Review

Assays for Monitoring Cellular Immune Responses to Active Immunotherapy of Cancer

Timothy M. Clay, Amy C. Hobeika, Paul J. Mosca, H. Kim Lyerly, and Michael A. Morse

Departments of Surgery [T. M. C., A. C. H., P. J. M., H. K. L.], Medicine [M. A. M.], Pathology [H. K. L.], and Immunology [H. K. L.], Duke University Medical Center, Durham, North Carolina 27710

Abstract

Numerous cancer immunotherapy strategies are currently being tested in clinical trials. Although clinical efficacy will be the final test of these approaches, the long and complicated developmental pathway for these agents necessitates evaluating immunological responses as intermediate markers of the most likely candidates for success. This has emphasized the need for assays that accurately detect and quantitate T cell-mediated, antigen-specific immune responses. This review evaluates the currently used in vivo and in vitro methods of assessing T-cell number and function, including delayed-type hypersensitivity, tetramer analysis, ELISPOT, flow cytometry-based analysis of cytokine expression, and PCR-based detection of T-cell receptor gene usage or cytokine production. We provide examples of how each has been used to monitor recent clinical trials and a discussion of how well each correlates with clinical outcome.

Introduction

The development of strategies for actively stimulating immunological rejection of tumors, previously an elusive goal, has been accelerated by demonstration of the prerequisites for antigen-specific immunity that has revealed numerous avenues for delivering antigens for presentation to T cells. Whole tumor vaccines mixed with adjuvant, gene-modified tumors, tumor antigen-encoding viral vectors, protein and peptide antigen, and dendritic cells loaded with antigen are all being studied in clinical trials. To promote a candidate to an evaluation in a large-scale clinical trial, it is usually necessary to demonstrate that the treatment has a significant impact on an intermediate predictive of clinical outcome. For cytotoxic agents, this marker is typically tumor regression. For agents not expected to cause tumor regression but that still may have a beneficial effect, a biological marker is usually chosen based on the presumed mode of activity. For immunotherapy, such a marker would be stimulation of a tumor antigen-specific immune response detectable by one or more immunological assays. Although effectors such as monocytes, natural killer cells, and antibodies may have an important role in antitumor immunity, most consider it vitally to use assays that evaluate the number and function of CD8+ CTLs that directly recognize tumor peptides presented by MHC molecules on the surface of a tumor cell as a trigger for direct cytolysis, and CD4+ helper T cells, particularly T helper type 1 responses, that lead to CTL generation. A number of assays show promise as methods for quantifying and characterizing the T-cell response to immunizations and for serially monitoring these responses. These tests of immunity include in vivo functional measures, in vitro phenotypic assays, and in vitro functional assays. In this review, we will initially discuss these assays and how they have been used thus far in clinical trials. Subsequently, we will compare their performance as intermediate markers of clinical benefit and conclude by reviewing important considerations for choosing immune assays.

In Vivo Measures of Antigen-specific Immunity

DTH.3 In the DTH test, antigen in the form of soluble protein alone or as antigen loaded onto antigen-presenting cells is injected intradermally, and the diameter of erythema or induration after 48–72 h is measured. CD4+ T helper cells that recognize the antigen presented on local antigen-presenting cells mediate the response by releasing cytokines that increase vascular permeability and recruit monocytes and other inflammatory cells to the site. Less frequently, a similar response may be mediated by CD8+ T cells (1). The cutoff for a positive response has not been standardized nor has the dose for DTH testing, although protein antigens are generally administered as 10–50 μg in 0.1 ml. This low dose is considered small enough that it does not induce a systemic immune response or cause excessive skin toxicity but is of a sufficient magnitude to induce a detectable local response.

DTH remains one of the most frequent immune tests performed in immunotherapy studies (2–4), but several issues must be taken into account. The first is whether the DTH response is truly antigen specific. Thurner et al. (5) vaccinated patients with peptide-loaded DCs and detected induration and erythema at the injection site in 7 of 11 patients but also found similar results for DCs not loaded with any antigen. Conversely, in our own studies, some patients without obvious induration or erythema had infiltrates of T cells at the injection site in skin biopsies

1 Supported by NIH Grants U01CA72162-03, 5-P01CA47741-08, and 1-P01-CA78673-01A1 and the C. Douglas McFadyen Fund. M. A. M. is a recipient of an American Society of Clinical Oncology Career Development Award and is supported by NIH Grant M01RR00030.
2 To whom requests for reprints should be addressed, at Box 3233, Durham, NC 27710. Phone: (919) 681-8350; Fax: (919) 681-7970; E-mail: m.morse@cgct.duke.edu.

