Biomarkers of tolerance: searching for the hidden phenotype

Esperanza Perucha[1,2,3], Irene Rebollo-Mesa[2], Pervinder Sagoo[1,3,4] and Maria P. Hernandez-Fuentes[1,2,3]

[1]Section of Experimental Immunobiology, Division of Transplantation Immunology and Mucosal Biology, King's College London, London, UK; [2]Medical Research Centre for Transplantation, King's College London, Guy's Hospital, London, UK; [3]NIHR Biomedical Research Centre, Guy's and St Thomas' Hospital and King's College London, London, UK and [4]Section of Immunoregulation and Immune Intervention, Division of Transplantation Immunology and Mucosal Biology, King's College London, London, UK

Induction of transplantation tolerance remains the ideal long-term clinical and logistic solution to the current challenges facing the management of renal allograft recipients. In this review, we describe the recent studies and advances made in identifying biomarkers of renal transplant tolerance, from study inceptions, to the lessons learned and their implications for current and future studies with the same goal. With the age of biomarker discovery entering a new dimension of high-throughput technologies, here we also review the current approaches, developments, and pitfalls faced in the subsequent statistical analysis required to identify valid biomarker candidates.

Kidney International Supplements (2011) **1,** 40–46; doi:10.1038/kisup.2011.11

KEYWORDS: biomarkers; immunosupression; renal transplantation; tolerance; validation

TO CITE THIS ARTICLE:
Perucha E, Rebollo-Mesa I, Sagoo P *et al*. Biomarkers of tolerance: searching for the hidden phenotype. *Kidney Int Sup* 2011; **1**: 40–46.

Correspondence: *Maria P. Hernandez-Fuentes, Medical Research Centre for Transplantation, Guy's Hospital, Great Maze Pond, London SE1 9RT, UK. E-mail: maria.hernandez@kcl.ac.uk*

Solid organ transplantation is the preferred treatment for end-stage kidney failure, as it substantially improves patient survival and quality of life over continued dialysis therapy. However, in order to avoid rejection, graft recipients have to undergo lifelong immunosupression (IS). This poses a difficult challenge for patients and clinicians as a delicate equilibrium needs to be reached. On one hand, the transplant needs to be protected with sufficient IS, whereas on the other hand the patient needs to maintain an immune system that is healthy enough to fight-off infections and cancer. Moreover, immunosupressive agents have metabolic side effects (hypertension, hyperlipidemia, and hyperglycemia) that contribute to the high level of morbidity in patients post-transplantation.

As a consequence, the 10-year overall survival rate of a transplanted kidney is just above 50% in Europe,[1] which puts a yearly increased strain on the number of people waiting for further transplants. Causes of transplant failure are diverse, uncontrolled activity of the inflammatory, and immune systems being one of the main contributors.[2–4]

Ideally, the best clinical situation for a transplant recipient would be the development of donor-specific immunological tolerance. This would allow long-term patient and graft survival without the need for IS. Thus, there is a pressing need to establish the correct level of IS in each patient, such as to individualize their therapy. In this process, the lack of analytical parameters able to indicate the correct treatment for each patient has prompted a wide search for biomarkers of clinical use in transplantation. In the early-phase post-transplantation, the ability to predict acute rejection would allow us to adjust IS accordingly, whereas in the late-phase post-transplantation, our inability to detect whether a patient has developed a degree of tolerance to their graft is absolute. This has led us to use the term 'the hidden phenotype'. Currently, there are no clinically validated tools to dictate whether to increase or decrease the level of IS that each patient needs to be maintained with. In this review, we will discuss the progress made to date and the future avenues in the development of biomarkers of immunological tolerance in kidney transplantation.

THE PAST: THE INDICES OF TOLERANCE STUDY

Despite promising studies reporting the induction of tolerance in experimental models of solid organ transplanta-

