics in Oncology

October 2015

Pieter De Richter,
Head of Global Molecular Diagnostics Portfolio,
Ipsos Healthcare
Introduction: beyond the microscope

Companion diagnostics have had a profound impact on the way in which cancer is treated. As an increasing number of drugs gain approval for patients whose cancers have certain molecular characteristics, the promise of truly personalised medicine is edging closer towards reality. The realisation that every patient’s cancer is different – and that these differences can be detected with in vitro diagnostic tests and exploited with tailored therapies – has led to a rapid fragmentation of the anti-cancer drug landscape. This, in turn, has led to a significantly more complicated decision-making process for oncologists in many cases.

Molecular diagnostics testing is at the far end of a long and evolving spectrum of narrowing down the nature of a disease. This spectrum starts with the clinical observation of (often visual) symptoms, moves on to more advanced visualization techniques involving x-rays and other imagining machinery, pulls in pathologists and their microscopes to examine the overall cell structure, involves dyes and stains to quantify the expression of certain markers on cell membranes, and finally moves deep within the cell nucleus to look for specific changes in the cancer’s DNA.

It is no longer sufficient to classify a patient’s cancer as being an adenocarcinoma of the lung; rather, it is now crucial to know if that particular patient’s adenocarcinoma of the lung has certain mutations in the gene that encodes the EGF receptor and, if so, whether those mutations are sensitising mutations as opposed to resistance mutations. This is not something that an oncologist can assess by going through the patient’s symptoms in detail. Neither is it something that a radiologist can see on an X-ray, nor can it be seen by the surgeon during a biopsy, nor even the pathologist upon examination of the biopsy material under a microscope. DNA changes occur at the molecular level, and can only be detected with sophisticated machinery that rely on such advanced techniques as real time polymerase chain reactions (PCR) and, increasingly, full gene sequencing.
Predicting response: perfection or educated guesswork?

The first question concerning companion diagnostics is perhaps the biggest one of all: do they increase a drug's usage due to its demonstrably superior effects in a specific patient population, or do they limit a drug's uptake due to the exclusion of (often large) groups of patients who lack the required biomarker status? In order to answer this complex question, we must first embark upon a scientific journey to better understand what exactly it is a companion diagnostics test is measuring, starting with immunohistochemistry testing for HER2.

Breast cancer cells, like most high-incidence tumours, are a heterogeneous lot. Apart from the histology, which distinguishes between ductal, lobular and other cell lines, we’ve known for some years now that there are a number of surface proteins that are expressed to varying degrees on the cell membrane of these cancer cells. Three cell surface receptors in particular stand out: the oestrogen receptor (ER), progesterone receptor (PgR) and human epidermal growth factor receptor 2 (HER2). Whereas the hormone receptors are overexpressed by the majority of breast cancer cells, the HER2 receptor is said to be overexpressed by around 25% of breast cancers. So what exactly does this mean?

The key point to realise is that HER2 expression (and protein expression in general) is not a binary, black-and-white parameter. The majority of cells—not just breast cancer cells—express HER2 to some extent; it is the exact levels of expression that matter. Normal breast tissue, and breast cancers that do not overexpress HER2, typically have 20,000 or fewer HER2 receptors per cell. Some breast cancers, which are said to be strongly HER2-positive, may express more than 2,000,000 receptors per cell—a hundred-fold increase. Importantly, however, many breast cancer cells fall somewhere in between these two extremes. Immunohistochemistry (IHC) tests for HER2 help pathologists to assess the levels of HER2 expression (the stain in the test kits will show up more strongly the higher the concentration of receptors) and subdivide them into 4 categories—but even within those categories, there is significant variation.
So what does this mean from a HER2-targeted drug point of view? The answer, again, is complex.

Several of the key early trastuzumab trials clearly showed a correlation between HER2 overexpression levels as measured by IHC, and the response rates to this HER2-targeted therapy. Importantly, the response rates were significantly higher in the IHC 3+ group, demonstrating that the cells that express the largest number of HER2 receptors are most likely to respond to HER2-targeted drugs. However, some IHC 2+ tumours were also shown to respond to trastuzumab; protein overexpression is certainly not an absolute marker for predicting response in this case.