3 The abbreviations used are: DTH, delayed-type hypersensitivity; DC, dendritic cell; IL, interleukin; PBMC, peripheral blood mononuclear cell; TCR, T-cell receptor; CMV, cytomegalovirus; CDR, complementarity determining region; V-D, variable-diversity; D-J, diversity-joining; ELISPOT, enzyme-linked immunospot; LDA, limiting dilution analysis.
taken after DTH testing with DCs loaded with carcinoembryonic antigen peptide (6). Other components of the immunizing agent may also contribute to the DTH response. For example, intradermal granulocyte/macrophage-colony stimulating factor, a component of some vaccine strategies, by itself, may induce a granulocyte/macrophage-colony stimulating factor-specific DTH reaction (7). Some authors have grown the cells to be used for immunization in fetal bovine serum, which contains proteins that may serve as potential immunogens. The second issue is whether the DTH reaction should be measured as an all-or-none end point or whether it may possess a “dose-response” relationship with the immunizing agent. Schreiber (3), for example, observed that the diameter of erythema and induration at an autologous, unmodified tumor injection site increased with each immunization (except the last) and was greater with higher cell doses, suggesting the possibility of correlating dose and immune response. The third issue, the concordance of DTH with other immune assays, has not been clarified entirely. Disis et al. (8) observed that DTH induration of ≥10 mm (but not 5–9 mm) correlated with a positive HER-2/neu-specific T-cell proliferative response (stimulation index, >2.0) in patients immunized with HER2/neu peptides. Fourth, more data are needed to support a correlation of clinical outcome with DTH response because the number of patients studied in any one trial has been small. Nesle (9) immunized melanoma patients with DCs loaded with MAGE-1, MAGE-3, MART-1, gp-100, or tyrosinase peptides along with keyhole limpet hemocyanin by direct intra-lymph node injection. Nine of the 12 immunized patients developed DTH responses to peptide-loaded DCs, and 5 had diameters >10 mm; 4 of these patients had clinically detectable tumor regression. Clinical trials of autologous colon cancer plus bacillus Calmette Guérin have suggested that those who develop DTH reactivity have a greater chance of remaining disease free than those who do not (10). One obvious criticism, though, is that the ability to mount a DTH reaction to an antigen is merely a marker for a patient with a better performance status and greater likelihood of more favorable outcome than an individual who cannot mount a response. Critical to answering this criticism will be the inclusion of control antigens in immunizations strategies. Patients who respond to the control antigen but not the tumor antigen could then be compared with those who respond to both to determine whether outcome is more closely related to tumor antigen-specific response.

In our opinion, DTH testing is useful as a preliminary screen for whether more detailed immune analysis is likely to be fruitful, and because it is relatively straightforward to perform and may possibly serve as an in vitro measure of the trafficking of lymphocytes to sites of tumor antigen, DTH testing will likely remain part of the immune analysis of cancer vaccines. The ability to isolate, expand, and assess the phenotype or function of antigen-specific T cells from skin biopsies of DTH sites may serve as an additional strategy to evaluate antigen-specific immune responses (11).

**In Vitro Measures of Immune Response**

**Sources of Lymphocytes for in Vitro Immune Analysis.** In vitro immune analyses require an adequate source of T cells with a frequency and functional activity reflective of the true immune response to the immunization. Clearly, peripheral blood is the most convenient source of T cells, but at least one study has questioned whether peripheral blood T-cell activity correlates with clinical response (12). Lee et al. (12) vaccinated patients with gp100 peptide with or without IL-2 and observed that despite detecting antigen-specific T cells in peripheral blood of some individuals immunized with gp100 alone, none had clinical signs of tumor regression. Conversely, although no antigen-specific T cells could be cultured ex vivo from the PBMCs of gp100 plus IL-2-treated patients, these were the only individuals in whom tumor regressions occurred. One possible explanation is that the antigen-specific T cells had migrated out of the peripheral blood, perhaps into tumor or other tissues. Although tumors may contain antigen-specific T cells (13), the detection of a lymphocytic infiltrate in a tumor has not uniformly correlated with an improved prognosis in cancer patients, and in one study, tumor-infiltrating lymphocytes were shown to have defects in the expression of the TCR-associated molecule CD3, specifically the ζ chain (14). Regional lymph nodes draining the immunization site may contain the most recently stimulated T cells, but it has been shown that even healthy, non-tumor-bearing individuals may have lymph nodes harboring MART-1-specific T cells (15). Finally, T cells specific for the antigen of interest have been cloned from DTH sites, and although this may serve as a surrogate for tumor infiltration, the conditions at a skin injection site not infiltrated with tumor are likely to be different from tumor tissue itself. Therefore, despite the theoretical concerns, sampling of peripheral blood lymphocytes has remained the standard. Important considerations for peripheral blood sampling include the timing of collection before and after immunization and whether to perform the analyses “real-time” on fresh specimens or simultaneously on cryopreserved cells.