tion, few of these approaches have been translated to clinical transplantation. The Indices of Tolerance (IoT) study was initiated as a necessary shift in the paradigm to approaches of inducing transplant tolerance. If we could identify the specific immunological characteristics of tolerance in human subjects, then we could use more targeted and informed approaches to tolerance-induction therapies, with the potential to manipulate the immune compartments responsible for generating donor-specific immune regulation. In 2010, we reported on the findings of the IoT multicenter study, which was conducted over a period of 6 years, and culminated in the description of several immunological characteristics uniquely associated with the tolerant state in renal transplant patients.[5] As an investigator-led study, based within an academic institute, we were faced with the logistic challenges of conducting a clinical study from a primarily bench-based laboratory perspective. Overcoming this was a steep learning curve for the scientists and clinicians involved alike. In fact, the first challenge we faced on initiating the study was the recruitment of tolerant renal transplant subjects. As described later, the risk of renal allograft loss due to IS cessation, based on clinical experience, is thought to be high, and as a consequence, operational tolerance is estimated to be a rare event. Nevertheless, we were able to locate 11 tolerant renal transplant recipients throughout Europe. A number of pathways had led to this outcome, from medical non-compliance to withdrawal due to malignancies. We also encountered some exceptional cases, such as a renal allograft recipient who was unable to gain access to his IS for several weeks, because of severe flooding, which left him stranded on his roof. This latter example seems serendipitous, but may also defy the traditional dogma that the establishment of renal transplant tolerance is a rare occurrence and needs to occur over a period of time. As discussed later, currently, this cannot be tested by any ethically sound approach. By comparing parameters of the immune system in the identified operationally tolerant recipients (stable renal transplant recipients who have ceased all immunosuppressive drugs for more than a year) with control groups (patients with stable renal function maintained on immunosuppressive drugs, patients with biopsy-proven chronic rejection, and healthy individuals), we identified a biomarker signature of tolerance. As indicated later, we adopted both cellular and molecular analytical approaches in our attempt to identify tolerant patients. One of the hypotheses tested was to find an absent functional or active anti-donor immune response. Analysis of cellular immune components revealed elevated numbers of B and NK cells in peripheral blood, diminished numbers of recently activated $CD4^+$ T cells, and donor-specific hyporesponsiveness of $CD4^+$ T cells. Further, we identified a discrete set of genes with altered expression, and a high ratio of FoxP3/α-1.2-mannosidase gene expression in peripheral blood. These findings were then validated on an independent cohort of tolerant renal transplant patients as part of a collaboration with the Immune Tolerance Network (ITN) in the USA, which subsequently led to a parallel publication showing highly comparable findings.[6] The cross-platform approach we adopted toward biomarker identification highlighted a particularly prominent role for B cells within transplantation tolerance, which have very recently re-emerged with newly defined roles within inflammation and immunity.[7–9]

THE PRESENT: THE VALIDATION OF THE BIOMARKERS OF TOLERANCE

Although a set of biomarkers associated with tolerance had been defined at the completion of the IoT project, more questions were then raised: First, would newly identified tolerant patients show the same biomarkers? Second, would the biomarker fingerprint of tolerance also be detectable within groups of renal transplant patients displaying stable allograft function while undertaking immunosuppressive regimens? If so, could the fingerprint be used to predict which patients are more likely to develop tolerance and, as a consequence, would it be possible to successfully minimize their IS under supervision? Finally, could the fingerprint also define patients at the other spectrum of tolerance; that is, patients undergoing chronic rejection? In which case, a set of biomarkers could potentially be used as a monitoring tool for kidney function. To address these questions we initiated our current project: Genetic Analysis and Monitoring of Biomarkers of Immunological Tolerance.[10]

We approached this project with the advantage of having learned many lessons from the IoT. We have put into place a network of several participating centers within the United Kingdom and Europe, together with a continuing firm collaboration with the ITN. We have recruited a technically skilled team that includes not only scientists and clinicians but also statisticians, data managers, and project administrators. We also now have the benefit of new developments in high-throughput techniques, which, because of the NIHR and the UK government investments in translational medical research, have allowed us to establish the infrastructure, facilities, and capacity to perform cutting-edge research. In addition, we are also participants of the International Solid Organ Transplant Registry, a renal transplant registry aiming to establish a multicenter database registering patients who underwent solid organ transplantation, in whom operational tolerance has developed.[11] International Solid Organ Transplant Registry represents a meeting place for patients, families, and physicians, thus facilitating the enrolment of interested patients into appropriate studies of transplant tolerance.

GENETIC ANALYSIS AND MONITORING OF BIOMARKERS OF IMMUNOLOGICAL TOLERANCE

The aim of the Genetic Analysis And Monitoring Of Biomarkers Of Immunological Tolerance study is essentially to assess the biomarkers of tolerance in a larger cohort of kidney transplant patients. This primary aim is approached from several angles.

