ELISA / Grand Serology Analytical Validation — Mt Sinai CIMAC

Version 2.1

Original Date: May 9, 2018
Date revised: July 25, 2019

Performance Lab:

Human Immune Monitoring Center
Miriam Merad, MD, PhD, Director
Sacha Gnjatic, PhD, Associate Director
Seunghee Kim-Schulze, PhD, Facility Director
Adeeb Rahman, PhD, Technology Development Director
1470 Madison Avenue, New York, NY 10028

The table below lists the nature of initial results produced (primary output), QC steps and analyses, and eventual output.

Assay type	Primary assay outputs	Pre-processing/ Normalization /QC	Initial analyses	Derived data outputs
ELISA and grand serology	Antigen-specific antibody titers	Check titration curves from internal controls, extrapolation based on internal controls	Normalized antibody titer after extrapolation.	Antigen specific antibody titers and fold change scores based on pre-determined cutoffs

ELISA and GRAND SEROLOGY	
(i) accuracy	Determination of titration based on extrapolated curve, providing values with reproducible >90% within a <2x range.
(ii) precision	Ability to detect antigen specificity is dependent on format of antigen chosen (peptide, protein, denatured vs. native form). Nevertheless, high level of precision shown with comparing closely related antigens to irrelevant targets. Intra-assay and inter-assay CV calculated for NY-ESO-1 at 1.1% and 1.2% respectively. CVs are larger for titers near the limit of detection.
(iii) analytical sensitivity	Titers range from 1/1 (non significant) to greater than 1/1,000,000 and are dependent on number of dilutions tested.
(iv) analytical specificity including interfering substances	Specificity is determined by comparing reactivity to several antigens, including negative controls such as DHFR. Potential for “stickiness”, i.e., low-level titers (range 10-500) to a majority of antigens tested is observed in some samples, particularly after chemotherapy (possibly due to hemolysis in part). Procedures are in place in SOP to minimize this rare (<5%) occurrence.
(vi) reference intervals (normal values) with controls and calibrators	Positive controls and expected OD ranges (<300 for negative control sera at 1/100, >1000 for positive controls sera at 1/100) and are specified in SOPs for each antigen tested in each plate.
(vii) standardization, harmonization, reproducibility and ruggedness	Assay was harmonized with Roswell Park, NordWest, and biotech company Serametrix, using reference specimens coordinated by the Gnjatic lab.
(viii) establishment of appropriate quality control and improvement procedures	Use of negative and positive serum control with known reactivity in each plate for each antigen, including healthy donor pools used as negative reference and known sera with reactivity as positive.
(ix) any other performance characteristics required for assay performance	Serum and plasma from either peripheral blood or bone marrow were found to be interchangeable for antibody titer determination. Urine and other fluids such as culture supernatants may be used but starting with undiluted material.

1. Purpose of assay

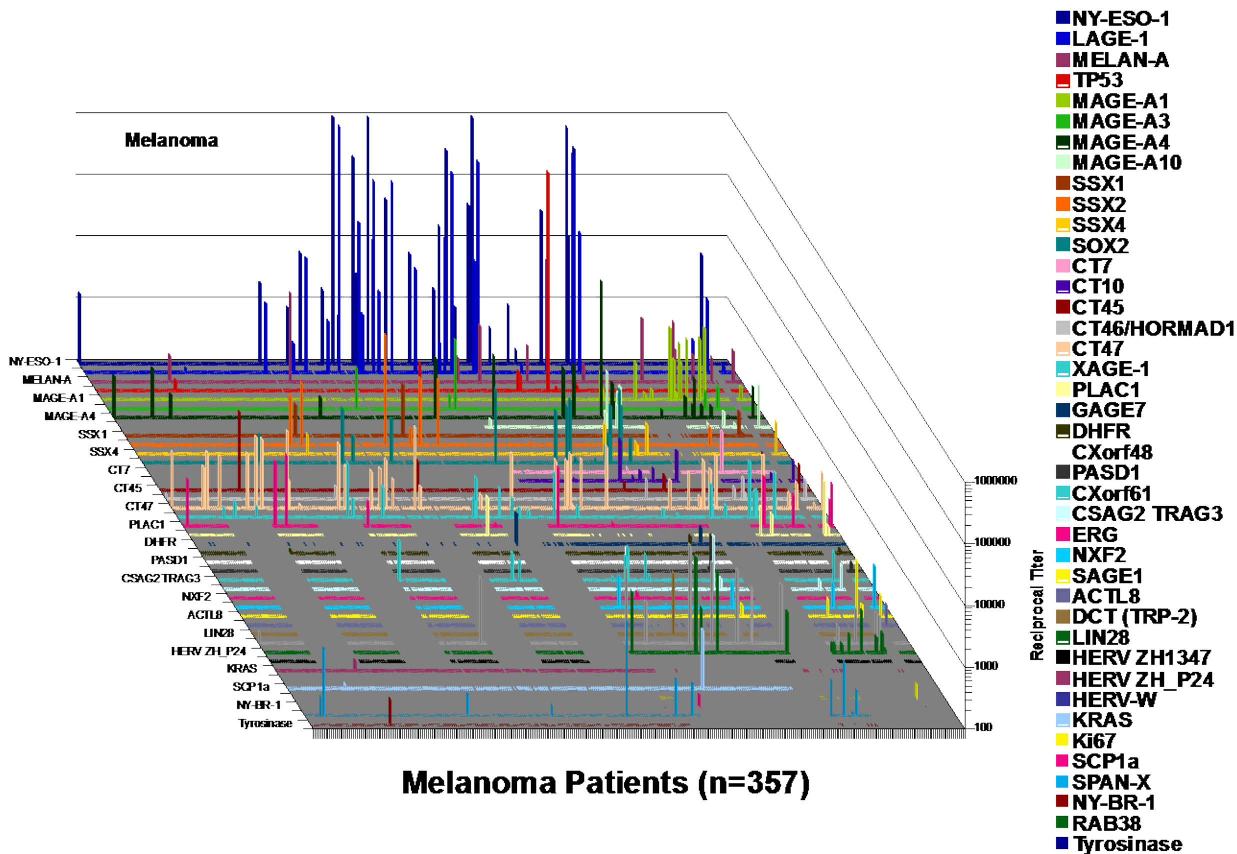
Analysis of humoral immune response in peripheral serum or plasma. One approach to test whether tumors become more immunogenic as a result of immunotherapy is to measure naturally occurring antibodies in patient serum or plasma. Autologous serotyping has resulted in a growing list of immunogenic human cancer antigens[1], including mutational, overexpressed, oncogenic viral, differentiation, and cancer-testis antigens. These antigens, though usually intracellular, reflect the immune system's capacity to detect abnormalities associated with neoplasia. The presence of antibodies to tumor antigens can be used as a marker of tumor presence or progression, but may also help generate T cell responses via immune complexes with cognate antigen. In some cases, such as the cancer/testis antigen NY-ESO-1, serum antibodies are associated with spontaneous T cell responses in peripheral blood. We propose to test the hypotheses that immunotherapy leads to induction or increase in immunity to locally expressed antigens by assessing serological changes pre-post treatment*, and that the serological repertoire in patients may be useful as a prognostic or predictive tool.*

Two levels of repertoire analysis can be offered, both based on classic ELISA:

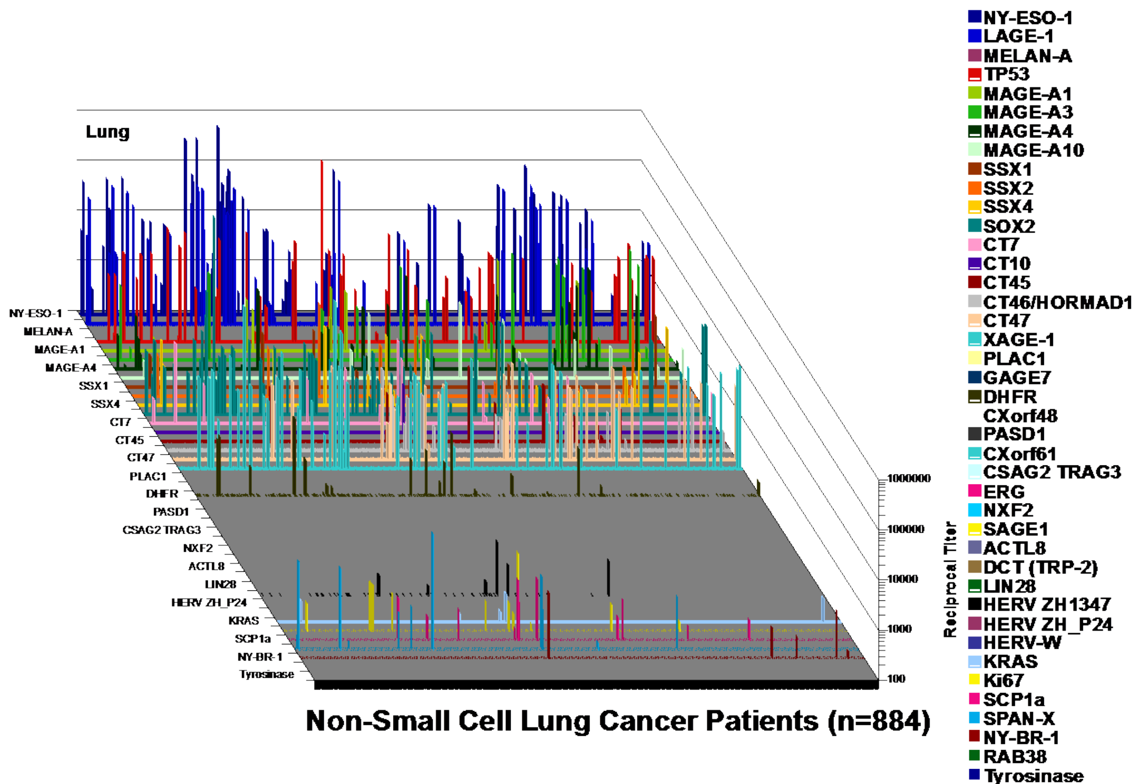
- **(i)** **Grand Serology**. We have shown that presence of baseline serum antibodies to NY-ESO-1 was correlated with clinical benefit to ipilimumab in metastatic melanoma patients[2], and that changes in NY-ESO-1 antibody titers also correlated with clinical events in a patient with abscopal response[3]. More generally, using serological markers as a surrogate for presence of antitumor immunity is proposed as a quick and affordable way to assess tumor immunocompetence and response. A series of known tumor antigens will be assessed in a hypothesis-driven manner for their capacity to elicit autoantibodies in treated patients. Using ELISA as previously described[4], we routinely test a series of 25 tumor antigens, including mutational, stem-cell, and cancer-testis antigens such as TP53, NY-ESO-1, SOX2, PRAME, WT1, MAGE-A3, SSX2, etc., most of which already have demonstrated immunogenicity in various solid and hematologic tumors. Grand Serology is also ideally suited to test for potential for antigen spreading, i.e., development of seroreactivity to antigens unrelated to immunogens, which is a useful measurement to assess in immunotherapy. *We propose to use this assay systemically as a unifying measurement throughout all NCI-supported immunotherapy trials, to define baseline immunogenicity, quantify changes or induction in peripheral immunity with local or systemic treatments, and probe potential prognostic value.*
- **(ii)** **Standard ELISA**. This test is ideal once a smaller number of known antigens has been defined from Grand Serology, or if there is a preexisting hypothesis such as response to a defined vaccine target. ELISA offers the flexibility to address immunoglobulin isotype (IgG, IgA, IgE…) or subclass (IgG1, IgG2, IgG3, IgG4) usage, to assign functionality to antibody responses, to measure immunoglobulin (Ig) titers to defined sequences, and map epitopes.

Below are examples of frequency of detection of IgG to NY-ESO-1 by ELISA in patients with various solid tumors, as well as frequency to other antigens tested in Grand Serology in several different cancers.

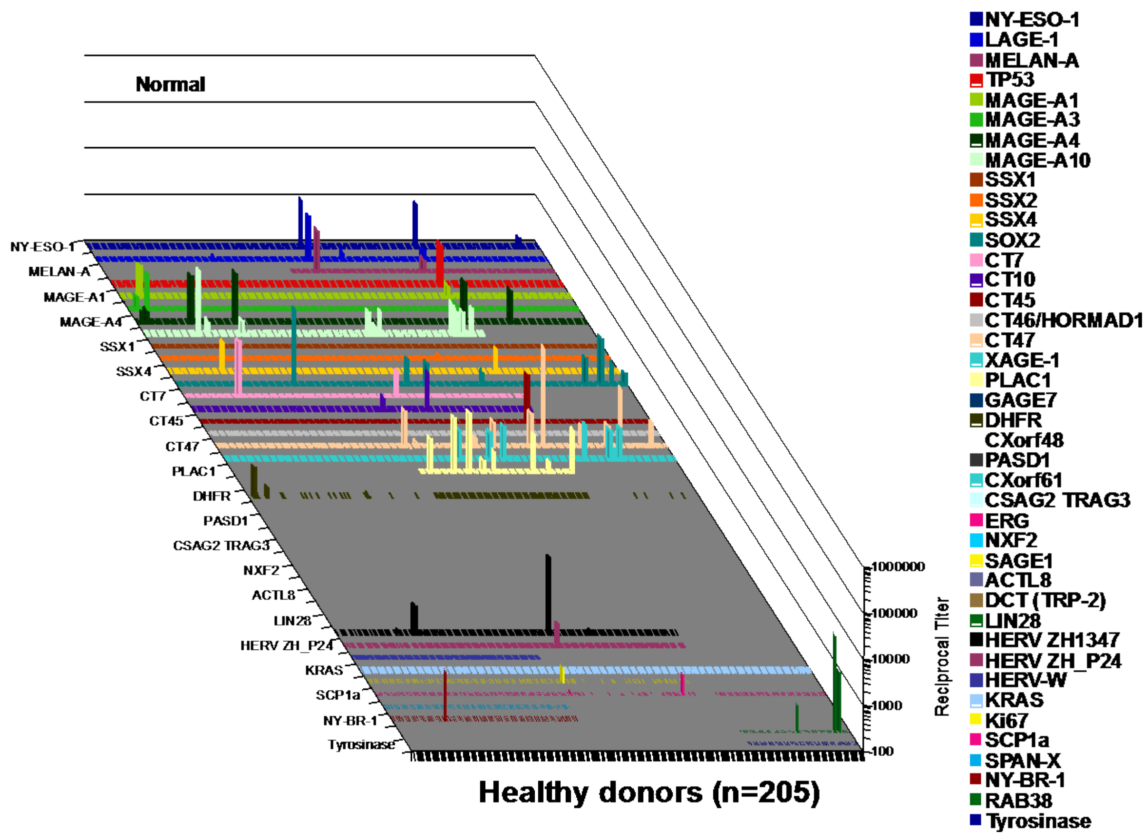
Ab titers by Grand Serology against multiple antigens in melanoma patient sera



Ab titers by Grand Serology against multiple antigens in non-small cell lung cancer patient sera