**In Vitro Phenotypic Measures of Antigen-specific Cellular Immune Responses**

**Analysis of T-Cell Receptor V Region Usage.** The magnitude of an antigen-specific immune response may be determined by enumerating T cells according to a phenotypic marker such as the TCR using flow cytometric or PCR-based techniques. Expansions of TCRs expressing particular variable (V) region-α or V-β chains may be detected by flow cytometry using antibodies that recognize different variable or joining region subfamilies of the TCR α or β chains. An increase in the number of cells expressing a particular J-α, J-β, V-α, or V-β chain would indicate development of oligoclonality, a possible sign of induction of a specific immune response (16). This approach has limited value for a number of reasons: (a) only a minority of T cells expressing a particular J-α, J-β, V-α, or V-β combination will be specific for a particular antigen; (b) the response to most antigens is quite diverse and uses many different joining and variable regions; (c) monospecific antibodies are not available for all J region or V gene subfamilies, and therefore, this analysis is incomplete at best. Nonetheless, if the antigens that are the target of the immune response are unknown, this method may still have some usefulness.

**Peptide MHC Tetramers.** More recently, it has become possible to visualize antigen-specific T cells under flow cytometry by using soluble, fluorescently labeled, multimeric MHC-peptide complexes (17) that bind stably, specifically, and avidly
to antigen-specific T cells. During flow cytometric analysis, one can gate on the CD8+ T cells and look for expression of antigen-specific TCRs. The largest body of data regarding the usefulness of tetramers is derived from studies of viral epitope-specific CTLs. Analysis of peripheral blood T cells specific for CMV and EBV demonstrated that between 0.2 and 2.5% of circulating CD8+ cells were specific for peptides representing these antigens. Some authors have found correlation of MHC tetramer positivity and cytotoxicity in traditional microcytotoxicity assays (17), and the intensity of staining of CD8+ T cells with peptide MHC tetramers appears to correlate with T-cell avidity for the antigen (18), but tetramer positive cells occasionally fail to kill targets expressing the specific epitope (16). Several recent studies have demonstrated the utility of flow cytometric analysis using peptide MHC tetramers to quantitate CD8+ T cells specific for tumor antigens or control antigens used frequently in immunotherapy protocols (12, 19, 20). Dunbar et al. (21) have used peptide MHC tetramers to allow the selection of antigen-specific T cells from peripheral blood or lymph nodes by cell sorting. These selected T cells were cloned for further analysis and were shown to respond to specific antigen by cytokine production.

Although peptide MHC tetramers are powerful tools, they have certain limitations. They can only be used to detect immune responses to known antigens, because the peptide of interest must be loaded into the peptide MHC tetramer and thus must already be known and synthesized. Additionally, only class I MHC tetramers have been available routinely for widespread use, although class II tetramers have been described. Finally, because of the exquisite sensitivity of peptide/MHC tetramers for quantitating antigen-specific T cells, an interesting question has been raised regarding whether CD8(+) cells that bind to peptide/MHC tetramers are naïve or antigen-experienced (“memory” T cells). Pittet et al. (15) observed that 10 of 13 melanoma patients and 6 of 10 healthy individuals had high frequencies (≥ 1 of 2500 CD8+ T cells) of Melan-A-specific cells in the peripheral blood. All of these Melan-A-specific cells from the healthy individuals and seven of the patients displayed a naïve CD45RA(lo)/RO(-) phenotype. In three of the patients, “memory” CD45RA(lo)/RO(+) Melan-A-specific cells were observed. In contrast, influenza matrix-specific CTLs from all individuals exhibited a CD45RA(lo)/RO(+) memory phenotype. One patient was observed to have an evolution of the Melan-A-specific cell phenotype over time. This suggests that in addition to simply detecting peptide MHC-positive cells, it may be important to assess whether antigen-specific cells are naïve or memory T cells to determine whether the detected antigen-specific T cells have been stimulated by the immunization strategy.