Tolerant patients

We aim to analyze an independent set of tolerant renal transplant recipients in an observational longitudinal study. Working with samples from newly recruited operationally tolerant patients identified by the collaborating centers around Europe, we are performing the assays described in the IoT fingerprint[5] to assess their probability of being tolerant. This will provide further important information on the validation of the previously identified biomarkers. In addition, because of the acquisition of new equipment and expertise within the research team, a selection of new flow cytometry parameters and a combination of new genes have also been included into the analyses. Finally, we are performing further screening of the biomarkers to be conducted over a period of 3 years in previously identified tolerant patients, testing the important aspect of the stability of the fingerprint. To keep assessing the specificity of the signature, we will also recruit healthy volunteers, patients under different degrees of IS because of a transplant, and patients taking immunsosuppression treatment because of other immunological causes apart from transplantation, such as autoimmune diseases.

Long-term stable kidney function

This part of the study will compare the biomarker signature among patients with stable graft function under conventional IS with patients undergoing chronic rejection. We expect the signature to be present in a small percentage (5–10%, based on the IoT analyses) of patients with stable kidney function. Identifying patients under IS, who may be labeled 'tolerant', is a crucial step toward translating these biomarkers into the clinic.

IS withdrawal

We will assess the frequency of patients showing the tolerance signature in an observational, prospective study of kidney transplant patients who are undergoing steroid withdrawal within 1 year post-transplantation. This protocol will also identify whether patients under IS withdrawal acquire the signature over time, allowing us to test the usefulness of the fingerprint as a biomarker for successful drug weaning.

Basic research

In parallel to these translational projects, we will investigate the role of B cells in tolerance in more detail as a follow-up of the B-cell function enrichment found in the genetic signature. Indeed, since the IoT and ITN studies, the role of B cell in transplantation tolerance has been further confirmed by others.[12] The goal is to examine whether any unique phenotypic or functional characteristics can be identified within the subsets of these expanded cells during operational tolerance. An understanding of the mechanistic basis of the tolerant state would not only allow us to better identify it, but also permit the development of molecular or cellular targets for tolerance induction therapy in renal transplant patients.

OTHER MULTINATIONAL EFFORTS IN DETECTING TRANSPLANT TOLERANCE

Detecting and/or inducing tolerance requires large multicenter networks that facilitate patient recruitment and ensure the necessary technical and scientific expertise. Several institutions are running projects with similar aims (Table 1). Immediately after the IoT study commenced, the RISET consortium was formed with similar aims: the development of reliable tests to predict tolerance, with a view to assess patients enrolled in RISET pilot clinical studies. Further, RISET aimed to stimulate debate regarding ethical aspects of tolerance induction protocols and establish educational programs in transplantation tolerance, not only for scientists and clinicians but also for patients and their families.[18] The RISET consortium is currently compiling the final report to summarize their activities. The first study to emerge in the Framework 7th research program from the EU commission is THE ONE study. This study led by Dr Geissler in Regensburg is going to test a number of cellular therapies to induce tolerance in kidney transplantation in several European centers.

The ITN is presently funding several clinical trials in solid organ transplantation, as a continuum of the biomarkers of tolerance study. One of these studies, ARTIST (An Observational Study to Assess the Prevalence of a Tolerance Signature in Renal Transplant Recipients),[13] aims to examine a large cohort of renal transplant recipients in order to prospectively determine the frequency and stability of the tolerance signature in patients maintained on IS treatment. The ITN has also made a worldwide call to locate operationally tolerant patients in an effort to create an additional Registry of Tolerant Kidney Transplant Recipients. Finally, the Gradual Withdrawal of Immunosuppression in Patients Receiving a Liver Transplant (AWISH)[17] study aims to investigate whether liver transplant recipients can be weaned from immunosuppressive drugs under medical supervision. Another active group within the tolerance biomarker field, based at Stanford University in the United States, is headed by Professor Minnie Sarwal. Their work focuses on candidate gene expression studies in both adult and pediatric liver and kidney transplant patients.[15] In collaboration with Professor Jean Paul Soullilou's team, the group described a transcriptional profile, specific for operational tolerant patients in 2007.[19] In Barcelona, the group headed by Dr Sanchez-Fueyo is focused on the characterization of biomarkers predictive of tolerance in the setting of liver transplantation. Their findings to date have led to the description of the first protocol in clinical transplantation, in which a tolerance signature has been used to monitor disease and inform decisions on drug withdrawal.[20]