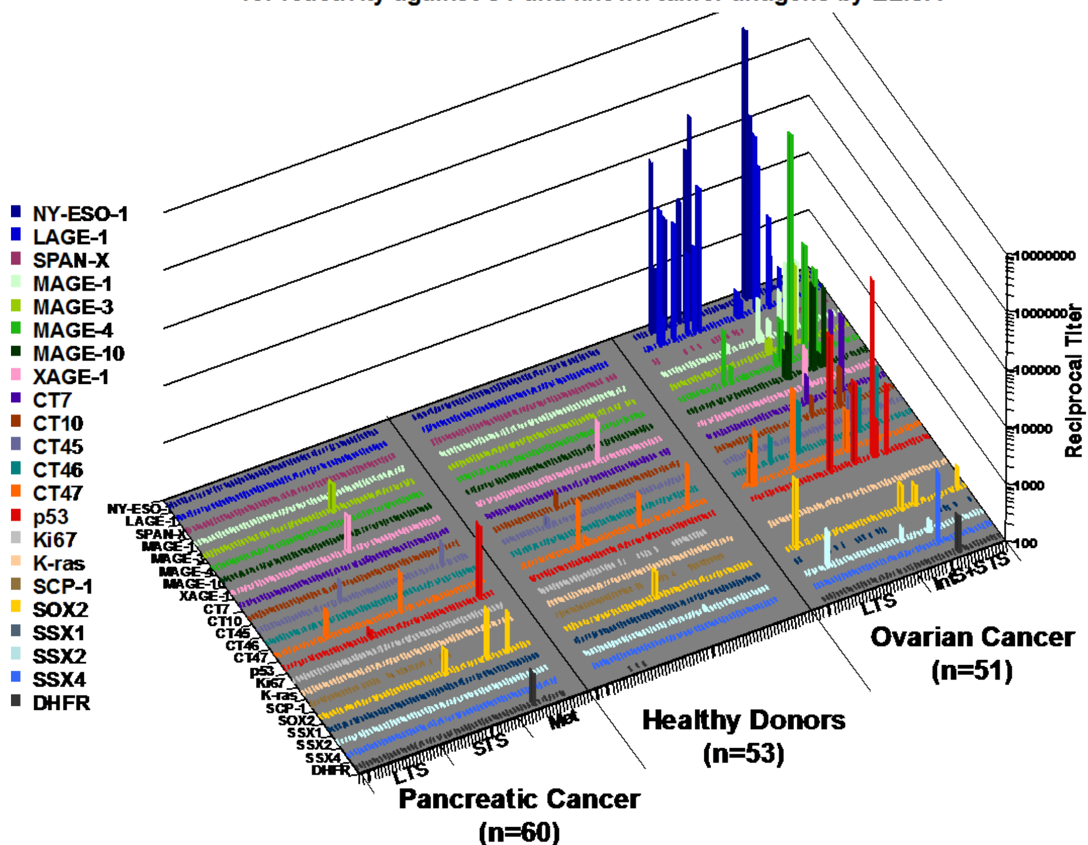


Ab titers by Grand Serology against multiple antigens in healthy donor sera

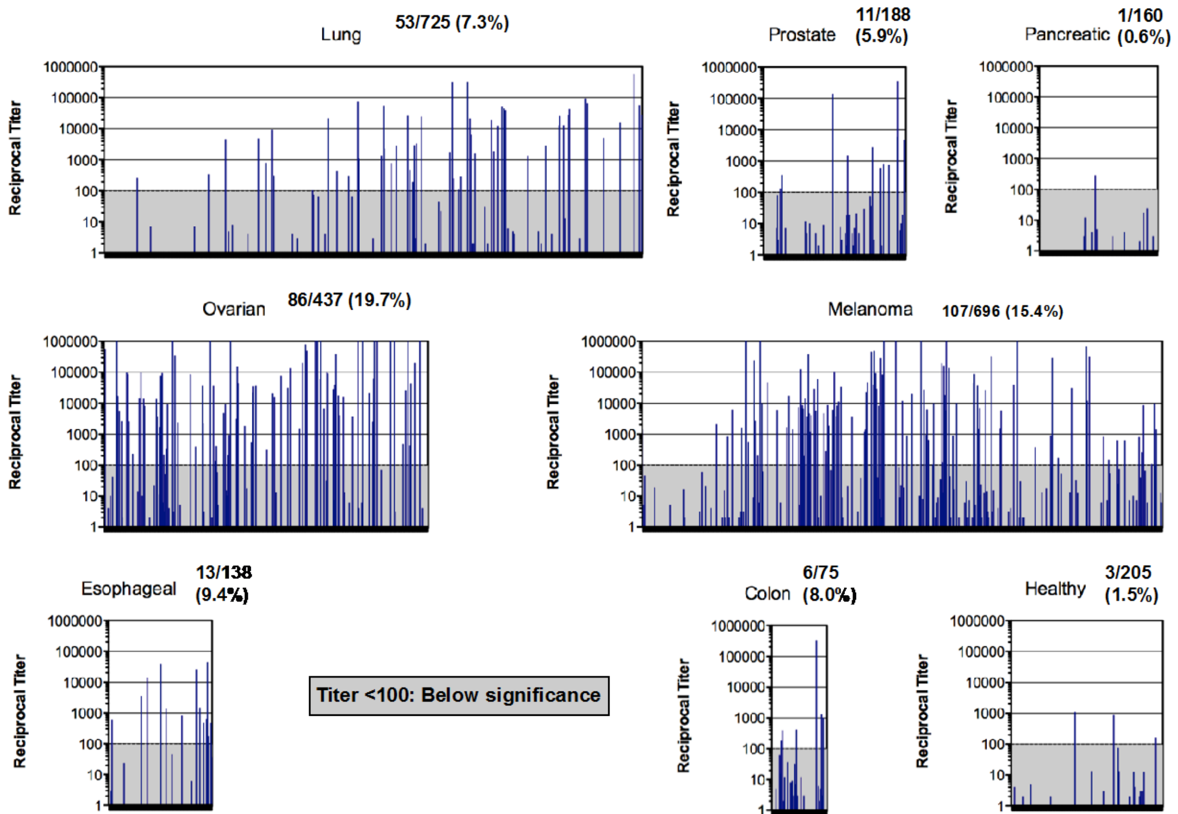


Grand Serology:

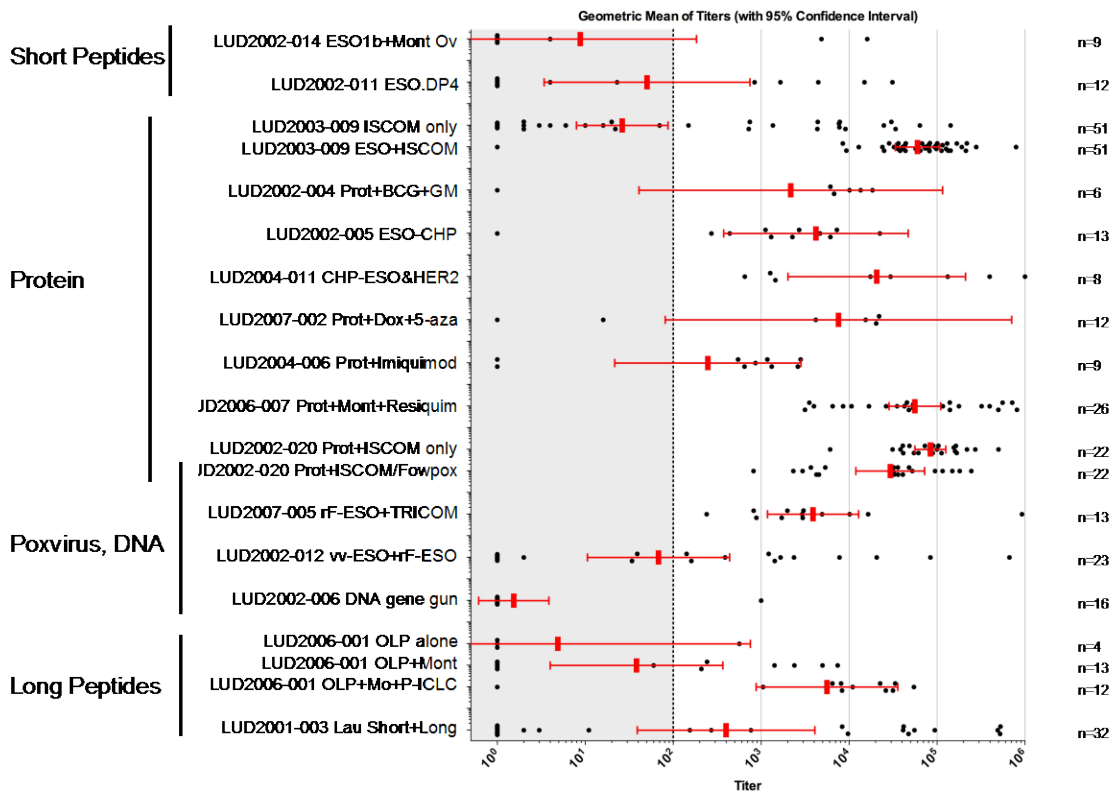
Comparing pancreatic cancer sera / healthy donor sera / ovarian cancer sera for reactivity against CT and known tumor antigens by ELISA