Researchers also investigated the impact of gene amplification, which is a measure of how many copies of the gene are present in the cell’s DNA. A technique called fluorescence in-situ hybridisation, or FISH, is used to determine this: cancers are either said to be FISH+ (more than the usual number of copies present) or FISH- (usual number of copies present) for a specific gene, though in reality this is also a sliding scale. Often, as in the case of HER2, protein overexpression and gene amplification are closely correlated; on a cellular level, gene amplification (too many copies of a particular gene) usually leads to protein overexpression, as all those extra copies make extra proteins. Hence, nearly all IHC 3+ cancers are FISH+ for HER2, and nearly all IHC 0 or IHC 1+ cells are FISH-. The interesting group of cancers are the IHC 2+ ones, which can turn out to be either FISH- (showing a low response to HER2-targeted therapies) or FISH+ (showing a much better response to HER2-targeted therapies). Therefore, FISH testing is recommended for all HER2 IHC 2+ cancers before a treatment decision is made.

Interestingly, however, more recent analyses have shown that even some FISH- patients may be able to benefit from trastuzumab therapy; this means that even when using both IHC and FISH testing to guide treatment decisions, some women may miss out on treatment that could potentially benefit them. In summary, while HER2 testing via different techniques is a powerful tool to predict treatment response to targeted therapies, linking HER2 protein expression and HER2 gene amplification to predicted outcome is by no means an exact science.
What about genetic mutations? Mutations are essentially binary events: either the mutation is present or it is not. So are mutations absolute predictors of treatment response? It turns out that this is far from the case. Let’s take EGFR mutations in NSCLC as an example. It has been known for some time now that the presence of certain somatic mutations in NSCLC tissue, such as exon 19 deletions, are associated with a better response to EGFR tyrosine kinase-inhibitors (TKIs)\(^5\), such as erlotinib and gefitinib. These mutations are oncogenic: they do not (generally) appear in healthy tissue, which is said to be ‘wild type’ for the EGFR gene. Just like healthy cells, many NSCL cancers are also wild type EGFR, while a significant subset (15-35%, depending on cell type, ethnicity, gender, smoking history, etc.) do display one or more of the aforementioned EGFR mutations. As studies have shown, EGFR TKIs are more likely to lead to a response in patients whose tumours exhibit one or more of these so-called sensitising mutations.

Some of these tumours will go on to develop resistance mutations, such as T790M, which helps the cancer cells overcome the effect of the EGFR inhibitors. This is one of the reasons why not all EGFR-mutated NSCLC patients respond to EGFR TKIs. However, the reverse is also true: a significant proportion of EGFR wild-type tumours nevertheless respond to EGFR TKI treatment\(^6\).

**Clinical utility of testing: a delicate balance**

As both the trastuzumab and gefitinib example have illustrated, using companion diagnostics to determine treatment choice may sometimes lead to a certain subset of patients not receiving potentially beneficial drugs. So why are those tests being used in the first place?

*Fig 4. Attribute associations with different biomarker tests amongst oncologists and pathologists in the US. The percentages shown denote the proportion of respondents associating each test with that specific attribute. The closer the attribute is to the outer rim of the chart, the more important the attribute is in making a decision whether to use a test or not.*

*Source: Ipsos Healthcare MDx Monitor, US, 2015 Q1/Q2 (research conducted online March – May)*
While there is no universal consensus, a large proportion of oncologists and pathologists agree that these tests have real clinical utility (are accurate predictors of treatment response) and there exists a large amount of clinical data supporting this clinical utility. This is in spite of the factors discussed earlier; in other words, a marker does not have to be perfect for it to be perceived as useful for making treatment decisions. In many doctors’ minds, the fact that some patients may miss out on a drug that could potentially work for them is more than compensated for by the fact that the patients they are prescribing the drug to are more likely to benefit as a result of their biomarker status. Targeted therapies carry a significant financial cost, and sometimes come with potential side-effects. Having a tool (biomarker) to help make a benefit vs. cost assessment is generally seen as a positive attribute.

It should be pointed out that, ultimately, this type of assessment is carried out at a much higher level by authorities such as the FDA when it comes to making a decision on drug approval. Based on all the available clinical trial data to date – including response rates in specific sub-types – the approval bodies must decide whether it is better for the drug to be approved for all patients (and hence risk additional expenses/toxicities from these drugs in patient groups that have low probability of gaining a benefit) or only in patients with certain mutations/overexpression profiles (and hence risk excluding a potentially large percentage of the population for whom a small, but non-zero, likelihood of benefit exists). In all this, it should be noted that the relative importance of each of these factors in this intricate balancing act can be very different depending on whose point of view you consider: the patient, the physician, the drug manufacturer, the payer or the government.

