
Ira Mellman1, Vanessa M. Hubbard-Lucey2, Matthew J. Tontonoz2, Michael D. Kalos3, Daniel S. Chen1, James P. Allison4, Charles G. Drake5, Hy Levitsky6, Nils Lonberg7, Sjoerd H. van der Burg8, Douglas T. Fearn9, E. John Wherry10, Israel Lowy11, Robert H. Vonderheide10, and Patrick Hwu4

Abstract

With the recent FDA approvals of pembrolizumab and nivolumab, and a host of additional immunomodulatory agents entering clinical development each year, the field of cancer immunotherapy is changing rapidly. Strategies that can assist researchers in choosing the most promising drugs and drug combinations to move forward through clinical development are badly needed in order to reduce the likelihood of late-stage clinical trial failures. On October 5, 2014, the Cancer Immunotherapy Consortium of the Cancer Research Institute, a collaborative think tank composed of stakeholders from academia, industry, regulatory agencies, and patient interest groups, met to discuss strategies for de-risking immunotherapy development, with a focus on integrating preclinical and clinical studies, and conducting smarter early-phase trials, particularly for combination therapies. Several recommendations were made, including making better use of clinical data to inform preclinical research, obtaining adequate tissues for biomarker studies, and choosing appropriate clinical trial endpoints to identify promising drug candidates and combinations in nonrandomized early-phase trials. Cancer Immunol Res; 4(4); 279–88. ©2016 AACR.

Introduction

In business, “de-risking” refers to the process of making economic investments safer by reducing the possibility that bad things will occur and money will be lost. Until recently, cancer immunotherapies were seen as risky investments for biopharma companies because the likelihood of success in terms of clinical utility and financial return was small compared with the high chance of failure. Several experimental agents that reached phase III testing, for instance, ultimately failed to live up to initial expectations (1). With the success of CTLA-4–blocking antibodies, the climate for investment in this area is changing. Indeed, cancer immunotherapy is increasingly recognized as a “breakthrough” approach, one with the potential to offer long-term, durable responses to patients with advanced cancer (2).

The recognition that anti–CTLA-4 and anti–PD-1/PD-L1 act as potent levers for moving the immune response has also led investigators to reconsider many interventions, such as vaccines, that have previously failed to demonstrate benefit in clinical studies. These interventions are now back on the table for consideration in combination studies to augment antitumor activity. The therapeutic arsenal of agents for immunotherapy is thus expanding rapidly. Nonetheless, many questions remain about how to choose combinations with the greatest likelihood of success in terms of efficacy, safety, and feasibility of administration. In addition, sustainability in terms of cost and pricing is likely to continue to be a factor affecting the immunotherapy development pipeline, especially given expectations that many of these agents will be used in combination. In this context, it becomes imperative to increase the likelihood of success and reduce the rate of late-stage clinical trial failures.

On October 5, 2014, the Cancer Immunotherapy Consortium of the Cancer Research Institute, a collaborative think tank composed of stakeholders from academia, industry, regulatory agencies, and patient interest groups, met to discuss strategies for de-risking immunotherapy development, with a focus on integrating preclinical and clinical studies, and conducting smarter early-phase trials, particularly for combination therapies. Several recommendations were made, including making better use of clinical data to inform preclinical research, obtaining adequate tissues for biomarker studies, and choosing appropriate clinical trial endpoints to identify promising drug candidates and combinations in nonrandomized early-phase trials.
Panel 1: Role of Mouse Models in Immunotherapy Drug Development

The conventional approach to cancer drug development typically follows a linear, unidirectional flow from basic to preclinical to clinical research. Following an initial basic discovery, preclinical (typically mouse) in vivo experiments are conducted before moving the candidate drug into a clinical trial. Preclinical data are presumed to predict what will happen in humans, and serve as the basis for identifying promising new therapies.

This conventional approach has some notable successes to recommend it. Indeed, it was preclinical mouse models that Arlene Sharpe, Tak Mak, and colleagues used to confirm that CTLA-4 can function as a negative regulator of T-cell activation (3, 4), following the earlier in vitro work of Jeffrey Bluestone (5) and James Allison (6). Mouse models were also used by Allison and colleagues to establish checkpoint blockade as an effective approach to cancer immunotherapy (7). This preclinical work led to clinical trials of anti–CTLA-4 antibodies in humans, which culminated in the FDA approval of ipilimumab in 2011. Likewise, the potential synergy and therapeutic effectiveness of anti–CTLA-4 plus anti–PD-1 in treating advanced melanoma (8) was first demonstrated in a preclinical mouse model (9).