Frequency of NY-ESO-1 serum antibody responses in various cancers by ELISA



Comparison of NY-ESO-1 titers across Cancer Vaccine Collaborative clinical trials (geometric mean of titers after treatment) - Tested centrally by MS-CIMAC investigators



2. Analytical validation

- **Analyte(s)**

A collection of recombinant proteins representing known full-length tumor antigens such as cancer-testis antigens NY-ESO-1, MAGEs, SSXs, and mutational antigens TP53 or stem cell antigens SOX2 and WT1 are available for testing presence of antibodies in serum, plasma or other sources. These proteins are either purchased or made in collaboration with Dr. Nishikawa (National Cancer Center, Tokyo). Most of the antigens are made as recombinant proteins from E. coli, and harbor either a his-tag or a GST-tag or both. Some are made in yeast, pichia, or mammalian cells, in particular if post-translational modifications are thought to be important. The proteins may form secondary structures based on natural folding, though many of them are in urea inducing partial denaturation of tertiary structures and better accessibility to a greater variety of potential epitopes. Finally, long synthetic peptides are also routinely used for some antigens to check for either mapping of individual epitopes or to confirm recognition of linear sequences when used in pools.

Though the composition of the panel may change over time due to protein availability, new lots/batches of antigens are systematically tested in comparison with old ones to ensure accurate titration of antigen and reproducible range of expected data with known controls.

Table 1. List of some of the current antigens available for testing in grand serology. Three different sources of NY-ESO-1 exist, along with not shown peptide pools for NY-ESO-1, p53, and MAGE-A3.

NY-ESO-1 (Nishikawa)	NY-ESO-1 (Immune Design)	NY-ESO-1 (LICR)	p53	MAGE-A1	MAGE-A3	MAGE-A4	MAGE-A10
SOX2	SSX4	MAGEC2/CT10	CT47	MELAN-A	CT46/HORMAD1	SURVIVIN	SURVIVIN-2B
DHFR	HERV-K	UBTD2	XAGE	WT1	PRAME	ERG	GAGE7

It is very easy to customize assays to include any other antigen, including neoantigen peptides for example, based on robustness of assay and if positive controls may be available such as monoclonal antibodies.

- **Technical platform(s) and scalability**

The ELISA / Grand Serology assay is performed in a semiautomated fashion thanks to a Biotek ELx405 Microplate Washer fitted with a Biostack Multiplate Stacker for washing steps, and a Biotek Synergy L Microplate Fluorescence Reader also fitted with Biostack Multiplate Stacker for fluorescent readout of ELISA. These instruments are serviced and recalibrated once to twice yearly, along with automated pipetters. Assays are run on half-volume 96 well plates, thereby ensuring minimal use of biospecimens – typically less than 10µl required for a complete run of 25 antigens. Semi-automation ensures capacity is increased to allow the handling of samples. Scalability of assays is dependent on numbers of targets used for Grand Serology/ELISA. Several hundreds of sera can be assessed in less than a month to one antigens, while throughput for 25 antigens is at least 100 sera per month.

- **Reagents, controls, and calibrators**

Secondary reagents to detect immunoglobulins and enzymatic fluorescent dyes are controlled within each experiment based on expected range of reactivity to known antigens such as NY-ESO-1. New lots are adjusted to ensure similar range of reactivity.

Ability of antigens to coat plates can be facilitated by the use of specific monoclonal antibodies expected to react with assays.

Positive control sera are used in each experiment and on every plate, and they are typically pools of previously tested sera known to react with most of these proteins. One simple control is using sera from patients who received a his-tag protein vaccine and developed antibodies to the tag, which then react with most tagged proteins in the ELISA. Preferred controls however are antigen-specific and used in combination or pooled, so that a large number of antigens can be controlled at the same time. A challenge is obtaining sufficient amounts of sera from such positive controls, knowing these are usually cancer patient sera with limited amounts. Therefore, positive controls need to assessed on a regular basis and compared to previous batches, in a customized fashion according to panels tested.

Negative control sera are used in duplicate <u>in each experiment on each plate</u>, because they are the basis for extrapolating titers. See more details below in assay validation, under calibrators.

- **Pre-analytical variables.**

Antibodies in plasma or serum typically have a long half-life and are fairly resistant to storage conditions and tolerant of tube collection. As part of trials under CIMAC consideration, an umbrella collection protocol dictating collection of plasma from sodium heparin-based tubes, along with processing and freezing recommendations, should minimize variability due to specimen acquisition. Samples should be processed within 24 hours of collection, to avoid potential for hemolysis and other interfering substances. Nevertheless, in the 20 years of running the ELISA assay, samples were obtained from a wide variety of sources and maintained frozen, thawed over 30 times, or just kept at 4°C for a prolonged time, and results still appeared generally consistent for a given serum with high titers. To be clear, it is recommended to minimize freeze-thaw cycles and maintain and ship plasma or sera at -20°C or -70°C whenever possible. Realiquoting specimens after thawing should be done in small amounts, to prevent too much manipulation. The recommended type of tubes to be used are sodium heparin for plasma and red-top for serum.

- **Specimens & methods for specimen acquisition, fixation, stabilization, and processing.**

The ELISA / Grand Serology platform uses only 3-15µl per sample (depending on number of antigens tested) of a variety of soluble biospecimen sources: peripheral blood plasma or serum, bone marrow plasma or serum, ascites supernatant, spinocerebral fluid, cell culture supernatant, urine, etc. The use of biospecimens however may have different characteristics based on processing (concentration, addition of cell culture medium, etc).

Nevertheless, we have tested matched serum and plasma (Figure 1), as well as bone marrow samples (Figure 2) to a variety of antigens and have not found significant differences in antibody titers detected, indicating that at least <u>serum and plasma from blood and bone marrow can be used interchangeably</u> (assuming collection and processing was done in a similar fashion). This is likely because antibody circulates at similar concentrations throughout the body and other factors from plasma or serum do not appear to influence ability to detect immunoglobulins. We recommend however only using one specific collection and processing method for a study, to avoid any unforeseen confounder.

Sera or plasma should be spun prior to each test to remove potential leftover non-soluble contaminants such as platelets. Hemolyzed sera should be noted, as they may in some extreme cases lead to non-specific low-level reactivity to a majority of antigens.