---

**Fig 5. Simplified representation of assessing the relative pros and cons of a drug being approved for all patients vs for biomarker-high patients only.**
In the example shown in figure 5, there are a number of factors that have a clear impact on which way the balance of an all-patients approval vs. a biomarker-specific approval will shift:

1. The proportion of patients who are biomarker high vs. biomarker moderate/low (as represented by the orange vs. yellow/grey patients)
2. The proportion of patients with biomarker low status who benefit from the drug despite being biomarker low (as represented by the two grey patients who benefit vs. the six who don’t)
3. The proportion of patients with biomarker moderate status who do not benefit from the drug despite being biomarker moderate (as represented by the three yellow patients who don’t benefit vs. the four who do)

This is, however, a very simplified view. In reality, a number of additional factors also play an often highly significant role:

1. The cost of the targeted therapy, i.e. how much money is ‘wasted’ on patients who do not benefit from the drug if it receives a blanket approval?
2. The cost of testing, i.e. how much money is spent on finding out biomarker status if the drug receives a restricted approval? Note that this is not just about financial cost of the test itself, but also laboratory resources, time spent waiting for the test results, potential extra biopsies required, etc.
3. The likelihood of toxicity, i.e. how likely are patients who do not benefit from the drug to develop ‘unnecessary’ toxicities?
4. The duration of response / efficacy of the drug, i.e. how much incremental benefit does the drug provide in patients who do benefit, and does this differ by biomarker status? In other words, it’s not just about response rates, but also PFS / OS improvements

This overarching question – whether a drug should be approved for patients who exhibit a specific biomarker status only, even if it has been shown to (sometimes) work for those who don’t – has perhaps never seemed more important than right now, with the advent of immuno-oncology agents. The PD-1 and PD-L1-inhibitors, in particular, have shown great promise for NSCLC patients whose cancers exhibit overexpression of PD-L1. However, much like in the HER2 example, PD-L1 expression is very much a sliding scale and, importantly, even some patients with low PD-L1 expression have been shown to benefit. An extra complicating factor is that there is currently no consensus on how to define PD-L1 overexpression, a factor which we will address in further depth in the next section.
Different toolkits: how you measure matters

We saw in the previous section that, not surprisingly, not all markers are the same. Testing for HER2 overexpression is important in breast cancer (even if, as we saw, it does not give us an ideal picture), whereas testing for the same protein in lung cancer has very limited clinical utility (at least at present). Conversely, testing for EGFR mutations in breast cancer is currently not recommended; it is not seen as an actionable mutation. In short, what doctors test for matters.

The story does not end there, however. For many of these markers, there are many different ways to measure the level, presence or absence of the same marker using different approaches. Previously, we alluded to the difference between IHC and FISH for HER2 testing: the former measures the level of protein expression, the latter measures the level of gene amplification. As HER2 genes encode HER2 proteins, it is no surprise that there is a strong correlation between the two; generally speaking, the more strongly the gene is amplified, the higher the resultant protein levels will be. However, (1) this correlation is not absolute and (2) positivity by FISH has been shown to be a more direct way to predict treatment response to HER2 therapy than positivity by IHC. This is why, in those cases where a tumour has a ‘borderline’ HER2 IHC score of 2+, it is the FISH result that determines the most appropriate treatment choice. Simply testing with IHC tests, which are cheaper and faster than FISH tests, is not always sufficient for making the most informed choice.

On a more fundamental level, there are even different ways of testing for the presence or absence of specific mutations. The exon 19 deletions in the EGFR gene, and related mutations mentioned earlier are often tested through the use of RealTime PCR-based assays. However, sequencing techniques such as Sanger sequencing, pyrosequencing and – increasingly – Next Generation Sequencing (NGS) can also be used. There are also a growing number of variations on the PCR theme, including digital droplet PCR and PNA-mediated PCR clamping. Going into the exact differences between all these techniques is beyond the scope of this paper. One key difference between PCR and sequencing is that the former techniques are only able to look for a list of pre-determined mutations (e.g. T790M, L588R, etc.), whereas the latter can (in theory) sequence the whole gene, flagging any and all mutations (including rare ones, and those of uncertain significance). Next Generation Sequencing panels often include a wide range of different genes, so doctors can potentially get a report of a list of different mutations across a series of different genes, as opposed to a simple yes/no status for a limited number of specific mutations in one specific gene. Due to speed and efficiency advantages, as well as a continued reduction in cost, the use of NGS is on the increase; it is, however, still significantly lower than the use of more traditional, PCR-based assays.