METHODOLOGICAL CHALLENGES IN THE STUDY OF BIOMARKERS OF TOLERANCE

The major methodological challenge associated with the study of tolerance in kidney transplantation is the 'hidden' nature of the phenotype. The tolerance state is, as yet,

Table 1 | Examples of current international efforts towards conducting translational research and biomarker discovery in solid organ transplantation

Transplant tolerance-related clinical study	Organ	Lead institution	Country	Ref.
Genetic analysis and monitoring of biomarkers of immunological tolerance (GAMBIT)	Kidney	King's College London	United Kingdom	10
An observational study to assess the prevalence of a tolerance signature in renal transplant recipients (ARTIST)	Kidney	ITN	USA	13
ITN registry of tolerant kidney transplant recipients	Kidney	ITN	USA	14
Prediction and mechanisms of transplantation tolerance	Kidney	Stanford University (Sarwal lab)	USA	15
Effect of immunosuppression drug weaning on hepatitis C virus (HCV)-induced liver damage after liver transplantation	Liver	Hospital Clinic of Barcelona	Spain	16
Gradual withdrawal of immunosuppression in patients receiving a liver transplant (AWISH)	Liver	ITN	USA	17

Abbreviation: ITN, Immune Tolerance Network.

unknown. The patients voluntarily or directed are weaned off IS for a long period of time, and this is not accompanied by rejection; therefore, we assume are tolerant. It is important to point out that any molecular changes may have already occurred because of the tolerance state; however, it may also be partially due to the absence of immunosuppression. For this reason, biomarker studies of tolerance to date have been purely cross-sectional, where patients at all stages of the post-transplant period are allocated to defined clinical groups. As a consequence, it is unknown when the tolerant signature arises and how it develops from the time of transplant. For example, it is unknown whether the signature detected by these studies is detectable if the tolerant patients are still under IS. This raises the problem of the appropriate control comparison group for tolerant patients. Enrolling healthy volunteers can serve as a control for the absence of immunosuppressive drugs, whereas stable patients or chronic rejectors are the ideal clinical comparison groups. It is expected that long-term immune monitoring studies will shed light into this question, making use of a more appropriate longitudinal design, and providing the most valuable type of data for this purpose; that is, the molecular differences between tolerant and stable/chronic rejection patients before successful weaning. Unfortunately, longitudinal studies currently rely on the independent decision of the patients to withdraw their medication, which is a rare and unsafe event. The rarity of the event puts limits to the statistical power of tolerance studies, whereas the high risk limits experimental allocation of patients to a 'tolerant' condition, as it is not currently ethically appropriate to encourage a patient to undergo weaning. Another interesting, and probably more feasible, design is to detect and follow up patients who are partially weaned off immunosupression for clinical reasons. Comparing the molecular changes in patients who succeed the process with no rejection episodes and those who suffer rejection as a consequence would help to further disentangle the tolerance signature from the drug's signature.

To confront the problem of statistical power, most studies establish multicenter and often international collaborations, which necessarily introduce noise due to stratification effects in genetic variation, and differences across countries and centers in clinical protocols. A clear definition of the phenotypes, that is, Stable vs. Tolerant vs. Chronic Rejector, and a careful choice of inclusion and exclusion criteria can help minimize the noise in the data. Thus, although a predictive model is better validated when replicated in miscellaneous populations,[21] the original development should be carried out using a restrictive phenotype, for example, 'A tolerant patient is that who has ceased all medication for at least 1 year, has not suffered any episodes of acute rejection or displayed indications of deteriorating/chronic kidney dysfunction, and whose kidney function has changed $<15\%$ in the previous year' is a better candidate than 'A tolerant patient who has ceased all medication and has not suffered allograft failure'. Generalization to a broader definition of tolerance can be tested subsequently through external validation in an independent patient sample, and the test can be updated when necessary for its application in significantly different populations.[22] In addition, the effects of known confounders can be tested explicitly by adding clinical information to the multivariate predictor, for example, donor type, number of transplants, time since transplant, age of donor and recipient, human leukocyte antigen mismatch, and so on.

WHAT MAKES A GOOD BIOMARKER STUDY?

We will now consider the component parts of a well-designed biomarker study. In a recent review, Naesens and Sarwal[21] argued that 'omics' data are currently being used in clinical studies for two different purposes: predictive biomarker discovery, and elucidation of pathophysiological processes to identify therapeutic targets. The ultimate goals of biomarker studies are prediction and prognosis, and this needs to be recognized at all stages of the study, from design planning to data analysis and reporting (see glossary for a description of methodological terms in this section).