These significant breakthroughs notwithstanding, it is instructive to note that approximately 80% of phase II and 50% of phase III trials fail, primarily due to lack of efficacy (10). This low success rate is due, in part, to a large number of drugs with suboptimal preclinical validation moving through the pipeline (11). It also reflects the limitations of mouse models, which—although crucial to understanding basic biology—are not always useful for making predictions about clinical effectiveness in human (12). Mouse models may fail to accurately predict response in patients with cancer because they do not fully recapitulate the human tumor microenvironment. Genetically engineered mice (GEM) are an example of a model that is often not physiologically relevant, as although these models look like human cancers histologically, they contain few mutations, unlike their human counterparts. In addition, the equilibrium between the tumor and the immune system cannot be wholly recapitulated in mice. The tumors in mice only interact with the host immune system for a short period of time, on the order of months versus the multiple years we see in the clinical setting. Other differences include differences in pharmacokinetics and pharmacodynamics between mice and humans (differences that are exacerbated by the need for targeting the mouse ortholog of the relevant target with surrogate antibodies). Finally, mouse models may fail to reflect the incredible propensity for human tumors to respond adaptively (i.e., become resistant) to almost any new drug agent used. Although the human and mouse immune systems are fundamentally similar, a number of differences affect the regulation of the immune response and the cell types involved, which complicates direct translation from mouse to human.

For these reasons, we shall argue that the conventional approach of beginning with mouse models and working up to humans is not necessarily the best way to develop new, particularly combination, immunotherapy drugs. Considering the recent successes being achieved in the clinic, and as immunotherapy is increasingly turning toward combinations, we propose that our current model of immunotherapy drug development is backward. Rather than beginning with preclinical mouse models, might it not be more advantageous to start at the other end of the spectrum, with human clinical trial data? Given the impressive successes in the clinic, human data may now afford a more detailed and rapid analysis of the correlates and determinants of tumor response.

In this white paper, we call for a more dynamic and iterative approach in which clinical and preclinical studies are conducted simultaneously, and in which there is continual back and forth translation between the two. The following sections consider in more detail the rationale for better integrating preclinical and clinical studies and how such integration can be achieved, with particular attention paid to current questions being explored in the field.

Cautionary tails: Limitations of mouse models in predicting clinical responses in humans

Although mice are clearly necessary and useful for discovering new biology and dissecting biologic mechanisms (13, 14), they have some limitations when it comes to predicting the efficacy of potential immunotherapies, identifying predictive biomarkers, and selecting indications, among other variables (15, 16).

To take an example from the contemporary immunotherapy literature, Powles and colleagues have shown that PD-L1 expression is predictive of response to treatment with atezolizumab (anti–PD-L1; Roche/Genentech) in a phase I trial in metastatic urethral bladder cancer (17). Patients with higher PD-L1 expression both pretreatment and during treatment showed overall better responses to atezolizumab, with the amount of PD-L1 staining correlating directly with clinical response. This is in stark contrast to mouse models in which PD-L1 expression does not correlate with response to anti–PD-L1 treatment; for example, some high PD-L1–expressing cell lines do not respond to anti–PD-L1 treatment, whereas some low PD-L1–expressing lines do (Genentech, unpublished data).

This difference between mouse and human is most probably due to the fact that many syngeneic tumor mouse cell lines given to mice in in vivo studies often constitutively express PD-L1, whereas the expression levels in humans may be more dynamic. Evidence from tissue samples from patients being treated with agents that target the PD-1/PD-L1 axis demonstrates that PD-L1 expression is upregulated in response to IFNγ, a phenomenon termed adaptive immune resistance (18, 19). It has since been found that IFNγ-induced PD-L1 expression occurs in many tumor cell lines and in tumor-infiltrating immune cells (particularly myeloid cells) both of human and mouse origin, but PD-L1 expression is common in mouse models regardless of whether they respond (20). Recent data from the clinic have demonstrated that PD-L1 expression in humans could be useful for predicting whether or not to give a combination of anti–CTLA-4 and anti–PD-1 therapy (21). Nevertheless, one could not have predicted from mouse data that PD-L1 would be an effective biomarker for human response to anti–PD-1/anti–PD-L1 treatment.

In both bladder and non–small cell lung cancer (NSCLC), it was observed that PD-L1 expression by tumor-infiltrating immune cells was at least as predictive, if not more so, than PD-L1 expression by the tumor cells themselves (17, 22). This was an unexpected finding with implications for underlying mechanism and for the development of a useful companion diagnostic test. Indeed, a recent FDA approval included the mandatory use of a diagnostic test for PD-L1 in NSCLC for the anti–PD-1 drug pembrolizumab. A diagnostic test for the...
anti–PD-1 drug nivolumab for NSCLC was also approved, although testing for this drug does not mandate biomarker testing. If the best results are coming from humans, why not devise the techniques to use them?