TCR Complementarity Determining Region 3. Antigen-specific T cells may also be phenotypically detected by PCR techniques for detecting a restricted TCR repertoire (22) by sequencing the third CDR (CDR3) of the TCR. The CDR3 region encodes the highly polymorphic portion of the TCR responsible for recognizing peptide-MHC complexes. For the β chain, the CDR3 region encodes the V-D segment and D-J segment junctions, whereas for the α-chain, it encodes the V-J junction. Using V, D, or J region subfamily-specific PCR primers, PCR may be performed to detect the development of restricted TCR gene usage (23, 24). Some studies in melanoma patients (25, 26) and renal cell carcinoma patients (27) have detected a restricted TCR gene usage. However, other studies in melanoma have found unrestricted TCR gene usage (28, 29). It is too soon to determine the role of this technology in monitoring immune responses in clinical trials, and more studies are needed. In particular, it has been used primarily as a qualitative measure of skewing of the T-cell repertoire, and it may not be possible to easily correlate its results with clinical outcome. Nonetheless, its advantages include the small amount of specimen required, the ability to perform the analysis from cells directly isolated from the blood to avoid introducing biases caused by ex vivo expansion, and the reproducibility and internal controls that permit analysis of samples collected at different times. Recently, a more automated and rapid fluorescence-based method for CDR3 length analysis of expressed TCR gene families that was able to distinguish between polyclonal, oligo-clonal, and monoclonal CDR3 distributions has been developed (30).

In Vitro Functional Measures of Antigen-specific Immune Responses

- T-cell number and function may be monitored by assays that detect T cells by an activity such as cytokine production, proliferation, or cytotoxicity.
- Lymphoproliferation Assay. The ability of T cells to proliferate in response to antigen has traditionally been used as an indicator of the presence of antigen-specific CD4+ helper T cells. Typically, the specimen of purified T cells or PBMCs is mixed with various dilutions of antigen or antigen in the presence of stimulator cells (irradiated autologous or HLA matched antigen-presenting cells). After 72–120 h, [3H]thymidine was added, and DNA synthesis (as a measure of proliferation) was quantified by using a gamma counter to measure the amount of radiolabeled thymidine incorporated into the DNA. A stimulation index can be calculated by dividing the number of cpm for the specimen by the number of cpm in cells incubated without antigen as a control.

The proliferation assay has been used frequently in clinical trials to compare T-cell responses before and after immunization (3, 4, 31–34). Depending on the immunization strategy, a small percentage of patients (31, 32) to as many as half (33) or all (34) patients have been found to respond by proliferation assays. Its major advantage is the ability to perform the assay directly on peripheral blood samples, giving a picture of the T-cell activity present in vivo (although the in vitro culture period can introduce artifacts in the results.) Its drawbacks are that it does not measure activity with direct mechanistic relevance to tumor rejection, it has not yet been convincingly correlated with clinical outcome (3, 31), it can be influenced by the nonspecific immune function of the patients, and the stimulation index does not necessarily correlate with the number of antigen-specific T cells present in vivo. High levels of proliferation by a few cells or low levels of proliferation by many cells would give a similar stimulation index. A recent flow cytometric assay measuring

---

4 T. Clay, unpublished observation.
distribution of cell membrane dyes into daughter cells produced during proliferation permits the number of antigen-responsive cells in a stimulation assay to be determined (35).

Detection of Secreted Cytokines by ELISA and ELISPOT Assays. Cytokine secretion by T cells in response to antigen may be detected by measuring either bulk cytokine production (by an ELISA) or enumerating individual cytokine producing T cells (by an ELISPOT assay). In the ELISA assay, PBMC specimens are incubated with antigen (with or without antigen-presenting cells), and after a defined period of time, the supernatant from the culture is harvested and added to microtiter plates coated with antibody for cytokines of interest such as IFN-γ, TNF-α, or IL-2. Antibodies ultimately linked to a detectable label or reporter molecule are added, and the plates are washed and read. Generally, a single cytokine is measured in each well, although a recently described modification permits up to 15 cytokines to be measured in a single sample (36). In this procedure, antibodies to cytokines of interest are covalently bound to microspheres with uniform, distinctive proportions of red and orange fluorescent dyes. Detection antibodies conjugated to a green fluorescent reporter dye are added, and flow cytometry is performed. By gating on a particular orange/red fluorescence indicating a particular cytokine of interest, it is possible to quantitate the amount of cytokine that is proportional to the amount of green fluorescence. ELISA has been used for monitoring in several clinical trials (3, 4, 33), although the definition of a positive result differs (e.g., an amount of IFN-γ that is 2 times greater than control wells). Because this is an assay of the cytokine production of a population of cells, it does not give information about individual cells and cannot be used to enumerate the antigen-specific T cells. Furthermore, it does not measure the actual cytokine profile of these cells in vivo but rather, the ability of the cells to secrete cytokine when exposed to an antigenic stimulus. The ELISA assay can also be used to determine the levels of cytokines in serum or other body fluids. Although this may give a broad picture of the inflammatory state of a patient, it cannot be used to evaluate the cytokine profile in the milieu of the actual tumor cells. Flow cytometry-based assays (as will be discussed later) have also been found to be more reliable by some authors (37).