Design

Prospective studies are preferred over cross-sectional studies, and the selection of the time points of interest is a major decision that cannot be taken lightly. It is our experience that

involving a multidisciplinary team in the planning stages of the study is vital to reach an optimal design and ensure a smooth running of the study; for example, clinical team, translational research team, laboratory technicians, data manager, and statisticians. An ideal biomarker or set of biomarkers should show significant differences between tolerant and non-tolerant patients either in mean expression at a specified time point after transplant, or in the slope of expression estimated across a number of repeated measures.

Data analysis

Naesens and Sarwal[21] propose an 'integrative omics approach' for translational research. Although it is clear that 'cross-platform' multivariate predictors that incorporate different types of 'omics' data are the future, this imposes a serious statistical challenge. The issues associated with the processing and analysis of high-dimensional microarray data have been broadly discussed elsewhere.[23,24] In the context of biomarker studies, it is important to use a data-analysis protocol that places classification accuracy as the top priority. The feature selection method should ideally:

(1) use misclassification error or other measures of classification accuracy as criteria for feature selection (genes, molecules, and so on), rather than simply *P*-values, for example, predictive analysis of microarrays, significant analysis of microarrays, supervised principal components;
(2) consider the performance of the features in combination rather than individually in the selection process, for example, classification trees, principal components analysis, regularized regression;
(3) be sensitive to correlations and/or interactions between the features, for example, random forest or elastic net (see Hastie *et al.*[25] for an extensive review of data mining methods).

Subsequently, different classification algorithms can be used to build the final predictive model, and the most appropriate will depend on the data, the proficiency of the analysis team, and the complexity of interactions and correlations present in the selected features. Recently, the MicroArray Quality Control (MAQC) consortium[24] demonstrated that 'good modeling practices were more important than the choice of a particular algorithm' (p 834). Four key points are to be taken from this report: (1) maintain a reasonable ratio of sample size to classifier complexity; (2) use internal validation (cross-validation or bootstrapping) throughout feature selection and development of the classifier, not only at the end; (3) perform external validation in a completely independent sample; and (4) all the steps and decisions taken during the classifier-building procedure should be carefully documented and justified. This last point brings us to the importance of a full report of the analysis process, as well as the use of appropriate measures of performance, to facilitate replication by other research teams.

Reporting

The value of a biomarker depends on its ability to discriminate different groups of patients, and in this case tolerant vs stable and/or chronic rejectors, but this should not be the only outcome. In a recent publication, Steyerberg *et al.*[22] describe classic and novel ways to evaluate the predictive value of a new marker. According to these authors, it is essential to quantify the value added by the new marker to currently available, easier to obtain (clinical) variables. Furthermore, accuracy measures such as sensitivity and specificity, or area under the receiving operating curve, provide incomplete information regarding the performance of the new test. Calibration measures are needed to compare outcomes and predictions, especially in the external validation stage; measures related to reclassification provide detailed information regarding the gain of adding a new marker; more importantly, a good sensitivity is not necessarily translated into clinical usefulness, and decision-curve analysis can help to evaluate the value of a marker, taking into consideration the drawbacks and benefits of false positives and true positives, respectively.

Overall, reports of discovery of new biomarkers, or validation of previous ones, need to focus around a comprehensive description of accuracy of classification and clinical usefulness added to existing clinical measures. A detailed disclosure of the final predictive model, including values of coefficients, is desired to allow for a complete replication and validation by others. Copyrights should be protected by patent applications, rather than by lack of transparency in scientific reports.

THE FUTURE: TRANSLATION OF BIOMARKERS INTO THE CLINIC

Advances in our understanding of human immunological processes and developments in new therapeutic and diagnostic agents make the detection and/or induction of graft tolerance a real possibility in the near future. Many therapeutic agents with potential tolerogenic properties have been described, and some of them are currently undergoing clinical trials (reviewed in St Clair *et al.*[26]). However, the lack of well-defined biomarkers of tolerance represents the most significant barrier to the development of tolerance therapeutics. Without these tools, studies will miss an appropriate clinical endpoint to define the operational tolerance state. Long-term prospective studies could address the generation of tolerance but only in the context of IS withdrawal protocols. The only chance for a weaning study to be successful will be a carefully designed one, in which weaning is considered in the presence of increased surveillance and using validated biomarkers of both rejection and tolerance (Figure 1). Notwithstanding, this will imply in the case of kidney allografts a real risk of graft loss. The current efforts in establishing validated and stable biomarkers of tolerance in combination with validated and widely accepted biomarkers of rejection will allow such study to be ethically acceptable.