With respect to checkpoint blockade, the first preclinical studies pointed us in the right direction: utilizing anti–CTLA-4 as an agent to promote tumor regression. These studies paved the way for success in the clinic (7, 9). Further studies have enabled us to dissect the biology of how this immunotherapy works (5, 23). However, in mouse models anti–CTLA-4 treatment was shown to cure 70% of the mice, whereas only ~20% of human melanoma patients have had durable responses. This finding, among others, suggests that we need to be careful about how we interpret results in mouse models, because the results will in large part reflect the design of the model. Thus, it will be important to better understand the limitations of particular models being used and to prioritize the models for testing hypotheses about basic biology derived from clinical data. This will entail developing specific questions and hypotheses that need to be asked and developing and using the most appropriate mouse models. For example, if the goal is to study an immunogenic tumor, using a mouse model where an immunogenic tumor can be utilized will be the most appropriate comparison, or alternatively utilizing a tumor with BRAF mutations in order to study a BRAF inhibitor (in conjunction with an immunotherapy; refs. 24–26). Syngeneic mouse models with transplantable tumors may be more relevant for evaluating checkpoint blockade, as over the years they have accumulated many more mutations, making them closer to many human tumors in this regard, but these, too, have limitations as discussed above.

Recent work in the area of optimizing the appropriate IgG framework for immunotherapy antibodies also highlights obstacles translating preclinical mouse work into our understanding of how these antibodies may function in humans (27, 28). For example, human IgG lacks an analogue of mouse IgG1 that contains the equivalent FcγRIIB binding ability. This naturally impedes the objective of translating the perfect preclinical antibodies into a drug to be used in the clinic and requires careful development of models to study them appropriately.

Lastly, studies of immunotherapy in mouse models cannot in many settings provide insights into the therapeutic index, which is the safety window in which to dose patients. In fact, whereas checkpoint blockade in patients frequently induces adverse events, these events are rarely observed in preclinical mouse models. As has been seen in dual checkpoint blockade, the safety and tolerability of immunotherapy in combination is a significant issue that can currently be fully understood only with careful studies in the clinic (21). It was only after immune-related adverse events were reported across multiple clinical trials with ipilimumab that a mouse model that was prone to autoimmunity was utilized to study hypophysitis in the mouse (29). This example highlights the next discussion, which includes the use of clinical data to guide appropriate mouse models for further study of biology.

The end is the beginning: Using clinical data to generate hypotheses about basic biology

An important question facing researchers in the field is why some cancer types are more responsive to immunotherapy than others. Melanoma, NSCLC, bladder, and renal cancer, for example, have proven to be relatively responsive to checkpoint blockade therapy, whereas pancreatic, colorectal, and ovarian cancers have not responded as well, if at all, to this approach (30–32).

Pancreatic ductal adenocarcinoma (PDA), in particular, has thus far proven to be extremely resistant to essentially all available immunotherapies, leading many researchers to wonder what explains this resistance. There are, in principle, several potential reasons why this might be the case. Perhaps PDA is a highly immunoeedited tumor, one that has already “escaped” surveillance by the immune system, though this seems unlikely based on the limited presence of host CD8+ T cells in tumor specimens. Or, perhaps tumor-associated T cells are suffering from exhaustion—albeit a type of exhaustion such that depletion cannot be rescued by existing checkpoint blockade antibodies. A third possibility is that T cells are simply not present in the tumor itself and may be “dominantly” excluded by other immunosuppressive cell types or proteins. Finally, it could be that the few mutations in PDA yield few immunogenic epitopes, which may limit the generation of an immune response.

Evidence to help distinguish these four possibilities comes from clinical data showing that increased numbers of T-cell infiltrates within the tumor microenvironment are associated with better survival in several tumors types, including colorectal cancer (33) and ovarian cancer (34). These data raise the possibility that diseases such as PDA may be resistant to current checkpoint blockade therapy because T cells are actively excluded from the tumor microenvironment. To investigate this possibility, Fearon and colleagues turned to the KPC mouse model of PDA, which appears to more faithfully recapitulate the pathophysiology and clinical profile of the human disease than transplanted models. Using this model, they showed that, indeed, T cells are restricted to the tumor stroma and do not enter the tumor microenvironment itself—consistent with the histopathology observed in human PDA. Further, one mechanism by which T cells are excluded from the microenvironment is through the overexpression of the homeostatic chemokine CXCL12 by stromal cells (35). This led Fearon and colleagues to wonder whether blocking the action of this chemokine could overcome the circumscribed movement of T cells. Administration of the FDA-approved AMD3100 (plerixafor), which antagonizes CXCR4, the CXCL12 receptor, along with anti–PD-L1, led to some evidence of tumor regression and an increase in the numbers of T cells at the tumor site in this mouse model. Expression of CXCL12 has been associated with poor prognosis in human cancer (36–38). Of course, other possibilities besides migration could be mediating the observed effect. As the tumor microenvironment in PDA is considered quite immunosuppressive, including the presence of indoleamine 2,3-dioxygenase (IDO), regulatory T cells (Treg), and the secretion of TGFβ, these factors could also be limiting the effect of checkpoint blockade (39, 40).

Another possibility is that the tumor-infiltrating T cells are kept at bay by co-infiltrating suppressive myeloid cells, present in great numbers in ovarian cancer and PDA (41–44). It is still unclear how all the effects of the tumor microenvironment act in concert on the adaptive immune response to tumors, and therefore much work is yet to be done. However, a phase I clinical trial of AMD3100 in pancreatic cancer is currently under way in the UK and will certainly inform our understanding of this pathway as it relates to PDA.