The basic steps of an ELISPOT assay (38) are: (a) coating a 96-well microtiter plate with purified cytokine-specific antibody; (b) blocking the plate to prevent nonspecific absorption of random proteins; (c) incubating the cytokine-secreting T cells with stimulator cells at several different dilutions; (d) lysing the cells with detergent; (e) adding a labeled second antibody; and (f) detecting the antibody-cytokine complex. The product of the final step is usually an enzyme/substrate reaction producing a colored product that can be quantitated microscopically, visually, or electronically. Each spot represents one single cell secreting the cytokine of interest. The antigen-specific T-cell precursor frequency is determined by dividing the number of spots (cytokine-secreting cells) by the number of cells placed into the well. The ELISPOT assay has been shown to reliably detect the number of antigen-specific T cells in experiments in which known quantities of antigen-specific T cells were added to bulk PBMC preparations (39). Miyahira et al. (40) have reported that the CTL precursor frequency provided by ELISPOT assay is comparable with that obtained by the limiting dilution analysis. Although rigorous statistical analysis has not been performed yet, there is interest in determining whether the ELISPOT assay correlates with survival (41). In a retrospective analysis of melanoma patients immunized with a polyclonal vaccine (42), MAGE-3 and Melan-A/MART-1-specific, IFN-γ-secreting T cells were enumerated by ELISPOT. Those who demonstrated antigen-specific T-cell secretion of IFN-γ had a longer recurrence free survival (>12 months) than nonresponders (3–5 months).

Because the task of counting the number of spots visually becomes difficult and time consuming with large numbers of spots (>100), computerized plate readers using digital cameras have been developed (43). In our opinion, the computerized methods provide superlative discrimination of antigen-specific responses from background and make the ELISPOT an excellent choice for an immune monitoring assay in a large-scale study. Further modifications that may increase the usefulness of the ELISPOT include a dual color method for evaluating two different cytokine release patterns at a time (44) and the use of PBMCs loaded with poxvirus vectors encoding the antigen of interest as stimulators so that individuals of any HLA type (instead of just well-known HLA types) may be included in analyses (45).

Detection of Intracellular Cytokines by Multiparameter Flow Cytometry. It was first demonstrated in murine models that different patterns of cytokine secretion could be used to differentiate between memory/effector T cells with different immune functions (46). The two patterns, T helper 1 with secretion of IL-2, IFN-γ, and TNF-α, and T helper 2 with secretion of IL-4, IL-5, IL-6, IL-10, and IL-13, also appear to have human counterparts. Thus, it is possible to monitor immune responses in humans by characterizing the cytokine secretion pattern of T cells in peripheral blood, lymph nodes, or tissues by flow cytometry (reviewed in Ref. 47). Most methods involve a short period (4–6 h) of in vitro T-cell activation (using antigen or mitogens) and source of stimulator cells (autologous antigen-presenting cells or PBMCs). For the last 3–4 h of stimulation, cytokine secretion is prevented by the addition of brefeldin A. After this stimulation period, most protocols recommend staining with fluorochrome-conjugated anti-CD4 and anti-CD8 antibodies to allow gating on T cells and anti-CD69 to monitor activation of T cells. The cells are then fixed and permeabilized and stained with an antibody to the cytokine of interest (e.g., IFN-γ or IL-2). We have found that it is also possible to fix and permeabilize the cells followed by staining for surface and intracellular proteins. Three- or four-color flow cytometry is performed to enumerate the percentage of CD4(+) or CD8(+) CD69(+) cytokine(+) T cells. This assay has been modified so that it may be performed on PBMCs (48) or whole blood (49). Typical T-cell percentages for various antigens range from 0.1% (e.g., measles or mumps antigen) to 5% (CMV antigen; Ref. 47) or more.

Some studies with serial analysis of intracellular cytokine induction have demonstrated correlation with clinical outcome. In a Phase I/II study of immunization with SRL 172 in patients with stage IV malignant melanoma, lymphocyte activation was assayed prior to each vaccine administration using a fluorescence-activated cell sorter-based intracellular cytokine assay (50). Induction of intracellular IL-2 production was associated
with improved survival. Surprisingly, induction of IFN-\(\gamma\) or both IL-2 plus IFN-\(\gamma\) was not associated with improved survival, demonstrating the complexities in choosing surrogate markers. Reinartz et al. (51) followed intracellular cytokine production in T cells obtained at various time points during immunization of ovarian cancer patients with the anti-idiotypic vaccine ACA125. Early in the immunizations, predominantly IL-2 and IFN-\(\alpha\) were observed, but later a T helper 2 pattern was observed. This correlated with generation of anti-anti-idiotypic antibodies and prolonged survival.