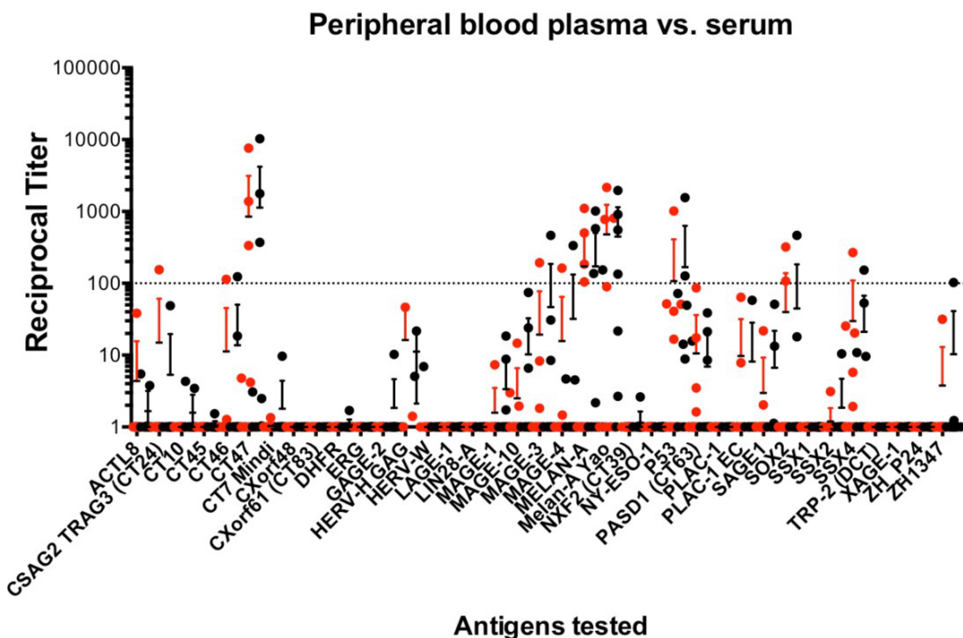


Figure 1. Comparison of peripheral blood plasma (red) vs. peripheral blood serum (black) in 11 paired sera from 7 breast cancer patients tested in grand serology to a series of 44 known tumor antigens. There was no significant difference observed in titers between plasma and serum, even for reactivities falling under the limit of significant detection indicated by the dotted line (<100).

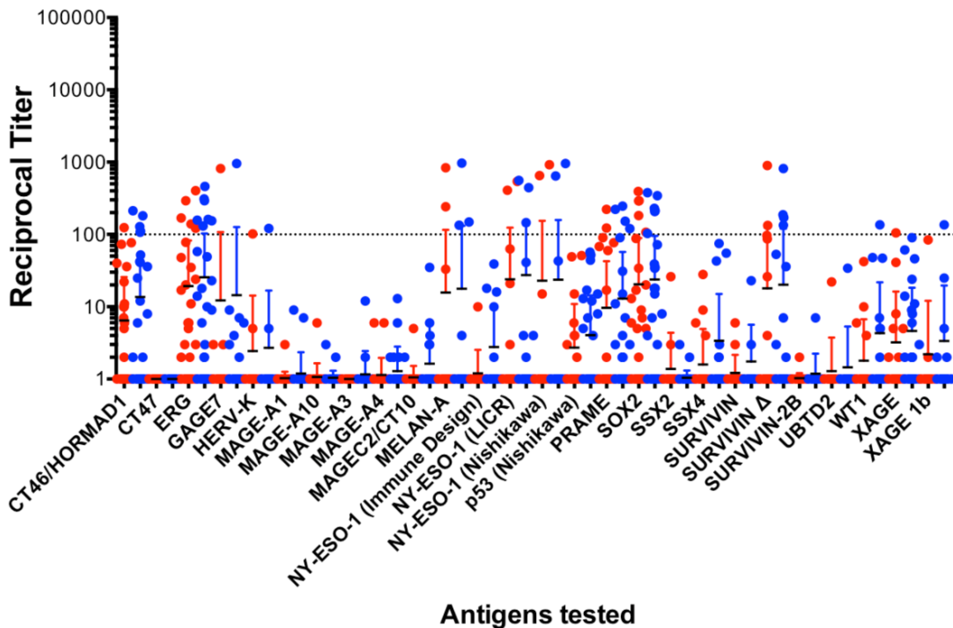


Figure 2. Comparison of bone marrow (blue) vs. peripheral blood plasma (red) in 76 paired samples from 65 smoldering multiple myeloma patients. There was no significant difference observed in antibody titers between bone marrow and peripheral plasma, even for reactivities falling under the limit of significant detection indicated by the dotted line (<100).

- **Interfering substances**

Some plasma and serum samples exhibit various degrees of hemolysis, as observed visually by color, in particular when collection was not properly handled or if patients had strong cytotoxic in vivo treatment regimens. A note should be made next to ELISA results from samples with hemolysis to potentially flag them in downstream analysis and prevent errors in interpretation.

Samples also sometimes contain remnant platelets that could potentially interfere with ELISA. The recommendation for the latter is to always spin at high speed specimens just prior to sampling and realiquote them if necessary (in case of visible pellet).

Less than 5% of all tested sera exhibit what is considered “stickiness”, i.e., reactivity to more than 50% of antigens tested in the panel. This occurs independently of hemolysis and is discussed below in the section about assay validation. While the source of stickiness is not known and could be due to lipids, bilirubin, or hemoglobin, attempting to remove them by purifying IgG or by using a sandwich ELISA approach would modify our assay too much and make interpretation across samples impossible. We therefore propose to flag samples that show stickiness to ensure they are not overinterpreted.

Secondary antibodies used in ELISA are cross-adsorbed for other species and should minimize potential HAMA effect of patients exposed to non-humanized reagents.

As discussed below, attention should be paid to the type of treatment received to adjust potential expectations of results. If the patients received a recombinant protein vaccine with a tag, using the same tagged-proteins in ELISA should be avoided, or at least considered as a potential source of tag-specific Ab reactivity. Also, treatment with intense chemotherapy can in some cases exacerbate “stickiness”.

Lastly, plasma/serum concentration of proteins, in particular IgG, could be a significant factor in particular for multiple myeloma where paraprotein could skew representativeness of the repertoire. While one could consider adjusting starting dilutions based on IgG concentration, we prefer recommending post-hoc adjustments in titer calculations if needed, to prevent changing analytical variables themselves.

- **Assay validation data**, including:

a. Current status of studies defining the **sensitivity, specificity, accuracy, precision, reportable range, reference ranges/intervals** (normal values), **turn-around time and failure rate** of the assay as it is to be performed in the trial.

The reportable range depends on the number of serum/plasma dilutions used in titrations. Typically, assays are run in 4-fold dilutions from 1/100 to 1/6400, allowing to extrapolate reciprocal titers with precision

from 25 to 25,000. If titers fall in the higher range of 25,000 to 1,000,000 or higher, 6 or more dilutions are recommended to extrapolate precise titers (they would otherwise possibly be overestimated).

Specificity of reactivity to a given antigen is determined by comparison with known standards, comparison with other formulations of the given antigen, as well as comparison with unrelated antigens. Accurate detection of antigens and epitopes is dependent on antigen formulation (peptides will only be useful for linear sequences while protein will also allow for some conformational epitopes).

Turn-around time for testing 20-40 sera to ~25 antigens is typically less than a week, with most assays performed over a 2-3 day period. Flexibility is allowed at several steps (blocking, coating) between overnight or shorter incubations without affecting results, in case deviation from SOP is required due to external factors.

Failure rate for ELISA is <1% and is mostly due to technical errors (omission of steps). Failure to interpret data due to biological reasons occurs infrequently (<5% of samples) and are generally related to quality of biospecimen or prior treatments in patients such as extensive chemotherapy. These can sometimes lead to "stickiness", which can be recognized because a majority of antigens show low-level reactivity. While these are then flagged to prevent false positive interpretation, true reactivity to specific antigens can still be determined if titers are >1000 and at least 1 log higher than the stickiness reactivity.