Also in PDA, Vonderheide and colleagues have provided evidence that tumor-induced immune suppression can be reversed...
by activating CD40 on macrophage populations (13). In one clinical trial where patients were given the fully human CD40 agonist antibody CP-870,893, 21 patients experienced a median progression-free survival that was nearly double that seen in patients who received gemcitabine alone (5.6 months vs. 2.3 months). Posttreatment biopsy samples taken from patients whose tumors regressed showed that numerous macrophages were present (and not many leukocytes). This prompted further studies in mouse models of PDA in order to fully understand the underlying biologic mechanism of anti-CD40 agonism. Administration of anti-CD40 agonist antibody in mice that lacked CD4+ or CD8+ T cells resulted in tumor regression, whereas blockade with clodronate-encapsulated liposomes did not result in tumor regression. It was determined that CD40 activation could promote antitumor immunity via the infiltration of CD40-activated macrophages. These examples serve as an illustration of how human clinical data can be used to formulate a hypothesis about biologic mechanism, which can then be appropriately and productively tested in a preclinical mouse model that recapitulates known features of the human disease.

Not mutually exclusive: Using mouse models to guide our clinical approaches

One of the most significant and looming questions within the field of cancer immunotherapy is why some patients within a given tumor type respond to immunotherapy whereas others do not. For example, only about 20% of patients with advanced melanoma respond to treatment with ipilimumab (45), and a similar proportion of NSCLC patients respond to pembrolizumab or nivolumab (31, 46). Identifying surrogate biomarkers that could be used to tailor treatment to those most likely to benefit are greatly needed.

Research from Wherry and colleagues suggests that mouse models are useful in identifying potential predictive biomarkers to anti–PD-1 therapy. Prior work from this lab had established phenotypic and transcriptional differences between functional memory and exhausted T cells in the context of chronic viral infection (14, 47), including the delineation of inhibitory receptor coexpression patterns that can guide combination therapies for reversal of exhaustion (48). Recent research has centered on identifying a transcriptional ‘signature’ reflective of T-cell exhaustion and extracting information from these signatures that can provide insights for reversal of exhaustion (49, 50). Such a signature can be defined by the changes in activation state and proliferation of key subsets of exhausted T cells defined by coexpression patterns of T-bet or eomesodermin (Eomes) in combination with PD-1. Wherry presented data showing that one subset of exhausted T cells defined as T-bethiPD-1–/– represents the majority of cells that get reinvigorated by PD-1 blockade, but gains rise, through extensive proliferation, to an EomeshiPD-1–/– subset that is intrinsically resistant to rescue. Changes in markers of recent proliferation and induction of this EomeshiPD-1–/– subset can therefore be used as a surrogate of reinvigoration of the T-bethiPD-1–/– pool. Indeed, evidence of expression of granzyme B and Kif7 (markers of activation and recent proliferation) in EomeshiPD-1–/– TILs was predictive of response to PD-1 therapy in mouse models. Data from a combination ipilimumab + radiotherapy clinical trial in patients extended this analysis and illustrated that such surrogate markers of changes in exhausted T-cell subpopulations may predict clinical outcomes. This research suggests that relative expression levels of Eomes and PD-1 may be a relevant functional biomarker for those cancer patients who may benefit from anti–PD-1 therapy.

Recent work by Gajewski and colleagues has added to our understanding of biomarkers for immunotherapy. In clinical studies, they found that melanoma patients who had a type I Interferon transcriptional profile and CD8+ T cells were more likely to respond to immunotherapy. However, even if we enrich for this population, most likely less than half of those patients would respond to the therapy. They found that a subset of melanoma patients with T-cell–inflamed tumors had an increased number of Tregs in addition to a high expression of IDO and PD-L1. Using mouse models, they performed mechanistic studies revealing that the upregulated expression of IDO and PD-L1, as well as the recruitment of Tregs, was dependent on CD8+ T cells (20). In addition to demonstrating that the expression of immunosuppressive molecules is immune intrinsic (likely the result of immune-mediated feedback loops), these results also indicate that the presence of several immunosuppressive factors could be a barrier to response to immunotherapy, and potentially targeting these additional barriers could have therapeutic potential. A number of clinical studies (NCT02178722, NCT02318277, NCT02298153, NCT01604889, and NCT02073123) are under way to determine the effect of combining IDO blockade with checkpoint blockade. It is important to note that these studies have been complemented by a number of studies that were conducted several years prior, demonstrating the relevance of IDO expression in human tumors (51–53).