The major drawback of this approach is that the cells detected are no longer viable, and thus cells cannot be sorted and cultured to produce clones. Recently, a novel method that uses flow cytometry to detect surface-expressed cytokines was described (52). Magnetofluorescent liposomes containing several thousand fluorescein molecules and colloidal magnetic particles and conjugated to antibodies specific for IFN-\(\gamma\) and IL-10 were shown capable of detecting surface expression of these molecules on 12.5 and 34.8% of T cells. Most of these cells were shown to have intracellular cytokine when they were permeabilized for analysis. The cells remain viable so they may be sorted for use in other assays. This method would not be applicable for all cytokines because some, such as IL-2, IL-4, and IL-5, cannot be detected on cell surfaces.

**Measurement of Cytokine mRNA Levels by Real-Time Quantitative RT-PCR**

Quantitative RT-PCR is a highly accurate molecular method for measuring the levels of transcripts of a gene or genes of interest in sample RNA (53). Kruse et al. (54) applied the technique to the analysis of cytokine mRNA from cryopreserved normal donor blood samples. Recently, Kammula et al. (55) used the technique in clinical trials of melanoma peptide-based vaccines to detect antigen-specific T-cell responses by comparing pre- and postvaccine samples from melanoma patients. Peripheral blood samples and tumor tissues obtained by fine needle aspiration were evaluated. For PBMC samples, the method entails thawing cells into fresh medium, allowing them to recover from thawing for 10 h, and then incubating the cells for an additional 2 h with either the peptide used in the vaccine or an irrelevant peptide, followed by total RNA isolation. Quantitative RT-PCR was then used to measure cytokine mRNA levels in the samples. Data were normalized to expression of a control gene, such as CD8. This study showed that quantitative RT-PCR can be used to detect antigen-specific T-cell responses in peripheral blood samples. Additionally, localization of antigen-specific T cells to tumor sites was demonstrated by analysis of biopsy samples without any in vitro stimulation step. Further studies are needed, in part to determine the relative sensitivity of this methodology, including comparative studies against other assay techniques, and to provide additional studies so that the reliability of the method may be assessed.

**Direct Cytotoxicity Assays.** Cytotoxicity assays are appealing because measurement of the ability of CD8+ CTLs to lyse tumor is thought to be a relevant marker for in vivo antitumor activity. The microcytotoxicity assay involves mixing the specimen containing T cells or PBMCs with antigen-expressing targets loaded with \(^{51}\)Cr or europium and measuring the release of the chromium or europium after target cell lysis. Because autologous tumor is often difficult to obtain, surrogate targets are often used, such as HLA-matched allogeneic tumor cell lines, and targets that can be loaded with the antigen of interest (such as autologous DCs loaded with peptide or genetic material encoding the antigen, or T2 cells loaded with peptide). Targets sensitive to natural killer cell lysis (K562 and Daudi) are also included to determine the level of nonspecific lytic activity. The percentage of lysis of the targets after incubation for 4 h is calculated by comparison with the maximum achievable lysis of the target. Using different E:T ratios, it is possible to derive a value for the potency of cytotoxicity measured in “lytic units,” the number of T cells needed to achieve a stated amount of lysis. Theoretically, lytic units can be used to compare various CTL preparations. Cytotoxicity assays have been used for immune monitoring in studies of passively delivered T cells (56) and active immunotherapy approaches (31, 57).

One drawback to the microcytotoxicity assay is its relative insensitivity. Although bulk CTL assays represent a useful technique to give high versus low or + versus − readouts, they are not particularly quantitative. Furthermore, there is a need to stimulate the CTLs multiple times before testing their lytic activity (31), because it is unusual to find antigen-specific lysis by cells directly isolated from the peripheral blood, even in vaccinated patients (58). These multiple stimulations may distort the composition and activity of the T-cell population from its original state. Also, as discussed above, because autologous tumor is difficult to obtain, other targets must be used that may not reflect the actual ability to lyse autologous tumor cells in vivo. For example, tumor cells may down-regulate their MHC molecules or up-regulate their own Fas ligand, causing T-cell apoptosis. It has also been shown that the CTL response is heterogeneous with different levels of avidity for the antigen (59). Because targets used for in vitro testing usually express high levels of antigen, lysis may not reflect the ability to lyse tumor in vivo if the in situ tumor expresses low levels of the antigen. Finally, questions as to the correlation with clinical response have been raised. In at least one study, clinical regressions were observed in two patients in the absence of CTL activity (57). Modifications to the cytotoxicity assay that make it simpler and more reproducible are being developed including flow cytometric techniques for separating dead (lysed) target cells stained with propidium iodide from live (unlysed) cells stained with a cyanine membrane dye.