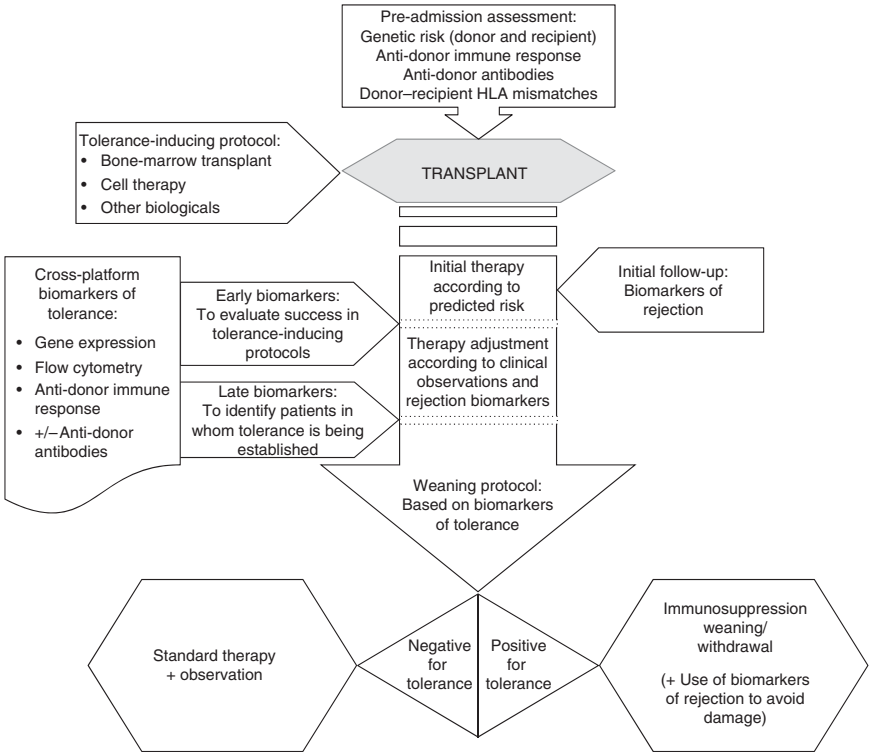

Figure 1 | Biomarker-led management of kidney transplants. Pretransplant assessment: risk is assigned according to anti-donor immune responses, anti-donor antibodies, human leukocyte antigen (HLA) mismatches, and genetic risk. Biomarkers of rejection will assess whether acute rejection will develop in the initial months post-transplantation, the time period depending on the depletion agent used. To maximize organ function, the treatment could be modified before tissue injury occurs. Two time points use biomarkers of tolerance: first, to establish the success of any peritransplant tolerance-inducing protocol. Later, to establish whether a tolerance state has been reached, the main aim would be to wean those patients off with positive biomarkers of tolerance. Throughout weaning, biomarkers of rejection could be used to stop the process before tissue injury is evident. Obviously, if a patient is negative for the biomarkers of tolerance, then the standard protocol of immunosuppression should be maintained for life.

The road to the discovery of biomarkers has been enormously boosted by the advances of new tools for high-throughput analysis, which enable us to interrogate the genome, epigenome, transcriptome, and proteome. Together with these advances, new methods for data analysis are also being developed to embrace this new multidimensional data in a holistic rather than reductionistic perspective. The discovery process implies the establishment of correlations between gold standard clinical parameters of disease and changes in biomarkers. To do so, there is a real need for a coordinated approach between international networks to create common biobanks associated with comprehensive clinical records. This is of particular importance in the process of biomarker validation, where large-scale prospective multicenter studies are required. New initiatives such as the Biomedical Research Centers[27] in the United Kingdom are of outstanding importance in bringing together both clinical and research expertise, which is an essential step to optimize the resources and funding that translational projects require.