One way in which to support the clinical development process is to include the use of pharmacodynamic (PD) biomarkers to aid in establishing clinical efficacy in a small phase I trial (Fig. 1). As an example, if there is a drug/antibody in development with the purpose of reducing Tregs within the tumor, testing would include utilizing this antibody in a couple of different mouse models (that have Tregs in their tumors). In this example, after administration of the antibody, the number of Tregs decreases; however, there may be marginal antitumor activity. The conclusion to be drawn from this example would be that there is no clinical efficacy (tumors did not decrease in size) but there is a PD biomarker that showed an effect (reduction in Tregs). Because the PD biomarker demonstrated an effect, under these circumstances the clinical development process may proceed to testing this antibody in humans. This may occur in a small phase I clinical trial in which careful preand posttumor biopsies were available for analysis to determine if the number of Tregs decrease. There may be a hypothesis as to what this agent could be combined with in order to provide clinical antitumor efficacy. Is there another relevant cell type that could also be playing a role in this indication? Would checkpoint blockade be needed here? These examples highlight the idea that mouse models can be used to guide our clinical approaches, but their usefulness must be validated with small sets of clinical data from phase I studies, particularly as we move into complex combination trials.

Panel 2: Getting the Most Out of Phase I Studies

With the discovery of new immune checkpoints beyond CTLA-4, and the development of additional immune modulating
agents by numerous biopharmaceutical companies, the immunotherapy landscape is changing rapidly. In addition to understanding the mechanisms of how these agents work as monotherapy, the race is on to discover the most effective combinations, particularly for those patients who do not respond to monotherapy alone. The impressive results of the combination ipilimumab + nivolumab trial for melanoma (8), in which 88% of patients in the optimal dose cohort were still alive at 2 years, has set a new bar for the potential therapeutic success of combinations in melanoma, and has focused attention on how to best evaluate new combinations in clinical trials.

Determining which combinations work best for which indications, and for which patients, while also minimizing toxicity, is a daunting task facing the field. In addition to the rapidly increasing number of available agents, the empirical evidence of clinical benefit is outpacing our understanding of the underlying immune biology. Without a deep understanding of the relevant biology, and without defined biomarkers to correlate with response in patients, our ability to design smart early-phase studies of combination agents is hindered. There is a great need for rational strategies to improve our ability to conduct smarter early-phase studies, and thereby reduce the likelihood that ill-fated combinations, doomed by either limited efficacy or unacceptable toxicity, will find their way into expensive phase II/III trials.

What counts as a winning combination? Approaches to identifying effective immunotherapy drug combinations

The rationale for combining immunotherapies with each other and with additional modalities is that combinations may achieve enhanced response rates, progression-free survival, overall survival, and perhaps most importantly, durable responses. Because of the complexity and multistep nature of the immune response to tumors, it may be necessary to intervene at different points in that process in order to bring about an effective anticancer immune response. The "cancer-immunity cycle" (54) conceptualizes the different steps involved (Fig. 2). The availability of multiple potential targets for immunotherapies presents both opportunities and challenges. On the one hand, multiple drug targets mean that we have numerous possible ways to turn a nonresponding tumor into a responding one. On the other hand, with so many such combinations possible, we need ways to assess the additive benefits of each intervention. This naturally leads to the question: How do we know if our combination is more clinically active than either single agent alone? Will we know it when we see it? This question is especially relevant in cases of early phase, nonrandomized studies utilizing active agents in combination. In such situations, due to statistics related to small studies, lack of comparator arms, and known background activity, biomarkers may prove useful. For example, if the combination of drug X (e.g., an anti–PD-L1 antibody) and drug Y (e.g., a myeloid-derived
suppressor cell (MDSC) depleting agent] is believed to be synergistic or additive, because drug Y sensitizes tumors to drug X, then one may be able to observe within the tumor tissue a depletion of MDSC, enhanced PD-L1 expression, or T-cell infiltration due to drug Y. Incorporating relevant biologic correlative data will enhance our ability to interpret complex and nuanced biologic responses to drug combinations and can help reduce the likelihood of taking a neutral (no added activity over either monotherapy) or antagonistic combination (less activity than either agent alone) through to phase III.

Another strategy for distinguishing the response of the combination from that of the monotherapy is to incorporate additional meaningful endpoints. For example, if the overall response rate (ORR) is 30% for a monotherapy, what response rate would you need to observe in order to have confidence that your combination is superior to the monotherapy alone? When performing small early-phase trials (for example, where n = 20), you would have to observe a very high response rate to feel confident to move the combination forward. Therefore, it may be more useful to use complete response (CR) rates as an endpoint in highly responsive tumors (assuming that CR rates to monotherapy are relatively low) to determine if the combination is promising. However, other biologic endpoints may be needed for tumors that have lower response rates. Incorporating meaningful endpoints that can differentiate the combination therapy from the monotherapy will be necessary to de-risk immunotherapy-based combinations.

A further consideration when performing early-stage combination trials is understanding who your patients are, including which prior therapies they may have received. Prior immunotherapy may influence response to the combination, which could also confound interpretation of the results. Is the combination generating a new immune response in a patient or just releasing an existing immune response from inhibition? Given that it is highly unlikely that a single immunotherapy treatment approach will work for all patients with cancer, understanding the “right patient” for the “right drug” may help us interpret activity observed in small early-phase studies of monotherapy or combination immunotherapies.