**Quantifying CTL Precursors by LDA.** LDA, a cumbersome but more quantitative assay for CTL precursor frequency, correlates T-cell number from a functional activity. LDA analyses involve the serial dilution of T cells in a very large number of wells, followed by an in vitro stimulation phase and target lysis phase. Poisson distribution analysis is applied to the results to determine the proportion of wells at a particular T-cell dilution that have \(>1\) antigen-specific precursor at the start of the stimulation. Analysis of the frequencies of positive wells in successively higher dilutions as a function of log (T cells/well) generates a line, the slope of which is proportional to the precursor frequency. In addition to being cumbersome, labor intensive, and extremely operator dependent, the LDA is also flawed by the intrinsic assumption that a single antigen-specific T cell can be expanded during the stimulation phase to generate a signal above a mathematically determined threshold.
Because the LDA assay is somewhat complicated and because a large number of cells are required for testing antigen specificity by LDA, it has been used in few published studies. Möller et al. (60) evaluated the T-cell response to immunizations of melanoma patients with IL-2 gene-modified autologous tumor cells using LDA and found that after vaccination, PBMCs contained an increased number of tumor-reactive proliferative as well as cytolytic cells. In three of six patients, the frequencies of antimalanoma cytolytic precursor cells increased between 2.6- and 28-fold. Two of these patients showed a minor clinical response. The same group demonstrated similar results with IL-12 gene-modified melanoma cells (61). D’Souza et al. (62) evaluated Melan-A/Mart-1-specific CTL precursors in melanoma patients using LDA and further demonstrated that they expressed a memory phenotype. More recently, Romero et al. (63) made additional modifications of the assay to increase the ability to quantify the number of precursors. Thurner et al. (5) used this modification, which involves one cycle of in vitro stimulation with antigen before testing the cells for cytotoxicity, to measure MAGE-3A1 peptide-specific immune responses in a study of immunizations with DCs loaded with MAGE-3A1 peptide. Eight of 11 patients were found to have increases in MAGE-3A1 CTL precursor frequency after immunization. Nonetheless, it is likely that this assay will be less frequently used in the future as newer assays that are more sensitive become accepted.

Comparison of the Assays

There is a paucity of studies that have directly compared the various assays for their performance in evaluating immune responses, and most of the data comes from studies of responses against viral antigens. Tan et al. (64) showed that estimates of CD8+ T-cell frequencies specific for EBV-related antigens varied considerably according to the method used. Values obtained from MHC-peptide tetramer staining were, on average, 4.4-fold higher than those obtained from ELISPOT assays, which were, in turn, on average, 5.3-fold higher than those obtained from LDA. Tetramer staining showed that as many as 5.5% circulating CD8+ T cells in a virus carrier were specific for a single EBV lytic protein epitope. Kuzushima et al. (65), using EBV-specific T-cell lines, confirmed that flow cytometric analysis is more sensitive than LDA for CTL precursors and ELISPOT in detecting IFN-γ-producing T cells. The results of direct T-cell staining using multimeric peptide-MHC complexes raise important questions about the meaning of precursor frequencies estimated from LDA. One possible explanation for this discrepancy is that only a fraction of cloned T cells are lytic; however, functional assays on sorted tetramer binding cells argue against this. Another major difference between the LDA and the direct detection assays, such as tetramer staining, is that the LDA depends on cell division. Greater than 10 divisions of a single precursor would be necessary during the stimulation phase of the LDA to register as a positive response. In cases of chronic viral infection, the precursor frequencies estimated by LDA appear to be closer to those estimated by direct staining with multimeric MHC-peptide. The LDA may therefore give a more meaningful figure of T cells with long-term growth potential.

Considerations for Choosing Immune Assays

Before choosing one or more immunological assays to monitor induction of antitumor immune responses, it is important to consider the performance characteristics of the assay in detecting immune responses. What magnitude of the immune response should be considered a positive response, and whether the assay results actually predict clinical outcome. Desirable performance characteristics of an assay for detecting T-cell responses would include: (a) adequate sensitivity, specificity, reliability, and reproducibility; (b) measurement of the true state of in vivo T-cell activity without introducing significant distortions; (c) simple and rapid to perform; and (d) requirement for only small quantities of specimens. As described above, tetramer analysis is highly sensitive, followed by ELISPOT. In our experience, the reliability of tetramers varies with some preparations, yielding no staining of T-cell clones, and also, because not all peptides form functional tetramers, tetramers with an untried peptide must be tested empirically. ELISPOT assays digitally analyzed have considerable reliability in our hands, but there is little published data on interlaboratory reproducibility. Tetramer analysis is quick to perform, whereas the other flow cytometric assays of T-cell cytokine production take ~8 h to perform and analyze. ELISPOT plates can be prepared in bulk in advance, and by using automated pipettes, plate washers, and plate readers, large numbers can be set up quickly and efficiently. The PCR-based techniques for detecting TCR gene usage or cytokine mRNA transcription require the smallest quantity of specimens.