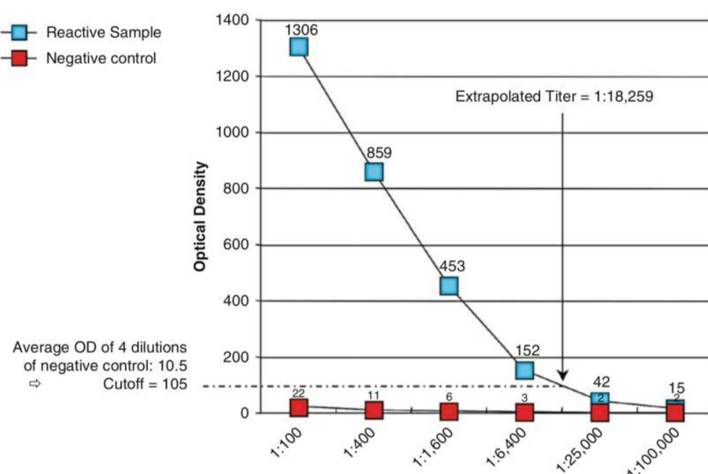
False positives can be detected in patients who previously received vaccination with recombinant tagged proteins and developed antibodies to these tags. This has been observed in patients receiving NY-ESO-1-his tagged protein immunizations, who eventually may end up reacting to a majority of other recombinant proteins sharing the same his tag. To ascertain specificity in these cases, use of non-his tagged antigens (peptide pools, untagged or gst-tagged proteins) is warranted unless titers are observed at >1 log higher levels than to the cognate vaccine target.

b. Use of positive and negative controls, calibrators, and reference standards.

Each plate includes a negative control pool of healthy sera pre-tested for reactivity to antigens used and available in sufficient amounts to be able to use them over a period of several years. This negative control serum pool is critical because it is used to extrapolate titers (as shown in Fig. 3) for all other samples. It is therefore used in duplicate on each plate, in case of artifact in one of them preventing analysis.

For ELISA, titers for a specific serum are extrapolated based on healthy donor serum pools as negative controls present on each plate for each antigen. It is recommended to batch all longitudinal samples from a patient and assay them simultaneously for titer comparisons, even though inter-assay reproducibility is >90%. This ELISA has been used in over 1,800 cancer patients in a variety of clinical trials, representing nearly 10,000 individual time points, with extremely high reproducibility and dynamic range. Significance is defined by titers >100, in comparison to irrelevant control antigen (such as DHFR), and changes across time points are considered specific if 4x different from baseline. For antigens showing humoral reactivity, mapping of linear epitopes may be assessed using overlapping peptide series covering the sequence of the antigen, to assess polyclonality and potential spreading over time. Batch effect and experimental variations are minimized by testing all longitudinal cryopreserved sera from a same patient simultaneously.

The most significant calibrator for our assays is the negative control healthy donor serum pool. This pool is the basis for calculating all titers, including those of positive controls. We have collected over 1 liter of this pool to ensure long-term usage and minimize variability over years of assays. This could also be used by other centers for harmonization. Positive control sera are pooled in some cases, as testing individual sera would use up precious specimens too quickly. As long as specificity is ascertained by comparing each protein and that the pool is within expected ranges, it is acceptable to change positive control pools on a regular basis and define parameters where OD are expected to fall within.



For grand serology, sera will be tested in 4x serial dilutions, starting from 1/100, for IgG reactivity against full-length E. coli-produced proteins of tumor and control antigens to assess specificity. Positive and negative control sera with known reactivity will be used on each plate to validate the assays. Reciprocal titers will be extrapolated and considered significant if >100. In samples with multiple serum collection time points available, significant changes in humoral response will be defined as antigen-

specific antibody going from undetectable (titer <100) to detectable (titer >100), or with titers at least 4-fold different between time points.

We do not generally observe any hook effect, likely due to the 1/100 starting dilution. In cases of extremely high-titered sera, we see a plateau at the first several dilutions, which would require adjusting the number of dilutions for a more precise estimate of titers. For the extrapolation of titers to be optimal, calculations should be made within the linear downslope range of the OD dilution curve. In some cases where titers are outside of the dilution range, they should be reported as over the limit of the extreme dilutions used.

Figure 3. Example of method used to extrapolate titer, based on the average of OD values from a negative control serum pool. The control serum pool is derived from 3 healthy donors tested for reactivity to a wide range of self-antigens, and each serum is available in >200ml volume to ensure that the same reference pool is used throughout multiple study over the course of several years.

c. How run-to-run variation (Coefficient of Variation; CV) was assessed and handled.

In the examples below (Figures 4-6), titers were calculated across a period of more than 1 year for NY-ESO-1 reactive positive control serum pools, changing over time but titrated to be in a similar detection range. For each test, up to 8 replicates were performed on the same day for intra-assay CV, while inter-assay CV was calculated based on reactivity of the same control over 5-10 assays. For Figures 4-5, CV values based on NY-ESO-1 protein using a high-titered control serum pool were 1.12% and 1.19% respectively for intra- and inter-assay variability. Figure 4 and 5 show the same data, on a log-scale (recommended), or on a linear scale to emphasize differences observed. Figure 6 shows a representative low titered serum tested against NY-ESO-1 long peptides, and as a result, showing also greater coefficient of variation both within the same assay and across assays (8% and 14% respectively on average). Levey-Jennings plots in Figures 5-6 indicate range of acceptable titers for controls, with only very few samples falling outside of the area in low titered sera (results associated with these time points can be flagged instead of outright rejected). Similar results were obtained with other antigens tested, although CV values are dependent on average titers observed. For Grand Serology, most antigens are expected to have a CV <5%, except if reactivity falls around the limit of significance and detection (titers of 1/100) in which case CV can be higher than 10%.

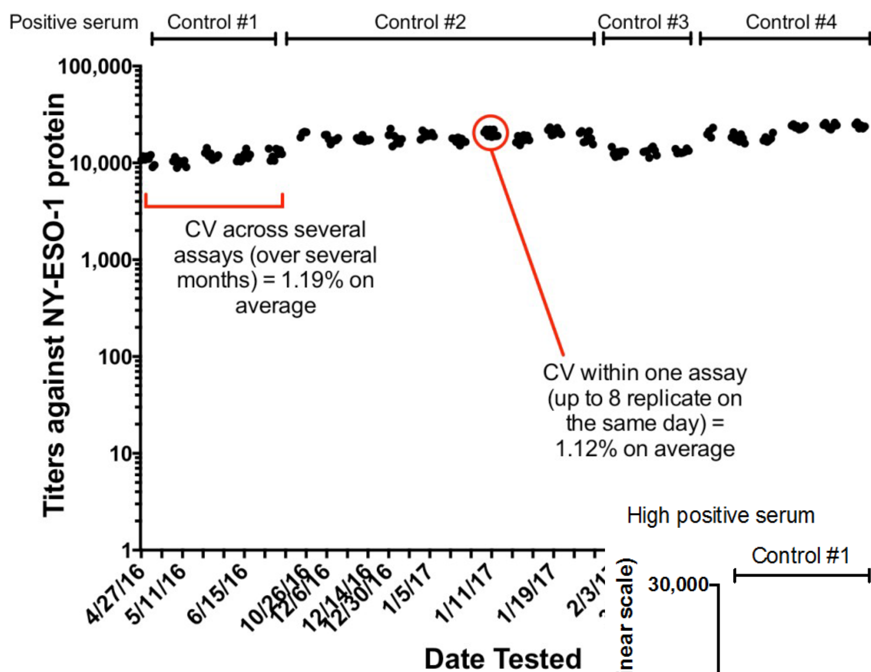


*Figure 4. **ELISA / Grand Serology analytical performance,** using four serum pools known to react against NY-ESO-1 protein, diluted to hit a predicted titer of 10,000-30,000. Coefficients of variance within assays and across assays are indicated.*