Drug developers should utilize strategic approaches to identify effective immunotherapy drug combinations. They can identify and design rational combinations for patients depending on patient characteristics. Once a clinical concept is designed, the treatment plan for patients should be approached depending on the various features of their tumor. If the patient’s tumor has limited antigens, potentially chimeric antigen receptor (CAR) T-cell therapy, a bispecific antibody, or potentially a personalized neoantigen vaccination, could be used to try to boost the limited potential

Figure 2. Cancer-immunity cycle. The cancer-immunity cycle is a multistep process that involves (1) release of cancer cell antigens; (2) cancer antigen presentation; (3) priming and activation; (4) trafficking of T cells to the tumor; (5) infiltration of T cells into tumors; (6) recognition of cancer cells by T cells; and (7) killing of cancer target cells. Possible intervention points can be identified by asking the following questions: A, Are there sufficient number of tumor-specific T cells? B, Is there trafficking and penetration of tumor by T cells? C, Are activated anticancer T cells inhibited in the tumor microenvironment (TME)? MDSC, myeloid-derived suppressor cells. Adapted with permission from Chen and Mellman (54).
responses. Many tumors have few T cells and pro-inflammatory antigen-presenting cells, creating a need to prime T cells (potentially with an OX40, 41BB, and/or vaccination combination if the antigen is known, and anti–CTLA-4, IL2, etc.). Many tumors have T cells in the vicinity, but because of the tumor microenvironment they may not be able to get inside the tumor. Some solutions to this problem could include intratumoral administration of an activating agent that promotes T-cell trafficking. In addition, within the tumor microenvironment specific cell types often directly inhibit the activation of T cells; therefore, a combination that includes inhibiting Tregs or MDSCs might be beneficial. Lastly, due to the highly immunosuppressive microenvironment, a combination that blocks IDO, TGFβ, or adenosine might prove useful.

High-throughput systems biology for immune monitoring: What, when, and how?

The successful clinical development of immunotherapies is hindered by our lack of comprehensive mechanistic knowledge about the agents being tested. This fundamental limitation precludes the truly rational development and testing of hypotheses about how a particular immunotherapy or combination is working and forms the basis for the argument for implementing comprehensive biomarker strategies in clinical trials. A comprehensive biomarker strategy would apply a diverse set of platforms to evaluate clinical samples and, at the same time, embrace the pursuit of hypothesis-generating efforts in the context of testing specific hypotheses (55).

One strategy to identify biomarkers that are relevant to the mechanistic understanding and treatment efficacy for immunotherapy trials is to approach the problem in a "systems biology" manner. This process requires coordinating biomarker efforts across the entire developmental process from product characterization, preclinical discovery assays, correlational studies, and clinical readouts, and applying high-throughput assays that capture large amounts of quantitative information from limited amounts of material. Multiparametric platforms and assays that capture diverse sets of information include cell-based platforms to interrogate immune cell phenotype and function, nucleic acid, soluble or proteomic assay systems that allow for simultaneous and quantitative analyses of large numbers of analytes from a single small sample; high-throughput T-cell receptor (TCR) spectratyping strategies to understand the breadth of antigen-specific T-cell responses; and high-throughput seromic strategies to evaluate in an agnostic manner antitumor antibody responses. As an example, CyTOF—short for cytometry by time of flight—allows investigators to look at up to 70 markers at a time in single cells in a given sample, and has permitted investigators to profile the immune system in a way not previously possible (56). All of these methods have the potential to provide mechanistic insights about when, why, and how an immunotherapy treatment is effective (55).

For instance, high-throughput and agnostic seromics-based studies on CAR T cell–treated patients suggested that there may be epitope spreading following therapy, indicative of a systemic tumor-specific adaptive immune response (57). In addition, the identification of new T-cell epitopes utilizing various bioinformatics tools and prediction algorithms can enable the identification of T cells that are reactive to these epitopes for further functional testing of their target cell killing in vitro (58). Application of agnostic cytokine analysis also revealed the unexpected but considerable elevation in IL6 indicative of severe cytokine release syndrome (CRS) in a subset of patients treated with CAR-based T-cell therapy. This observation resulted in the implementation of tocilizumab treatment as an effective mitigation strategy for this dangerous syndrome (59, 60). More recent implementation of broad and agnostic biomarker strategies has led to critical observations about the nature of the T-cell response to PD-L1 checkpoint therapy and has provided critical mechanistic insights about which patients might respond to this therapy and potentially why (19, 22).