CONCLUDING REMARKS

We aspire that validation of the biomarkers of tolerance will conclude in a clear and robust definition of the hidden phenotype of immune tolerance. This definition will have multiple benefits that will ultimately impact the management of kidney transplant recipients. We foresee a new paradigm in transplantation medicine, where patients will be stratified based on pretransplant-established risk factors such as human leukocyte antigen matching or cold ischemia times, and post-transplant monitoring of non-invasive biomarkers.[28] Together, the studies described here provide evidence that it might be possible to develop biomarkers capable of detecting operational tolerance in kidney transplantation. We expect that, in the near future, together with clinical and histological information, better characterized molecular and cellular markers will help us predict the outcome in kidney transplantation, guide personalized IS and weaning protocols, and even develop new therapeutic targets.

DISCLOSURE

MPH-F has filed biomarkers of tolerance as intellectual property (IP); however, this has not generated income or royalties to date.

ACKNOWLEDGMENTS

MPH-F, EP, and PS acknowledge financial support from the Department of Health via the National Institute for Health Research (NIHR) comprehensive Biomedical Research Centre award to Guy's & St Thomas' NHS Foundation Trust in partnership with King's College London and King's College Hospital NHS Foundation Trust.

MPH-F acknowledges financial support from the MRC (G0801537/ID: 88245) and Guy's & St Thomas's Charity (grant 080530). IR-M acknowledges financial support from the MRC. We thank Dr Mike Weale for his critical reading of the manuscript.

REFERENCES

1. Collaborative Transplant Study. 10 year survival. All transplants 1985–2009. 2011 (updated 01 February 2011, 06/03/2011), available from http://www.ctstransplant.org/.
2. Afzali B, Lechler RI, Hernandez-Fuentes MP. Allorecognition and the alloresponse: clinical implications. *Tissue Antigens* 2007; **69**: 545–556.
3. Baker RJ, Hernandez-Fuentes MP, Brookes PA *et al* . Loss of direct and maintenance of indirect alloresponses in renal allograft recipients: implications for the pathogenesis of chronic allograft nephropathy. *J Immunol* 2001; **167**: 7199–7206.
4. Hernandez-Fuentes MP, Lechler RI. Chronic graft loss. Immunological and non-immunological factors. *Contrib Nephrol* 2005; **146**: 54–64.
5. Sagoo P, Perucha E, Sawitzki B *et al.* Development of a cross-platform biomarker signature to detect renal transplant tolerance in humans. *J Clin Invest* 2010; **120**: 1848–1861.
6. Newell KA, Asare A, Kirk AD *et al.* Identification of a B cell signature associated with renal transplant tolerance in humans. *J Clin Invest* 2010; **120**: 1836–1847.
7. Porcheray F, Wong W, Saidman SL *et al.* B-cell immunity in the context of T-cell tolerance after combined kidney and bone marrow transplantation in humans. *Am J Transplant* 2009; **9**: 2126–2135.
8. Bouaziz JD, Yanaba K, Tedder TF. Regulatory B cells as inhibitors of immune responses and inflammation. *Immunol Rev* 2008; **224**: 201–214.
9. Kirk AD, Turgeon NA, Iwakoshi NN. B cells and transplantation tolerance. *Nat Rev Nephrol* 2010; **6**: 584–593.
10. Hernandez-Fuentes MP. Genetic analysis and monitoring of biomarkers of immunological tolerance. 2010, available from http://www.kcl.ac.uk/schools/medicine/research/diiid/depts/nephrology/indtol/gambit/.
11. Elinoff B, Nadler M, Mazariegos GV. International Solid Organ Transplant Tolerance Registry. 2010, available from http://www.transplant-tolerance.org/.
12. Le Texier L, Thebault P, Lavault A *et al.* Long-term allograft tolerance is characterized by the accumulation of B cells exhibiting an inhibited profile. *Am J Transplant* 2011; **11**: 429–438.
13. Newell K, Turka L, Chandraker A. An Observational Study to Assess the Prevalence of a Tolerance Signature in Renal Transplant Recipients (ARTIST). 2011, available from http://www.immunetolerance.org/studies/observational-study-assess-prevalence-tolerance-signature-renal-transplant-recipients-artist.
14. ITN. Immune tolerance network. Available from http://www.immunetolerance.org/.
15. Sarwal M. Stanford genomics and proteomics in organ transplantation. 2010, available from http://sarwal.stanford.edu/.
16. Sachez-Fueyo A. Effect of immunosuppression drug weaning on hepatitis C virus (HCV)-induced liver damage after liver transplantation. 2008, available from http://clinicaltrials.gov/ct2/show/NCT00668369.
17. Shaked A. Gradual Withdrawal of Immunosuppression in Patients Receiving a Liver Transplant (AWISH). 2011, available from http://www.immunetolerance.org/studies/gradual-withdrawal-immunosuppression-patients-receiving-a-liver-transplant-awish.
18. RISET-Consortium. Reprogramming the immune system for establishment of tolerance. 2007, available from http://www.risetfp6.org.
19. Brouard S, Mansfield E, Braud C *et al.* Identification of a peripheral blood transcriptional biomarker panel associated with operational renal allograft tolerance. *Proc Natl Acad Sci USA* 2007; **104**: 15448–15453.
20. Martinez-Llordella M, Lozano JJ, Puig-Pey I *et al.* Using transcriptional profiling to develop a diagnostic test of operational tolerance in liver transplant recipients. *J Clin Invest* 2008; **118**: 2845–2857.
21. Naesens M, Sarwal MM. Molecular diagnostics in transplantation. *Nat Rev Nephrol* 2010; **6**: 614–628.
22. Steyerberg EW, Vickers AJ, Cook NR *et al.* Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology* 2010; **21**: 128–138.
23. Allison DB, Cui X, Page GP *et al.* Microarray data analysis: from disarray to consolidation and consensus. *Nat Rev Genet* 2006; **7**: 55–65.
24. Shi L, Campbell G, Jones WD *et al.* The MicroArray Quality Control (MAQC)-II study of common practices for the development and validation of microarray-based predictive models. *Nat Biotechnol* 2010; **28**: 827–838.
25. Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning: Data Mining, Inference, And Prediction*, 2nd edn. Springer: New York, 2009.
26. St Clair EW, Turka LA, Saxon A *et al.* New reagents on the horizon for immune tolerance. *Annu Rev Med* 2007; **58**: 329–346.
27. NIHR.UK. Biomedical research centres. 2007, available from http://www.nihr.ac.uk/infrastructure/Pages/infrastructure_biomedical_research_centres.aspx.
28. Goldman M, Wood K. Transplantation research: will we ever reach the holy grail? *Transplantation* 2009; **87**(9 Suppl): S99–100.
29. Steyerberg EW. *Clinical Prediction Models: A Practical Approach To Development, Validation and Updating*. Springer: New York, 2009.