![Fig 6: EGFR mutation testing methodology shares](source: Ipsos Healthcare Oncology MDx Monitor, US, 2015 Q1/Q2 (research conducted online March – May 2015))
For any given methodology, there is still a choice of different tests for the same marker. Going back to our earlier example of EGFR mutation testing, there are several different PCR-based tests available. These include: Roche cobas EGFR Mutation Test, Qiagen therascreen EGFR RGQ PCR Kit, several other branded kits, as well as a limitless number of laboratory-developed/home-brew kits. The Roche cobas assay detects 41 specific mutations\(^7\), the Qiagen therascreen assay detects 29 mutations\(^8\); the former has been approved by the FDA as a companion diagnostic for erlotinib\(^9\), whereas the latter was approved to select patients for treatment with gefinitib and gilotrif\(^10\). This differential approval, where similar tests are approved alongside specific and different drugs, is an important point which we will come back to later. It is also important to note that doctors, and in many cases will, make their treatment decisions based on non-FDA-approved, lab-developed tests, such as the chart below illustrates:

The availability of different test kits allows for greater flexibility and means laboratories are less limited by the exact type or brand of hardware (e.g. PCR machines) that they have at their disposal. However, the lack of standardisation can sometimes be a potential concern. For the example above, several studies have shown a very close (though not 100%) concordance between test results from the Roche cobas and the Qiagen therascreen assay, as well as Sanger sequencing. Apart from the fact that some mutations are not covered by both tests, they will almost always return the same test results when they are used to test the same sample, i.e. they are functionally interchangeable. This is not always the case, however, and this is becoming a particularly important issue for the emerging suite of PD-L1 expression tests. The following are examples of tests that have been used in several recent trials for PD-(L)1 inhibitors:

- Dako IHC assay, clone 28-8: ≥5% tumour cells showing staining defined as PD-L1 positive\(^11\)
- Dako IHC assay, clone 22C3: ≥1% tumour cells showing staining defined as PD-L1 positive in phase Ib trial\(^12\); ≥50% tumour cells showing staining defined as PD-L1 positive in other phase I trial\(^13\)
- Ventana IHC assay, clone SP263: ≥25% tumour cells showing staining defined as PD-L1 positive\(^14\)

As is obvious from the above, the different tests all use very different cut-off points to define PD-L1 positivity/overexpression. A patient that tested positive (e.g. 6% staining) with the Dako 28-8 assay may have tested negative with
the Ventana assay (note that not only the % cut-offs used in the trials were different, the assays have different sensitivities too, resulting in weaker or stronger staining).

At the time this article was prepared for publication, the FDA had just approved pembrolizumab to be used for second-plus line advanced NSCLC patients whose tumours overexpress PD-L1, as determined by the PD-L1 Dako IHC 22C3 pharmDx companion diagnostic, with a cut-off of 50%. Soon thereafter, another PD-1-inhibitor, nivolumab, was approved for this patient population, regardless of PD-L1 expression levels; however, the FDA noted that the level of PD-L1 expression in NSCLC tumours may help identify patients who are more likely to benefit, and approved the PD-L1 IHC 28-8 pharmDx test alongside the drug. In light of this, it will be interesting to see how future testing guidelines will reflect these recent developments, and provide some degree of standardisation against this very complex backdrop.

The future: accelerating change and multiple paradigm shifts

Despite the increasing fragmentation and growing number of clinically meaningful markers, companion diagnostics testing in oncology (and indeed, in medicine in general) is still in its infancy. HER2 testing has been around for nearly 20 years, and hormone receptor testing started many years earlier. Since then, the fundamental principles haven’t changed much: tissues are extracted from solid tumours, sent to a lab, analysed with single assays – or at the most, small panels – and results are fed back to physicians, who then make a decision based on a small number of actionable proteins or mutations. Even in adenocarcinoma lung cancers, where a significant number of new tests have been launched recently, there are ultimately only a few different molecular sub-types that a doctor will differentiate between when it comes to making a drug therapy choice:

- EGFR wild type, ALK wild type: treat with chemo or chemo + Avastin
- EGFR mutated: treat with an EGFR-TKI
- EGFR wild type, ALK mutated: treat with an ALK-inhibitor

For patients in the first category, some doctors will now also test for a handful of other markers, such as ROS-1, PD-L1 or cMET. Ultimately, however, the number of different patient types remains limited. The state of molecular diagnostics in oncology cannot yet be considered as true personalised medicine.