One limitation with the broad implementation of high-throughput and systems biology-based biomarker strategies is the inherent variability of the biology being studied, which confounds efforts to apply quality-supported infrastructure to the studies. For example, intra- and interpatient sample variability in immune cell subset composition, interpatient genetic variability, and prior and concurrent treatment medications are all parameters that compromise our ability to establish validated assays to generate biomarker data sets. Establishing robust assays in terms of specificity, accuracy, and calibration between different institutions is also very often challenging, due to an inability to broadly standardize research-based efforts across laboratories. The inability to practically implement assays standardization in this setting has been addressed by the development and application of the concept of assay harmonization, an iterative process through which critical variables that impair assay performance are identified and eliminated, resulting in an operational standardization in the quality of data output from sites that use laboratory-specific protocols. The CIC/CRI Immunoassay Proficiency Program provides a useful model of how harmonization panels can be used to identify deficiencies in current procedures and sources of variability within and among institutions. This program has led to the establishment of protocols to optimize assay performance within several assay panels (Elispot, serum, CFSE proliferation, Luminox, intracellular cytokine stains, multimer, and flow cytometric gating strategies; ref. 61).

Ultimately, the expectation is that application of broad, agnostic, and quality-controlled biomarker support platforms will generate data sets that will enable investigators to generate comprehensive hypotheses related to functional and mechanistic aspects of immunotherapies, which can then be tested prospectively in more advanced clinical trials. Although it is clear that assays that are applied to clinical samples must be supported by quality-enabled infrastructure, a fair question to ask is whether a signal of clinical efficacy is necessary prior to the application of broad, agnostic, and quality assays to a particular immunotherapy trial. An argument can be made that in any clinical trial setting, the minimal additional resources required to perform biomarker studies is offset by the potential to understand not only how a particular treatment is working, but also, perhaps as importantly, why a particular treatment may have failed.

Vaccines and the importance of T-cell priming: It’s okay to say the “V” word

As impressive as checkpoint blockade therapy has been for some cancers, others have been notoriously resistant. Can we boost the responsiveness of these cancers to checkpoint blockade so that additional patients respond? One way to approach this problem is to consider again the nodes of the cancer-immunity cycle. For example, do appropriately activated effector T cells exist? If effector T cells are present in sufficient numbers, have they succeeded in migrating into the tumor microbed? Are the T cells tolerized by the tumor or can they recognize tumor antigens?
And most importantly, are they cytotoxic and do they secrete IFNγ and TNFα?

There are several possible scenarios that may help to explain why different tumor types respond to immunotherapies in different ways. First is the tumor against which a T-cell response has been generated, resulting in partial elimination, equilibrium, and eventually escape: resistance to immunotherapy is the eventual result in this scenario. Other tumors are targeted by a T-cell response that is subsequently suppressed by host/tumor immunosuppression: these tumors are good candidates for combination therapies such as checkpoint blockade therapy, along with inhibitors of other immunosuppressive factors such as IDO and adenosine, among others. Finally, are the tumors that exhibit strong immunosuppression (or are immunologically silent from the beginning due to lack of strong antigens such as those caused by mutation, for example) and against which a T-cell response is not generated. An emerging hypothesis is that these "immune privileged" tumors might be treated successfully with a combination of vaccine and checkpoint inhibitors (or other regulatory factors), but not by either agent alone. This is because effective clinical results from checkpoint blockade may require a preexisting T-cell response (19, 62). Therefore, administering a vaccine if the antigen is known, or activating CD40, could be necessary to generate the initial immune response that can then be sustained using checkpoint blockade therapy.

In the case of pancreatic cancer, it seems likely that inefficient T-cell priming, few candidate antigens for T cells to target (a consequence of the small mutational burden), and T-cell exclusion from the tumor microenvironment are parts of the explanation for why this cancer type has so far been resistant to checkpoint immunotherapy. This cancer type might benefit from a strategy that incorporates a vaccine if an appropriate target antigen can be identified (63).

What do we mean by "vaccine"? In the modern era, a vaccine is any modality that primes an immune response, including traditional tumor vaccines, personalized neoantigen vaccines, chemotherapy, radiotherapy, antitumor monoclonal antibodies, anti-angiogenesis drugs, T-cell therapies, oncolytic viruses, and even surgery. Anything that can overcome the host/tumor immune-mediated suppression and prime an immune response might enable checkpoint blockade to be more effective, particularly in tumor types where there have been limited responses. In other words, "It’s okay to say the V word." Naturally, we still have much to learn when it comes to understanding precisely how to vaccinate (systemic or in situ?), what to vaccinate with (TLR agonists, STING agonists, cytokines?), what antigens beyond neoantigens from mutations represent good candidates, and in which setting these situations may be applicable.

What we do not know at present is how much a lack of T-cell priming is really the problem when patients are nonresponsive to immunotherapies. Is it 5% of cases, 10%, or possibly much more? This remains to be determined, as well as determining what represent good targets, if the detected mutations do not encode MHC-binding neoepitopes in the host, and highlights the need to further explore combinations of traditional and non-traditional vaccines with checkpoint blockade in order to enhance T-cell priming.