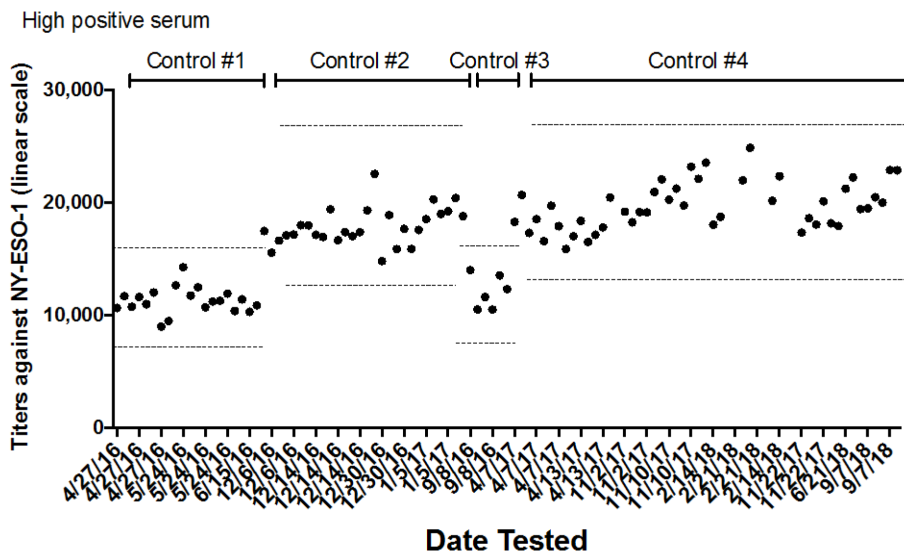


Figure 5. Same data as in figure 4, shown on a linear scale with acceptable titer range within each control used.

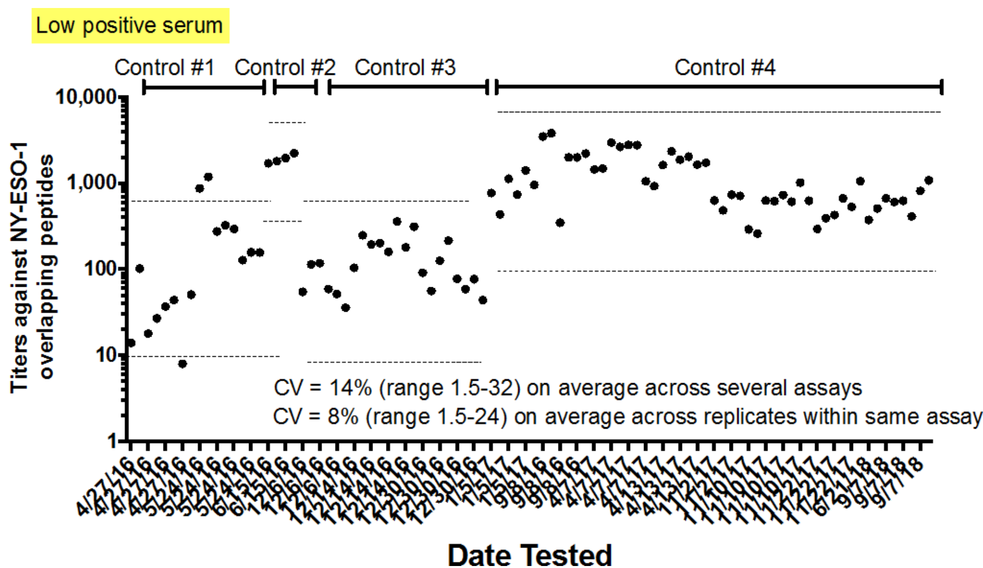


Figure 6. Calculating CVs across assays and within assay for a low positive serum pool against NY-ESO-1 overlapping long peptides.

d. **How inter-laboratory variability in the measurements was assessed** and how these sources of variation were minimized to maintain performance at all sites within acceptable limits and to prevent drift or bias in the assay.

We established many rules for immune monitoring of cancer vaccines through our collective experience with NCI and industry-sponsored trials and with the Cancer Vaccine Collaborative. To allow rapid comparisons of these multiple parallel Phase I/II trials, we applied harmonized or centralized immune monitoring to vaccines to quantify antibody, as well as CD4 and CD8 T cell induction, and to assess "integrated immunity", i.e., all three components of adaptive immunity. Because of high sensitivity, adaptability, and robustness, ELISA was chosen as the primary readouts for Ab titer determination. Implementation of serological immune monitoring has been successfully implemented in dozens of published studies, some of which are listed here[5-20]. This landmark work has helped us develop and validate the most robust strategy to assess antigen specific immunity against a variety of tumor antigens.

Determination of antibody titers induced by immunotherapy to drug-related antigens will be performed by ELISA as described[4], using a semi-automated platform. The assay offers flexibility in choice of input source (serum, plasma, urine…), antigens tested, formulations (proteins, peptides), subclass and isotype of immunoglobulins tested, and range of titration, while remaining fast and affordable. Assays are validated by the use of previously tested sera with positive reactivity to each antigen tested, included on each plate.

e. **Describe proficiency testing and results**.

At this time, no other CIMAC offers the range of proteins tested in Grand Serology, and therefore efforts to establish proficiency panels are still to be decided. It is possible that Grand Serology and ELISA assays will be centralized at MS-CIMAC in the beginning, thereby reducing the need for cross-validation between sites.

Nevertheless, we have previously established a proficiency test with the company Serametrix, by comparing titers with MS-CIMAC before they would offer NY-ESO-1 serological testing as a CRO. Sera were sent to Serametrix for testing while Serametrix sent a mammalian source of NY-ESO-1 used in their ELISA. Both sites performed titer calculations on the shared sera as well as on proteins used at each site. Sera were chosen to have a range of representative titers, from very high (>100,000), to high (~10,000), to intermediate (~1,000), to low (~100) to negative. Assays were highly concordant for both sources of proteins, and titer interpretations were similar across both sites, even though sensitivity was a little higher at MS-CIMAC for low titer sera. Concordance was assessed by comparing titer values, which were consistent (within 2x difference) between sites for high and medium titers, but differed for the low titered serum (CIMAC perfoming with higher sensitivity compared to Serametrix originally). By aligning protocols, we are currently in the process of ensuring that the dynamic range of each respective assay is similar to ensure that even low titers will be detected in a similar manner.

Besides Serametrix, it should be noted that our SOP was successfully established and transferred between Mt Sinai CIMAC, Frankfurt Nordwest Hospital, and Roswell Park Cancer Institute. However, harmonization efforts will eventually be needed between CIMACs once they start implementing ELISA.

f. **Scoring procedures and type of data to be acquired**:

- quantitative/continuously distributed
- semi-quantitative/ordered categorical
- qualitative/non-ordered categorical

Reciprocal titers are a continuous variable calculated as indicated above from dilution curves and plate-specific negative controls. There are however categorical cutoffs applied, where titers are considered significant if ≥100, which is the dilution at which tests are started for titrations. In some rare cases where titers are expected to be consistently low, a lower starting dilution may be considered allowing for potential significant reactivity in the range of 1/10-1/100.

Because titers are tested in 4x dilutions and taking CV into account, significant changes between Ab titers pre- and post- treatment are defined as being significant if greater than 4x different, or going from negative (<25) to positive (≥100). More than 99% of negative titers are expected to be at a value of 1, i.e., predicted undetectable in undiluted serum. In cases where low-level reactivity is seen but still below significance (titers 25-99), changes in titers still need to be at least 4x different and reach above 100 to be considered significantly increased.

For statistical analyses, both continuous titers or categorical titer such as present/absent (based on the 1/100 cutoff) can be useful depending on hypotheses tested. For example, vaccine studies to look for antigen-specific antibody induction should measure changes as a continuous variable, to ascertain that treatment had a quantitative effect. In studies of antigen spreading, a yes/no for presence of antigen response or presence of 4x change may be sufficient.

g. **Criteria and metrics for defining significant changes** (e.g., between timepoints, between responders and non- responders).