The cutoff that should be accepted as indicative of an effective level of immunological response is entirely unknown (and may vary for each assay), but it is necessary to make educated guesses. A reasonable starting point for trying to determine what constitutes a clinically relevant immune response is the experience in animal models. Cure of a murine sarcoma required infusion of 3 × 10^4 T cells with receptors specific for the rejection epitope, if the sarcoma had been established for 3 days. It is estimated that this represents 0.2–0.5% of the circulating leukocytes (67). Because previous studies of adoptive immunotherapy for malignancies in humans have used fairly nonspecific T cells, it is difficult to find similar human data. In one study of stem cell transplant recipients at risk for EBV-associated lymphoproliferative disorders, two to four infusions of as few as 10^7 EBV-specific CTLs/m², starting from the time of maximal virus load, resulted in a 2- to 3-log decrease of virus titers (68). In patients who develop lymphoproliferative disorders, infusion of a similar number of EBV-specific CTLs can eradicate the tumors (69). If there are approximately 1.5–4.5 × 10^9 CD8+ T cells in the circulation (70), then at the time of infusion, the EBV-specific T cells would represent as many as 1 of every 100 CD8+ T cells. Thus, we propose that a level of peripheral blood antigen-specific CD8+ T cells in the range of 1% may be necessary to cause tumor remission. Of course, antiviral T-cell responses tend to be of greater frequency and avidity than antitumor responses, and

5 A. Hobeika, unpublished observations.
thus, whether the T-cell response to viral antigens is relevant to antitumor responses is unknown.

Correlation with clinical outcome is the most critical issue for any intermediate marker. Markers along the pathway to the ultimate mechanism for clinical benefit are the most desirable. Next best are markers that correlate very closely with outcome. Just as tumor regression after administration of a cytotoxic agent does not always result in survival benefit, development of an immunological response may not necessarily predict clinical outcome. For immunological assays, the most likely reason for this is that the immune responses that cause tumor regression are not known with certainty, and the available immunological assays may not measure these mechanisms. Tumors may be destroyed by CTLs through insertion of perforins and delivery of granzymes, by Fas-Fas ligand interactions, or by cytokine-mediated toxicity. Standard cytotoxicity assays only measure the direct lysis. Modifications of cytotoxicity assays to measure target cell apoptosis may be necessary to measure Fas-Fas ligand-mediated interactions. It is possible that none of the currently available assays actually measures a function with direct relevance to how tumors are actually attacked immunologically in the body. This demonstrates the importance of collecting data on correlation of immune response with clinical outcome whenever possible.

Conclusions and Areas for Further Investigation

The development of assays for detecting immunological responses to cancer vaccines is essential if these strategies are to be optimized. Standards are needed for performing the assays and interpreting the results. Agreement is needed on whether to analyze samples directly isolated from blood or lymph nodes or after a period of in vitro stimulation. Because the various assays yield estimates of antigen-specific T cells that sometimes differ in magnitude, it is critical to compare the immune response detected by a particular assay in a particular patient with the immune response specific for a well-established, immunogenic antigen, such as EBV or CMV peptide. Reproducibility between laboratories and correlation among immune assays requires rigorous evaluation. Because most of the immune assays do not measure an activity with direct relevance to tumor cell killing by the immune system, the importance of rigorous statistical analysis to determine the assay with the greatest correlation with outcome is necessary. Currently, several different assays are necessary until it can be established which correlate the best with clinical outcome. In our opinion, although tetramer analysis has high sensitivity, ELISPOT is more versatile for monitoring clinical trials and is more readily performed, given the current limited availability of tetramers.

References


Downloaded from clincancerres.aacrjournals.org on October 16, 2017. © 2001 American Association for Cancer Research.
1134 Immune Responses to Immunotherapy


Assays for Monitoring Cellular Immune Responses to Active Immunotherapy of Cancer


Updated version  Access the most recent version of this article at: http://clincancerres.aacrjournals.org/content/7/5/1127

Cited articles  This article cites 64 articles, 28 of which you can access for free at: http://clincancerres.aacrjournals.org/content/7/5/1127.full#ref-list-1

Citing articles  This article has been cited by 17 HighWire-hosted articles. Access the articles at: http://clincancerres.aacrjournals.org/content/7/5/1127.full#related-urls

E-mail alerts  Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions  To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions  To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.