Historically, cancer vaccines have targeted shared cancer antigens, such as the differentiation antigen Melan-A and the cancer/testis antigens MAGE and NY-ESO-1 (64). However, it is now recognized that neoantigens unique to each patient may be an important target for immunotherapy as well. Several recent papers make a strong case that neoantigens play a central role in tumor immunity. Utilizing exome sequencing, Yadav and colleagues found only a few candidate neoantigens that were immunogenic in vivo: these mutations were encoded by genes that had nothing to do with cancer initiation and progression, suggesting that passenger mutations could alter immunogenicity (65). Further supporting this idea, Schreiber and colleagues used exome sequencing to identify two neoantigens that underlay an effective response to checkpoint blockade therapy in a mouse model of cancer. When utilized in a vaccination protocol along with the adjuvant poly(I:C), these neoantigens were capable of inducing tumor rejection at levels comparable with checkpoint blockade itself (66).

Vaccine formulations with novel adjuvants are needed that will allow innate immune cell activation to effectively prime the host against specific antigens. The HPV16-induced premalignant lesions and cancers may serve as an example here. Extensive studies show that still only 40% of the HPV-infected patients display intratumor T-cell response against the virus, despite the presence of immunogenic viral oncoproteins in every tumor cell (67), and most patients have either no or weak responses. Vaccines may be able to overcome this discrepancy, although the HPV antigens have not proven to be highly immunogenic, and many inhibitory mechanisms may need to be targeted with other modalities to reproducibly achieve therapeutic benefit (61, 68).

This role for vaccines in combination therapy also suggests that a more appropriate criterion for gauging the "biological success" of a cancer vaccine under development is the induction of a change in the composition, magnitude, and functionality of infiltrating T cells, rather than pure clinical measures of single-agent activity that may be thwarted by adaptive resistance mechanisms. Simply put, each component of a combination therapy "needs to do its job," achieving a necessary, but insufficient step along the path to tumor rejection.

Conclusions

In the next 5 to 10 years, investigators in the field of cancer immunotherapy will have access to an unprecedented number of agents and potential combinations—more than we are likely to know what to do with. In this context, it is imperative to have a smart approach to designing both preclinical work and early-phase clinical trials.

Historically, the direction of drug development research has flowed from preclinical mouse work to human clinical trials. With the explosion of new data available from immunotherapy clinical trials, we now have the importunity to use this clinical work to inform our basic research in ways not previously possible. Our contention is that, in this environment, preclinical mouse work and human clinical data need to work together simultaneously, to guide each other, and to fill in mechanistic gaps that will aid in the development of better immunotherapy drugs, combination therapies, and clinical trials. The process should be a mutually bidirectional process, where well-thought-out preclinical studies aid in clinical development, and carefully performed correlative clinical research studies (that could be further studied in a mouse model) are performed prior to a phase II/III study. Given the current successes being achieved in the clinic, the time may be right to use the human data to generate hypotheses, which can then be tested in mouse models.
We also need to improve the way we design early-stage clinical trials. Choosing meaningful endpoints is one way to design better phase I trials of combination immunotherapies. When sample sizes are small, for example, it may be more useful to perform biopsies and look for biomarkers as an endpoint to determine if the combination is worth taking forward through the drug development pipeline.

Understanding the basic biology of different treatments can enable us to incorporate relevant biomarkers into early-phase studies and thereby avoid taking a neutral or antagonistic combination through to later-phase studies. Identifying better biomarkers will require a comprehensive biomarker strategy that utilizes a standardized variety of phenotypic and biochemical methods to perform immunomonitoring.

In addition, as the field progresses, it will be important to combine targeted therapies and immunotherapies and to bring molecular biologists together with immunologists so we can uncover how the genetic pathways within tumors influence the immune microenvironment. In this context, it is sobering to realize that, despite billions of dollars spent on genetic tumor profiling, we know of no somatic mutation that predicts response to immunotherapy. Such a predicament indicates that with regard to drug discovery we cannot engage in business as usual. We must bridge the silos that exist not only between basic and clinical researchers, but also between cancer immunology and the rest of the oncology community. Improved communication and deeper partnerships are fundamental to allow the promise of cancer immunotherapy to revolutionize cancer care.

Disclosure of Potential Conflicts of Interest

C.G. Drake receives commercial support as a principal investigator for Aduro Biotech, Bristol-Myers Squibb (IloN), and Janssen; has patents licensed with Potenza Biotech, AZ Medimmune, and Bristol-Myers Squibb; and serves as a consultant/advisory board member for Bristol-Myers Squibb, F-star, Merck, Roche/Gercentech, Astellas, Novartis, Tizona, Potenza Biotech, and Agensys. D.T. Fearon has an ownership interest in Myostis. E.J. Wherry serves as a member of the Scientific Advisory Board of Surface Oncology. S.H. van der Burg reports receiving commercial research support from ISA Pharmaceuticals and is a consultant/advisory board member for ISA Pharmaceuticals. No potential conflicts of interest were disclosed by the other authors.

Published online April 1, 2016.

References