GLOSSARY OF METHODOLOGICAL TERMS

Confounder: a variable (measured or not) other than the predictor variables (biomarkers) potentially associated with the outcome variable. Failing to consider it can result in a biased estimate of the association between biomarker and outcome.

Feature selection: the process of selecting a subset of predictor variables (biomarkers) from a highly dimensional set (e.g. microarray) based on their association with the outcome variable (e.g., clinical groups).

Classifier/predictive model: a statistical model or algorithm used to predict the outcome variable (or probability of outcome) from the observed values in the predictor variable/s (biomarker/s).

Discrimination: the ability of a classifier to differentiate between individuals with and without the outcome. Measures of discriminative performance are the concordance (c) statistic for quantitative outcomes, and the area under the receiver operating characteristic curve (AUC of ROC).

ROC curve: a plot of sensitivity (true positive rate) against 1-specificity (false positive rate) for consecutive cutoffs for the probability of an outcome predicted by the classifier.

Internal validation: determining the reproducibility of the predictive model in the development (training) sample via:
Cross-validation: dividing the sample in random groups, leaving one group out to develop the classifier, and testing its performance in the remaining group. This procedure is repeated until all groups serve as test sample, and the whole process repeated a number of times to obtain a stable average estimate of performance.
Bootstrapping: drawing samples with replacement from the original sample a sufficient number of times to obtain a stable estimate of performance.

External validation: the process of testing the generalizability of the performance of a predictive (biomarker) model in an independent sample, from a 'plausibly related' population.

Calibration: the agreement between observed outcomes and predictions. Calibration-in-the-large is the difference between average predicted outcome and average outcome in an independent sample. The calibration slope is related to the strength of the predictors' effects, and how those replicate in an independent sample.

Clinical usefulness: evaluation of the performance of a predictive model taking into account the relative importance of false-positive and false-negative decisions in a clinical decision context. Decision curve analysis is the application of the ROC curve method to choose an optimal cutoff that considers a weighted error rate.

Note: see Hastie *et al*. [25] and Newell *et al*.[6] for a more detailed description of these terms.