For ELISA/Grand Serology, titers for a specific serum are extrapolated based on healthy donor serum pools as negative controls present on each plate for each antigen. It is recommended to batch all longitudinal samples from a patient and assay them simultaneously for titer comparisons, even though interassay reproducibility is >90%. This ELISA has been used in over 1,800 cancer patients in a variety of clinical trials, representing nearly 10,000 individual time points, with extremely high reproducibility and dynamic range. Significance is defined by titers >100, in comparison to irrelevant control antigen (such as DHFR), and changes across time points are considered specific if 4x different from baseline. For antigens showing humoral reactivity, mapping of linear epitopes may be assessed using overlapping peptide series covering the

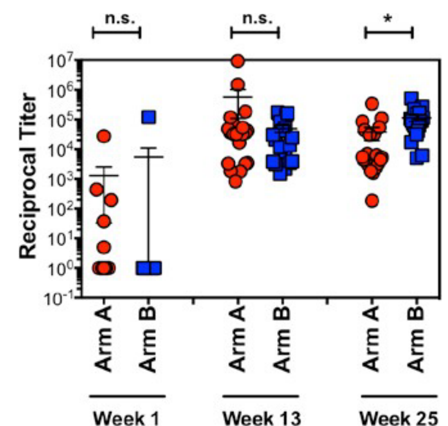


Figure 7. Titer differences for NY-ESO-1 at different time points in study LUD2002-002 comparing NY-ESO-1 protein vaccination in ISCOMATRIX for 24 weeks (Arm B) vs. the same for 12 weeks followed by 12 weeks of recombinant fowlpox NY-ESO-1 (Arm A). Titers were significantly lower in Arm A at week 25 (Mann Whitney test, p=0.0003).

sequence of the antigen, to assess polyclonality and potential spreading over time. Batch effect and experimental variations are minimized by testing all longitudinal cryopreserved sera from a same patient simultaneously.

For grand serology, sera will be tested in 4x serial dilutions, starting from 1/100, for IgG reactivity against full-length E. coli-produced proteins of tumor and control antigens to assess specificity. Positive and negative control sera with known reactivity will be used on each plate to validate the assays. Reciprocal titers will be extrapolated and considered significant if >100. In samples with multiple serum collection time points available, significant changes in humoral response will be defined as antigen-specific antibody going from undetectable (titer <100) to detectable (titer >100), or with titers at least 4-fold different between time points. The 4-fold difference was determined based on actual titration range in the assay and was chosen to exceed expected variability in titers from CV studies above (where up to a 2x difference can occur in repeat assay). Longitudinally collected specimens within an individual with seroreactivity typically show very stable titers across several months (unless an intervention such a vaccine or a treatment such as chemotherapy occurred).

As described above, longitudinal changes between timepoints are considered significant if greater than 4x or changing from negative to positive (or inversely). Antibody responses are long-lived and are not expected to change rapidly within a few weeks in the absence of intervention. To compare sera between cohorts, Mann-Whitney t tests are appropriate. An example is provided in Figure 7 comparing titers across several cohorts.

3. References

1 See //ludwig-sun5.unil.ch/CancerImmunomeDB/.

2 Yuan, J. *et al.* Integrated NY-ESO-1 antibody and CD8+ T-cell responses correlate with clinical benefit in advanced melanoma patients treated with ipilimumab. *Proc Natl Acad Sci U S A* **108**, 16723-16728, doi:10.1073/pnas.1110814108 (2011).

3 Postow, M. A. *et al.* Immunologic correlates of the abscopal effect in a patient with melanoma. *N Engl J Med* **366**, 925-931, doi:10.1056/NEJMoa1112824 (2012).

4 Gnjatic, S., Old, L. J. & Chen, Y. T. Autoantibodies against cancer antigens. *Methods Mol Biol* **520**, 11-19, doi:10.1007/978-1-60327-811-9_2 (2009).

5 Parikh, F. *et al.* Chemoradiotherapy-induced upregulation of PD-1 antagonizes immunity to HPV-related oropharyngeal cancer. *Cancer Res.* **74**, 7205-7216, doi:10.1158/0008-5472.CAN-14-1913 (2014).

6 Sabbatini, P. *et al.* Phase I trial of overlapping long peptides from a tumor self-antigen and poly-ICLC shows rapid induction of integrated immune response in ovarian cancer patients. *Clin. Cancer Res.* **18**, 6497-6508, doi:10.1158/1078-0432.CCR-12-2189 (2012).

7 Odunsi, K. *et al.* Efficacy of vaccination with recombinant vaccinia and fowlpox vectors expressing NY-ESO-1 antigen in ovarian cancer and melanoma patients. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 5797-5802, doi:10.1073/pnas.1117208109 (2012).

8 Tsuji, T. *et al.* Split T cell tolerance against a self/tumor antigen: spontaneous CD4+ but not CD8+ T cell responses against p53 in cancer patients and healthy donors. *PLoS One* **6**, e23651, doi:10.1371/journal.pone.0023651 (2011).

9 Gnjatic, S. *et al.* NY-ESO-1 DNA vaccine induces T-cell responses that are suppressed by regulatory T cells. *Clin. Cancer Res.* **15**, 2130-2139, doi:10.1158/1078-0432.CCR-08-2632 (2009).

10 Odunsi, K. *et al.* Vaccination with an NY-ESO-1 peptide of HLA class I/II specificities induces integrated humoral and T cell responses in ovarian cancer. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 12837-12842, doi:10.1073/pnas.0703342104 (2007).

11 Nishikawa, H. *et al.* Influence of CD4+CD25+ regulatory T cells on low/high-avidity CD4+ T cells following peptide vaccination. *J. Immunol.* **176**, 6340-6346 (2006).

12 Jäger, E. *et al.* Monitoring CD8 T cell responses to NY-ESO-1: correlation of humoral and cellular immune responses. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 4760-4765 (2000).

13 Jäger, E. *et al.* Recombinant vaccinia/fowlpox NY-ESO-1 vaccines induce both humoral and cellular NY-ESO-1-specific immune responses in cancer patients. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 14453-14458, doi:10.1073/pnas.0606512103 (2006).

14 Gnjatic, S. *et al.* Strategy for monitoring T cell responses to NY-ESO-1 in patients with any HLA class I allele. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 10917-10922 (2000).

15 Gnjatic, S. *et al.* Survey of naturally occurring CD4+ T cell responses against NY-ESO-1 in cancer patients: Correlation with antibody responses. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 8862-8867, doi:10.1073/pnas.1133324100 (2003).

16 Dutoit, V. *et al.* Multiepitope CD8(+) T cell response to a NY-ESO-1 peptide vaccine results in imprecise tumor targeting. *J. Clin. Invest.* **110**, 1813-1822, doi:10.1172/JCI16428 (2002).

17 Atanackovic, D. *et al.* Monitoring CD4+ T cell responses against viral and tumor antigens using T cells as novel target APC. *J. Immunol. Methods* **278**, 57-66 (2003).

18 Atanackovic, D. *et al.* Vaccine-induced CD4+ T cell responses to MAGE-3 protein in lung cancer patients. *J. Immunol.* **172**, 3289-3296 (2004).

19 Atanackovic, D. *et al.* Booster vaccination of cancer patients with MAGE-A3 protein reveals long-term immunological memory or tolerance depending on priming. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 1650-1655, doi:10.1073/pnas.0707140104 (2008).

20 Adams, S. *et al.* Immunization of malignant melanoma patients with full-length NY-ESO-1 protein using TLR7 agonist imiquimod as vaccine adjuvant. *J. Immunol.* **181**, 776-784, doi:181/1/776 [pii] (2008).