However, this is all about to change. This transition period of single-marker testing is a necessary, but likely brief, stepping stone towards mass profiling/sequencing of large numbers of genes, and making treatment decisions based on patients’ detailed genetic profiles. Next Generation Sequencing, alluded to in the earlier sections, has been gradually increasing in use over the past few years, and is getting close to being widely adopted. Also known as massively parallel sequencing or high-throughput sequencing, it allows for many different genes — often in the hundreds — to be sequenced in one run, making it much more efficient than PCR-based techniques or Sanger sequencing. There have been two major barriers to widespread adoption of NGS in the past: (1) its cost and (2) its clinical utility. The first barrier is rapidly disappearing, as the cost of NGS is decreasing significantly each month. The second barrier is harder to address, and comes down to the fact
that many of the genes profiled through NGS panels have not yet been extensively studied in clinical trials, leaving doctors unsure as to what to do with all that data (i.e. if there is a mutation in gene X but this mutation has not been shown to correlate with a better or worse response to any therapies, what is the clinical utility of this information?).

There are other revolutions taking place in the Oncology CDx testing market. For example, the increase in so-called liquid biopsies, which utilise blood or urine samples for testing biomarkers such as EGFR mutations, is a significant new development. Tissue biopsies are invasive, time-consuming and sometimes risky, and hence patients do not usually get re-tested to see if their biomarker status has changed; being able to do so through a simple blood draw or urine collection may significantly increase re-testing rates. On the one hand, this has the potential to make things a lot easier for both doctor and patient; on the other hand, the increase of re-testing, the tracking of biomarker status as a dynamic rather than a static factor, and the choice between solid and liquid samples add further layers of complexity to this ever-changing field.

Conclusion: from complexity to opportunity

Cancer is, ultimately, a genetic disease: it is always caused by genetic changes. We cannot hope to conquer cancer unless we fully understand the intricate relationship between the myriad genes and proteins that control the cell cycle and which, through random mutations, can and do lead to malignancies. As we uncover more and more of these genes and proteins, we are starting to realise that they hold the key to successful treatment: no two cancers are the same, and their subtle differences can only be identified by probing deep into their sub-microscopic worlds. The tools that enable us to do so are constantly evolving, allowing us to map those molecular differences more efficiently, in much larger numbers, less invasively, and more accurately.

It is easy to become overwhelmed by this flood of new data, new techniques, new methodologies, new targets, and even entirely new paradigms. This is a challenge not just faced by oncologists and pathologists, but also by those researching, developing and marketing new targeted anti-cancer therapies. Drug manufacturers now operate in a very different world than just a decade ago, and a solid CDx launch strategy is of vital importance. Such a strategy must ideally incorporate all the new advances, anticipate the upcoming revolutions, and ensure being a prescription driver – rather than a barrier – by navigating the ever-increasing complexity.

Cancer is complex, and an understanding of this complexity leads to opportunities: for doctors, for drug and diagnostics manufacturers and – most importantly – for patients.
About the Ipsos Oncology Center of Excellence

With over 100 people dedicated to oncology worldwide – and even more accredited through our Oncology Accreditation Program – Ipsos has the fluency in oncology to meet our clients’ every need. Our Oncology Center of Excellence includes a comprehensive suite of custom capabilities alongside market-leading syndicated services such as the Global Oncology Monitor, the Oncology MDx Monitor and the Oncology Lab Mapping Study. Our key differentiator is our ability to integrate all oncology solutions to offer in-depth insights that effect tangible change. With unrivalled strength and presence in both developed and emerging markets, we always deliver global expertise with local market knowledge.

Contact
To find out more, or to discuss your specific research needs, don’t hesitate to contact our Molecular Diagnostics team:

(Global) Pieter De Richter, Head of Molecular Diagnostics Portfolio
Tel: +601 2290 0315; Email: Pieter.DeRichter@ipsos.com

(US) Ayse Levent, Vice President, Global MDx Monitors
Tel: +1 646 313 7725; Email: Ayse.Levent@ipsos.com

www.ipsos.com